The Effects of Narrated Animation and Still Images with Text in Science Learning

Lianru Zhang
Graduation Thesis, January 2015
Media Technology MSc program, Leiden University
Supervisor: Maarten Lamers and Dan North
lianrutuantuan@hotmail.com

Abstract—*Present day multimedia offers new approaches to science learning. Some scientists discover animation is a better component to combine with a verbal medium (such as text, narration) to help people understand scientific knowledge in multimedia format (e.g. Yang, Andre, & Greenbowe, 2003), compared to a combination of still images and verbal form. Others argue that animation does not perform better than static pictures in learning science (e.g. Mayer et al., 2005).This study investigates a difference between narrated animation (NA) and static pictures with text (SP) in the effect of science learning. It assumes that SP may perform similarly to NA, if the presented information is considerably equivalent and the time duration and procedure between both are strictly controlled. An experiment was conducted to test this assumption: Subjects were randomly assigned to learn how coffee impacts on human beings' brain and some related knowledge via computer-based NA or computer-based SP. The final results showed that the difference between NA group and SP group was statistically insignificant (P > 0.05). It indicated that there was no statistically significant different effect of NA and SP in science learning, which was in consistent with the hypothesis in this study.*

Key Words— **Science education; multimedia; animation; static picture; computer-based.**

I. INTRODUCTION

With the development of technology, multimedia has offered new approaches to science learning. Multimedia is the use of text, graphics, pictures, video, and sound to present information (Najjar, 1996). Under three conditions, multimedia could become an efficient tool for science learning (Najjar, 1996): the medium follows the dual coding theory (Paivio, 1986); when information is presented with one medium supporting another one; learners have low prior knowledge or aptitude when they learn a specific field. Dual coding (Paivio, 1986) is a theory which indicates that people have verbal and non-verbal systems to process information. These two systems are independent but interconnected.

Learning involves words and pictures materials. They can be delivered using any medium (Mayer, 2005).To explore the relationship between words and pictures in multimedia learning, Mayer (1994) proposed a cognitive theory of multimedia learning (CTML) based on early studies from Paivio (1986), Baddeley (1998), Sweller (1999), Wittrock (1974).

CTML consists of three assumptions: dual channels, limited capacity of each channel, and active processing (Mayer, 2005). Dual channels mean an individual has two different channels to process information: a visual/pictorial channel and an auditory/verbal channel. And each channel processes limited amount of information at one time due to the limitation of working memory. When two systems process one source of information simultaneously, they complement each other and the learner receives more information than using one single channel. Active processing assumes that when learner involves his cognitive processing, it enhances his deep learning. The active cognitive processing includes attention, information organization, and integrating incoming information with prior knowledge (Mayer, 2005).

A series of studies conducted by Mayer and his colleagues (Mayer et al., 1996, Moreno &Mayer, 2000, and Mayer et al., 2001, Mayer &Moreno, 2002) do not only examine CTML, but also show some interesting findings in multimedia learning. First of all, words and pictures make people learn better than words alone. Moreover, animation with narration (in a spoken form) is the better combination to facilitate science learning, compared to other forms of mediums (on-screen text only, narration alone, animation with on-screen text, narrated animation with on-screen text.) In addition, some principles of designing animation with narration are summarized, such as animation should be presented with narration together and simultaneously, text need to be closed with picture in animation, extraneous words and sounds (including music) should not be adding in narrated animation, etc.

Furthermore, some studies discover animation is a better component to combine with a verbal medium (such as text, narration) to help people understand scientific knowledge in multimedia format (e.g. Large et al., 1994, Large et al., 1996, Yang, Andre, & Greenbowe, 2003, etc.), compared to a combination of still images and verbal form. Höffler and Leutner (2007) report an effect size of d= 0.37 after they analyze 26 studies regarding animation and static pictures from 1978 to 2003, which means an overall superiority of animations over static pictures. They also point out representational animation (animation explicitly represents the learning content) is more efficient than the decorational type (Animation doesn't directly display the learning content). Others argue that animation does not perform better than static pictures in learning science (e.g. Mayer et al., 2005, Höffler et al., 2010). Although the results from some studies (e.g. Baek & Layne, 1988, Large et al., 1996, Rieber,1990, 1991a, b) show that students learn scientific knowledge better via animation than still images, Tversky et al. (2002) doubt the validity of those positive results which support animation. They point out that the material is not presented equivalent between animated and static graphics in those studies (either more information displays in animation or the same information presents better in animation than still images). Those studies have contradictory views between animation and static images in science learning.

The definition of animation is a simulated motion image portraying movement of drawn (or simulated) objects, whereas an illustration or a still image is a static picture of drawn (or simulated) objects (Mayer &Moreno, 2002). Simulated objects are artificially created by some simulation methods, such as drawing (Mayer &Moreno, 2002). Animation and still image both are pictorial representations, so both of them can abstract the essential conceptual information (Tversky et al., 2002), especially they are designed in a low level of realism (Dwyer, 1978). Additionally, both formats can provide a complement to verbal information according to CTML.

However, due to the dynamic format, animation can visualize a process or a procedure, it also displays the information in a transient way (the frames of pictures changes over time), and those features conduce to less cognitive load and effort to construct a mental model (Höffler &Leutner, 2007). Moreover, these vivid visual representations could trigger learner's motivation and interests in learning (e.g., Lepper & Malone, 1987, Barak, Ashkar & Dori, 2011). In contrast, still images as a non-dynamic format allows learners to control the pace and order of incoming information; it emphasizes the key information and reduces irrelevant information, compared to animation (Mayer et al., 2005).

Since still images and animation both hold their own strengths in science learning, it is hard to tell which format is more effective in the process of understanding of scientific concepts. The informational equivalence is existing issue between static pictures and animation, when researchers conduct their experiments to compare both formats in their studies (e.g. Tversky et al., 2002). Although it is difficult to demonstrate information between two formats absolutely equal, it should be considerably equivalent as much as possible. Therefore, methodology needs to be carefully designed, when it compares animation with still pictures.

Static pictures with text and narrated animation both are commonly used in self-study in practice (Mayer et al., 2005), so my current study focuses on comparing these two formats in the effect of science learning.

Therefore, my hypothesis is that static pictures with text may perform similarly to narrated animation, if the presented information is considerably equivalent and the time duration and procedure between both are strictly controlled.

.

II. Relevant Studies

Yang et al. (2003) examine an understanding of targeted chemistry concepts between college students after they either watch animation with instructor- guided narration or still diagrams with instructor- guided narration during the lecture in the classroom. In their experiment, animation and still diagrams illustrate the same content of the principles of how batteries produce electricity. The results suggest that animation lead students understanding this electrochemistry concept better than still diagrams overall. However, instructor replays some animations several times at class due to the requirement of students in animation group. This may cause the outperformance of animation group. And another problem is that the information present between animation and still diagrams in their study is not considerably equivalent. Yang et al. (2003, p345) claim themselves "The computer animations may have had direct cognitive effects because they provided information not available to the SDG (still diagrams group)."

Mayer et al. (2005) conduct four experiments in four different scientific topics (how lighten works, the mechanism of a toilet tank, the process of ocean waves, the mechanism of a car's brake system) to compare which format leads college students learn better in those scientific topics between paper-based annotated illustration with text and computer-based narrated animation. The outcome shows that students in narrated animation group did not significantly get higher grade on posttests than students from the group of annotated illustration with text. However, there are two issues in their experiments. First, two formats deliver in two different media: the static pictures with text is printed on papers, whereas the narrated animation is computer- based. Second, in experiment 1 and 3, subjects in animation group only view the animation once, however, during the same amount of time, some subjects in annotated illustration group could repeat reading their papers more than once. These two unequal treatments may lead the group of still images with text score higher than animation group.

Höffler et al. (2010) test the effect of visual cognitive style between static pictures with text and animations with text. Höffler et al. (2010) test the effect of visual cognitive style when high-school students learn the main reactions in photosynthesis via either static pictures with text or animations with text. Their findings show that highly developed visualizers have better understanding of the knowledge through static pictures than animations. In contrast, less developed visualizers make no difference between these two learning formats. However, compared with still images, there is no movement indicators (e.g. 'arrows') are used in animation. This implies the imparity of presented information. That means subjects in the animation group don't only need to pay attention to the change of motions, they also have to read annotations in the pictures at the same time. This requires that participants from animation group make more cognitive efforts than that from still images group. In addition, the level of visual cognitive style is rated by using self-evaluated form between subjects. Although it does provide a standard to measure visual cognitive style, the accuracy rating of this self-evaluation is skeptical.

Höffler and Leutner (2011) measure that people receive the scientific knowledge of how surfactants works during the washing process under different learning environments – computer-based narrated animation and computer-based narrated still images. In their study, two experiments are conducted by the same topic but different type of subjects. In the first experiment, all subjects are graduated students, whereas high-school students participate in the second experiment. The learning outcomes claim that generally participants acquit knowledge via static pictures as same as via animation in their experiments, although subjects with highly spatial ability benefit from both animation and static pictures learning environment, whereas participators with low spatial ability poorly profit from the learning environment of static pictures. Nevertheless, Höffler and Leutner allow subjects to watch animation or pictures as long as they want during their experiments. If there is no time limitation in learning process, some students could spend more time to digest information deeper than others, this could impact the final results.

In addition, although some studies are concerned with individual difference (such as spatial ability) in multimedia learning (e.g. Yang, Andre, & Greenbowe, 2003, Höffler & Leutner, 2011), the results they get are contradictory. Because individual difference is very complex, it involves more variable. Therefore, my current study only focuses on the benefits between narrated animation and static pictures with text for people in general. Without considering individual difference, it may impact on the final findings.

III. METHODOLOGY

A. Subjects Selections

The participants of this study were enrolled through two ways: 1. from my own circle of friends; 2.online recruitment on Facebook. All of participants must speak English, which included native English speakers and non-native English speakers. The age of participant was set at 18 years old and above. The amount of participants was above 40 people in total. After the experiment, each of subjects were given €2, a bottle of beer or a box of chocolate biscuits as compensation or voluntarily contribute into this study, according to their options.

B. Conditions

In this study, subjects were presented a scientific topic about how coffee impacts on human beings' brain and some related knowledge via computer-based narrated animation or computer-based still images with text during the experiment.

The narrated animation was selected from one video from the science channel on YouTube called "as soon as possible science (Asap SCIENCE)". The name of the animation was "Your Brain on Coffee". The narration of this animation was in English. This animation was in MP4 version and lasted two minutes forty seconds in total. It could be viewed by any player on a computer.

There were three reasons to choose this animation as the learning material in this study. First of all, Asap SCIENCE is a professional group that produces online animations to explain specific scientific knowledge for the general public and it is very popular. This channel on YouTube has 3,206,546 subscriptions until December 2nd, 2014. Importantly, except for the music background, this animation follows the basic design principles in multimedia learning (Mayer &Moreno, 2002), such as annotation. In addition, the images in this animation are designed in a simplified style, which is in line with the images with less details allowing for more effectiveness to express the essential conceptual information (Dwyer, 1978).

It was also important for the narration in the animation to be the same as the text in the static pictures presentation and this was the case in this experiment that the text was a literal transcription from the video. In addition, the text was placed below the pictures, which followed the design principle of text needs to be close to the images (Mayer &Moreno, 2002). The still images with text contained 26 images and 32 sentences (537 words in total) and it supplied in a PDF file (Figure 1). The subject could view the PDF file only by scrolling up and down, which was similar with the function of progress bar in media player. This design function prevented the viewer from jumping between pictures, which might generate unequal treatment between animation viewer and viewer of static pictures. Therefore, the presented information between these two formats (narrated animation and static pictures with text) was considerably equivalent.

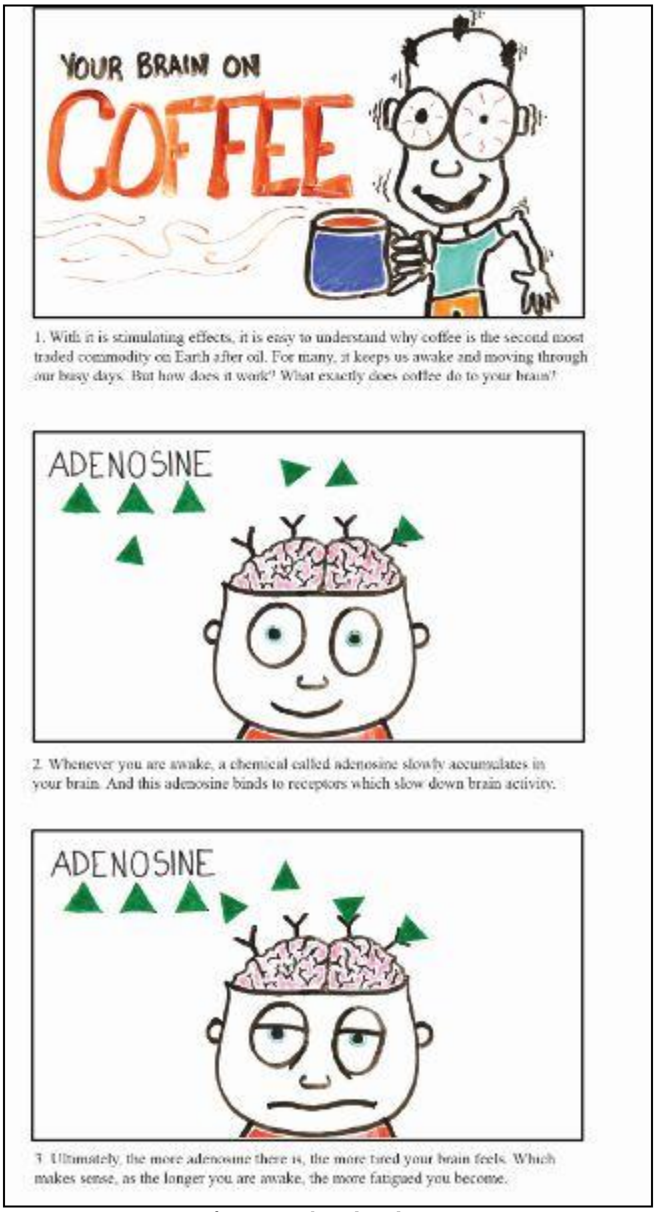

Fig .1. Section of computer-based reading version on Your Brain on Coffee

C. Procedure

Participants were randomly assigned to either narrated animation group (NAG) or static picture group (SPG). Due to the difficulty of organizing the experiment in a laboratory, the experiment was done by subjects using their own computers at home or somewhere else in two different methods:

- *Method 1(M1)*: Before the experiment, all of subjects received a digital folder via their Email. The folder consisted of learning material (MP4 animation file or FPD reading material), a procedure explanation PDF file (which included explaining the order of the experiment, matters need attention, etc.), a pre-test WORD file to collect demographics (which included some personal information, such as age, nationality, education background, etc.) , a WORD version of post-test which contained a short questionnaire and a screen recorder software (Free Camstudio was for windows users, QuickTime Player was for Mac users) to record the subject's whole process while doing the experiment on their computers.

  During the experiment, the viewer was first required to read the instruction guide that explained the procedure of the experiment. The second step was that subject installed the screen recorder software and then clicked start button to record the whole process. And then they filled out the pre-test less than 6 minutes. After that, they either watched animation or read material accordingly. The maximum time for this section was 7 minutes. Subjects felt free to repeat watching or reading any part of the material as many times as they wanted during this time. They also could go on to the next step before they reach the maximum time. The final step was to finish a post-test in 10 minutes and then a simple questionnaire. After they finished all the tasks, they clicked stop button on screen recorder software and then saved the video in their computers. After that, subjects could send their pre-tests, post-tests and recording video back to the experimenter via Email.
- *Method 2(M2)*: The second form was an online test. The content of this test was the same as the first one. The only two differences were that participant did the whole experiment in website called Surveygizmo; and in the website, each part of the test had a timer to count the required time down for participant during the test. When time was up in each part of the test, it automatically jumped into next step. This prevented participant from spending more time in the test than he was required in each part of the whole experiment.

Generally, the estimated maximum time duration of experiment was for each subject was around 23 minutes in total. It contained 6 minutes of pre-test, 7 minutes of viewing material and 10 minutes of post-test.

D. Measurement

In this experiment's design, there is no need for a control group (CG). The pre-test during the experiment substitutes for the CG need. In turn, participants in NAG and SPG do only one test after studying the material while the subjects in a CG will take the same test as participants in NAG and SPG do without learning anything before. However, doing a pre-test in NAG and SPG has the same effect as only doing one test between a CG, a NAG and a SPG. Thus the purpose of adding a CG or doing a pre-test between NAG and SPG are equivalent: to evaluate the average prior knowledge of caffeine between people without learning the study material. Furthermore, it is difficult to find large quantities of subjects to join this study in a short time, and a CG also demands a lot of participants.

Prior knowledge has an important influence on science learning. Najjar (1996) stresses that people with lower prior knowledge could learn scientific knowledge better via multimedia tools. Therefore, to identify participants' prior knowledge about coffee, a simple pre-test was administered. The pre-test (contained ten points in total) consisted of one multiple choice question (scored two points) and two open questions (scored four points for each).The content of pre-test was about energy drinks (Appendix 1), because energy drinks and coffee had similar features. If the content of pretest had a direct relation to the subjects previous knowledge, such as mentioned about word 'caffeine', it would lead subjects focus on certain points of the information when they learn material, that could impact on the validity of the posttest.

Retention and transfer processes are two main aspects of weighting the understanding of knowledge (Mayer, 2005). Retention means remembering the information from the presented learning material. This study focuses on testing a short- term retention due to the time limitation. The long term retention needs to be considered in the future study. And transfer here includes using the learned information to solve new problem (Mayer, 2005) and using the learned information to explain certain phenomena. So to measure the learning results of how coffee impacts on human beings, a post-test was conducted after subject views learning information. Ten questions were constructed in this test, which consisted of six retention questions and four transfer questions (Appendix 2). The form of questions were multiple choice questions format, open and fill-in questions. Each of the questions scored one point, ten points in total.

All the items from pre-test and post-test were consulted with a micro-biology master student at Leiden University. In addition, a simple questionnaire was asked to fill out after the post-test (Appendix 3). It was used to check whether the style of images would impact on subjects' learning effects and the difficulty level of the topic. It also tracked the actual time that participants spend on the learning material and the post-test, in case the time they used was beyond the requirement of the maximum number or if they had any problem installing the screen recorder. If subjects used more than the required maximum time, the test was eliminated.

A t- test is a statistic method to test the null hypothesis whether is supported between the means of two small samples, when the population standard deviation is unknown (Fisher Box, 1987). A t- test includes an unpaired t-test (or called a student's t –test) and a paired t–test. An unpaired t –test can be used to look at a difference between two sets of data from two independent, random groups (Altman, 1991; Armitage and Berry, 1994). A paired t-test is to compare one set of sample with another related set of sample under the matched conditions or units in one group have been tested two times (David and Gunnink, 1997). Therefore, after collecting data from the experiment, an unpaired t-test was used to analyze the data between NAG and SPG, and a paired t-test was applied to evaluate data in a small group. In addition, my hypnosis does not assume the effect between two data would go into a specific direction, so two tailed test would be added in the t-test to examine the left and right extreme end side of the distribution curve of the data after the experiment as well.

IV. RESULTS

The data was collected from December 12th 2014 to January 8th 2015. In NAG, a total of 23 participants completed the test, which contained 10 participants accomplished the experiment via M1 and 13 subjects finished the test by M2. However, 2 people exceeded the required time during the experiment via M1 and one person refilled her answers several times on Surveygizmo during the experiment due to the poor Internet connection at that moment. Therefore, 3 of participants were not qualified in NAG. In fact, the final NAG consisted of 20 adults aged 21-33 years old (mean age = 26.2 years old), which included 8 female and 12 male. Among those participants in NAG, 5 of them were employed and others were students. Except for one person, all subjects in NAG were not native English speakers. Through their self-evaluation, 10 people thought their English reached in an advanced level, 9 participants were in an intermediate level and only one person was beginner. (**Table 1**)

Table 1. Number of participants, sex, occupation, mother language and English in NAG.

<table>
<tr><th colspan="2">NAG</th></tr>
<tr><td rowspan="2">N total = 20
(Mean age = 26.2)</td><td>N female= 8;
N male =12</td></tr>
<tr><td>N student= 15;
N employed =5</td></tr>
<tr><td colspan="2">Mother Language</td></tr>
<tr><td colspan="2">N Chinese = 10; N Dutch = 2; N English=1; N Filipino = 1; N French =1; N Greek=1;N German=1; N Italian =1;
N Romanian =1; N Turkish =1</td></tr>
<tr><td colspan="2">English Level*</td></tr>
<tr><td colspan="2">Na =10 ; Ni = 9; Ne =1</td></tr>
</table>

*a= advanced; i= intermediate; e= elementary.

There were 32 participants in SPG joined experiment: 19 adults completed the test via M1 and the rest of them did the test via M2. But one subject misunderstood an order of the test via M1 during the experiment, he first did the post- test and then pre-test; due to the poor Internet connection, one participant retook the test several times on Surveygizmo during the experiment. Therefore, the final SPG has 30 participants (Mean age = 27.1 years old), with a range from 18 to 61 years old. There were 19 male and 11 female. Only one participant's mother language was English in SPG. 15 participants had an advanced level in English, 14 subjects in an intermediate level and one person in an elementary level via their self-evaluation. (**Table 2**)

Table 2. Number of participants, sex, occupation, mother language and English in SPG.

SPG	
N total = 30 (Mean age = 27.1)	N female= 11; N male =19
	N student= 21; N employed =8; N other=1
Mother Language	
N Albanian =1; N Amharic =1; N Arabic=4 ; N Chinese = 10; N Dutch = 11;N English=1; N German=1; N Spanish =1	
English Level*	
Na =15; Ni = 14; Ne =1	

*a= advanced; i= intermediate; e= elementary.

Therefore, there were 50 valid results in total. In NAG, the mean of pre-test was 2 and that of post-test was 4.2; in SPG, average score of the pre-test was 2.7 and that of post-test was 4.9.(**Table 3**)

Table 3. Means of score, each questions' average grade in pre-test and post-test between NAG and SPG

Pre-test		
Question No.	NAG	SPG
1	0.8	1.1
2	0.4	0.6
3	0.8	1.1
Sum of means	**2**	**2.8**
Post-Test		
1	0.6	0.7
2	0.7	0.7
3	0.5	0.6
4	0.4	0.5
5	0.3	0.3
6	0.4	0.6
7	0.4	0.3
8	0.3	0.3
9	0.2	0.4
10	0.4	0.5
Sum of means	**4.2**	**4.9**

In addition, most participants liked the content of the learning material. Except one person in NAG and four adults in SPG indicated that the style of images did not impact them on absorbing information, 19 participants in NAG and 22 participants liked the style of learning material, whereas only 3 people in SPG stated that they did not like the style of learning material. And one subject in SPG argued that the style of images can be improved. In addition, most participants (11 people in NAG and 17 people in SPG) felt that the content of material was not difficult but also not easy for them to learn, while 7 people in NAG and 11 people in SPG said it was easy for them to learn the material. In both groups only one participant thought that the learning content was difficult, whereas one subject stated it was very easy to learn. (**Table 4**)

Table 4. The appreciation of style and difficulty level of learning material in NAG and SPG.

The Appreciation of Style of the Learning Material		
	NAG (Number of subjects)	**SPG** (Number of subjects)
L*	19	22
M*	1	4
D1*	0	3
O*	0	1
Difficulty Level of the Learning Material		
VE*	1	1
E*	7	11
NDNE*	11	17
D2*	1	1
VD*	0	0

*L= like; M= it does not matter for me, as long as it display the content clearly; D1=dislike; O= other; VE= Very Easy; E=Easy; NDNE= Not difficult but not easy; D2=Difficult; VD =very difficult.

A. Pre-test between NAG and SPG

Since two set of data of pre-test between NAG and SPG were independent and identically distribute, a two tailed unpaired student's t-test was used to measure a difference in means between two groups. Due to unequal variance, the result revealed that a difference in means of the pre-test between the two groups (2,8-2=0,8) could have been caused by chance (P = 0.22). In other words, the prior knowledge that participants had between NAD and SPG was no significantly different.

B. Post-test between NAG and SPG

Since two set of data of post-test between NAG and SPG were independent and identically distribute, a two tailed unpaired student's t-test was selected to evaluate a difference in means between two groups. Due to unequal variance, the outcome showed that a difference in means of the post- test results between NAG and SPG (4,9-4,2=0.7) was statistically possible (and even likely) to have been caused by chance(P = 0.24). It meant that the difference between two groups was"statistically insignificant", because P > 0.05.

C. Compare pre-test with post-test in each group

Technically, a two tailed paired t-test can be used to compare a difference between means of pre-test and post-test in NAG and SPG separately to examine whether participants significantly improve their knowledge about how caffiene effects the human brain and its related knowledge after their learning than before. In fact, this study is not going to compare the data between pre-test and post-test via a two tailed paired t-test in each group, because the contents between pre-test and post-test were not same. In other words, the two sets of data from pre-test to post-test in each group were not exactly similar units, so they were not really comparable. However, a difference between a pre-test and a post- test still exists in NAG (4,2 – 2 = 2,2)and SPG (4,9 – 2,8 = 2,1) separately.

Overall, the results from the experiment displayed that there was no statistically significant different effect of narrated animation and static pictures with text in science learning, which was in consistent with the hypothesis in this study.

V. DISCUSSION & CONCLUSION

Is a dynamic format (involving computer-based narrated animation) in multimedia more helpful in learning science than a non-dynamic format (involving computer-based static pictures with text), or the other way around? My hypothesis predicts that the NAG and the SPG will perform similarly. In this study, participants who learned from computer-based narrated animation did not score significantly better on post-test than subjects who studied from computer-based static pictures with text, and vice versa. This result is in line with prior studies (Höffler and Leutner, 2011, Tversky et al., 2002).

Why did participants in NAG get similar results as subjects in SPG on learning how caffeine effects the human brain, and other related knowledge? The explanation is that both animation and static pictures belong to the pictorial format (or visual format), although they have several different features (which are mentioned in the introduction section); both narration and text represent the verbal system, so narrated animation and static pictures with text are both a combination of verbal and non-verbal format; the information is showed simultaneously via both formats during the experiment. These features abide by a cognitive theory of multimedia learning (Mayer, 2005) and dual coding theory (Paivio, 1968), so there is no significant difference between narrated animation and static pictures with text. In practice, this finding implies that content regarding scientific topics can be interpreted via either animation with narration or static pictures with text on CD-ROMs and websites as long as the content itself is well-structured.

Furthermore, this study found that in average participants in NAG and SPG had a low prior knowledge of how caffeine in coffee works on the human brain and body, but after learning, subjects in both groups scored higher than before. This could indicate that multimedia (either narrated animation or still images with text) can efficiently assist people with low prior knowledge scientific information. This outcome would reconfirm the view from former research (e.g. Najjar, 1996, Mayer and Moreno, 2002.): multimedia (when information is present with one medium supporting another one) is an efficient tool for people who have a low prior knowledge to study of scientific topics.

An alternative competing explanation could be that most participants might know how caffeine in coffee works in the human brain and body before they learned the material, but due to lack of knowledge about energy drinks, they did not link caffeine with energy drinks when they were doing their pre-tests. If most participants already had a lot of knowledge about caffeine in coffee, most of them would get a higher score in post-test. However, in post-test, only six out of fifty participants got a grade above 7 points, no subject scored higher than 8.5 points.

Except investigating the effects of narrated animation and static pictures with text, two interesting observations were made in this study. The first one is that English level plays an important role in science learning (Chapman, 2000). People with lower level of English may have difficulties in science learning: In this study, two participants with elementary English level scored poorly (below 2 points) in post-test. However, having an advanced English level does not mean that people learn better in scientific topics: two participants with a higher level of English scored lower than 2 points during post-test.

Secondly, 7 out of 18 (around 40%) participants who believed the material was easy to learn did not get their grades higher than 5 points in post-test, this means that the result from people who self-assess their performance is not really accurate compared to the outcomes of their actual performance (Tousignant and DesMarchais, 2002). Therefore, self-assessment should be used carefully when researchers or scientists design their experiments.

Although many studies (e.g. Yang, Andre, & Greenbowe, 2003, Mayer et al., 2005) hold different views on the effects of animated media and static media, a vital contribution to this experiment is that it is an empirical demonstration that generally, a static format combined still images with text is not superior to a dynamic system contained narrated animation in respect of learning results, and vice versa. In other words, none of these methods (narrated animation and static pictures with text) has advantages other.

The limitation and weakness of this experiment is that this study had its own time limitation, so considering the difficulty of finding a large number of participants in a short time, this experiment was not conducted in a lab. So a possibility might be that: when participants took the test via M1, some of them might use more than the required time during the experiment, because most subjects used self-reported method to write down how much time they spent on learning section and post-test section rather than recording the whole experiment via recording software. If this situation takes place, it could impact on the final outcome. Therefore, a laboratorial experiment needs to be conducted in future. Furthermore, this study needs to be replicated in different settings and featuring different scientific topics, with a large number of participants.

In addition, normally, setting up a scientific hypothesis is for comparing a difference between samples via a P value. If P value ≤ 0.05, it means that there is a significant difference between samples, otherwise there is no a significant difference between samples. In principle, a t-test is not designed to evaluate a similarity between data. So if there is not a significant difference between two sets of data, it does not mean that there is a significant similarity between two sets of data. However, according to the results from the experiment in this study, there is no reason not to believe that a similar effect of narrated animation and static pictures with text does not exist.

In conclusion, the findings of this study show that narrated animation and static pictures with text both proved to be a significant benefit for people with low prior knowledge who would like to learn more knowledge about science, no matter which format people choose to study.

ACKNOWLEDGMENT

I'd like to thanks to Maarten Lamers and Dan North for guiding my whole graduation study. And thanks a lot to Mark Takefman for revising my English grammar. Thanks to Mohamed El-Moursi, Jasper scheffel, Peiyao Liu, Zheren Zhang and Zhiwei Yang who helped me during my research and all the friends I know or I don't know who participated in my experiment. And I am especially grateful the support from my lovely parents during my whole study.

APPENDIX 1

1. Do you know what the main component in energy drinks (such as Red bull) is? (2 points: caffeine or other elements such as sugar, etc.)
A. Yes, I know. It is _________ B. No, I don't know.

2. Can you shortly explain why the energy drinks keep people being energetic? (4 points: explain the process of how caffeine makes people awake or the principle of how taurine supplies energy to human body.)

3. Can you briefly summarize the information you know about energy drinks? (4 points: mention 4 features of energy drink, such as mix with alcohol makes people awake, even they already get drunk.)

APPENDIX 2

1. What are the features of Adenosine? (1 point)
A. It is a chemical which increases sleepiness.
B. It is a chemical which slows down the activity of people's brain.
C. The dosage of adenosine decreases, when people sleep.
D. The more adenosines, people feel more tired.
E. Adenosines activate the brain activity of human beings.
F. More adenosines create more receptors.
(Each correct answer gets 0.25 point. Answers are ABCDE.)

2. What are the similarities between adenosine and caffeine? (1 point)
A. They are similar in structure.
B. They are able to bind with receptors.
C. They both make people sleepy.
D. They have no similarities.
F. They both can be produced by the human body.
(Each correct answer gets 0.5 point. Answers are AB.)

3. Except increasing wakefulness, what are some effects when people drink coffee? (1 point)
A. Increase the blood pressure
B. Increase the heart bates
C. Breathe more easily
D. Improve people's positive feeling
E. Stimulate Adrenaline
F. Increase active level of the muscle
(Each correct answer gets 0.2 point. Answers are ABCDE.)

4. What makes people moderately addicted to coffee? (1 point)
A. Dopamine stimulation
B. Adrenaline stimulation
C. decreasing adrenaline level
D. decreasing Dopamine level
(Each correct answer gets 0.5 point. Answers are AD.)

5. Is it possible for adults to die due to drink too much coffee? (1 point)
A. Yes.
B. No.
C. Yes, a person can die if he drinks large quantities of coffee at once.
D. No, people cannot die after drinking large quantities of coffee over a long period of time.
(Each correct answer gets 0.5 point. Answers are BD.)

6. Explain why coffee makes people awake? (1 point)
(If participant writes down one part of the answer below gets 0.5 point. Answer is "Caffeine and Adenosine are similar in structure, so caffeine replaces adenosine into receptors. Due to caffeine won't make people feeling tired, so it keep people awake.")

7. Explain why it is difficult for people to stop drinking coffee after they drink it in a long time? (or: How come people who drink coffee on a regular basis experience withdrawal symptoms such as increased tiredness after quitting from drinking coffee) (1 point)
(If participant writes down one part of the answer below gets 0.5 point. Answer is" With long term use of caffeine, the brain responds by creating more adenosine receptors- which means more caffeine is required to elicit the same response. It also means that when a person tries to quit drinking coffee or miss his daily intake, he might experience some withdrawal symptoms and feel more tired than he would have before he ever drank coffee.")

8. Why does coffee induce feelings of happiness in people? (1 point)
(If participant writes down one part of the answer below gets 0.5 point. Answer is "The intake of caffeine results in the release of adrenaline, which in term reduces the absorption of dopamine, which results in a feeling of happiness.")

9. If a person drinks a cup of coffee contains 120mg of caffeine, after 24h, how much caffeine still remains in his body? (1 point)(Answer is "7.5mg".)

10. What is the rough lethal dose of caffeine per kilogram in human's body? (1 point) (Answer is "150mg/kilogram".)

APPENDIX 3

1. Do you like the style of the presented pictures? ( )
A. Like
B. Dislike
C. It does not matter for me, as long as it display the content clearly
D. Other:__________________

2. How long you actually viewed the learning material? ________minutes.

3. How long did you actually finish the post-test?_________ Minutes.

4. Do you think the topic you learned is: ( )
A. Very Difficult
B. Difficult
C. Not difficult but not easy
D. Easy
E. Very Easy

REFERENCES

Altman DG. Practical Statistics for Medical Research. Chapman and Hall 1991.

Armitage P, Berry G. Statistical Methods in Medical Research (3rd edition). Blackwell 1994.

Baek,Y.K.&Layne, B. H. (1988). Color, graphics, and animation in a computer-assisted learning tutorial lesson. Journal of Computer-Based Instruction, 15,131–135.

Baddeley, A. (1998).Human Memory, Allyn and Bacon, Boston.

Barak, M., Ashkar,T., Dori, Y. J. .(2011). Learning science via animated movies: Its effect on students' thinking and motivation. Computers & Education, 56, 839–846.

Chapman, Orville L. (2000). Learning Science Involves Language, Experience, and Modeling. Journal of Applied Developmental Psychology 21(1):97–108.

David, H. A.; Gunnink, Jason L. (1997). "The Paired *t* Test Under Artificial Pairing". The American Statistician 51 (1): 9–12.

Dwyer, F. M. (1978). Strategies for Improving Visual Learning. State College, PA: Learning Services.

Fisher Box, Joan (1987). "Guinness, Gosset, Fisher, and Small Samples". Statistical Science 2 (1):45-52.

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

Höffler, T. N., Prechtl, H., & Nerdel, C. (2010). The influence of visual cognitive style when learning from instructional animations and static pictures.Learning andIndividual Differences. doi: 10.1016/j.lindif.2010.03.001.

Höffler, T. N., & Leutner, D. (2010). The role of spatial ability in learning from instructional animations – Evidence for anability-as-compensator hypothesis. Computers in Human Behavior,27, 209–216

Lepper, M. R., & Malone, T. W. (1987). Intrinsic motivation and instructional effectiveness in computer-based education. In R. E. Snow & M. J. Farr (Eds.), Aptitude, learning, and instruction: Vol. 3. Conative and affective process analyses (pp. 255–286). Hillsdale: Lawrence Erlbaum Associates

Mayer, R. E. (1994).Visual aids to knowledge construction: Building mental representations from pictures and words. In W. Schnotz and W. Kulhavy (eds), Comprehension of Graphics (Amsterdam: Elsevier Science B. V).

Mayer, R. E. (1996). Learning strategies for making sense out of expository text: The SOI model for guiding three cognitive processes in knowledge construction. Educational Psychology Review, 8,357-371.

Mayer, R. E. (Ed.). (2005).The Cambridge Handbook of Multimedia Learning. New York: CambridgeUniversity Press.

Mayer, R. E., & Moreno, R. (2002). Aids to Computer-Based Multimedia Learning. *Learning and Instruction*, v12 n1, 107-119.

Mayer, R. E., & Moreno, R. (2002). Animation as an Aid to Multimedia Learning. Educational Psychology Review, v14 n1, 87-99.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. Journal of Educational Psychology, 93, 187-198.

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. E. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. Journal of Experimental Psychology: Applied, 11, 256−265.

Moreno, R. E., & Mayer, R. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. Journal of Educational Psychology, 92, 117-125.

J. NAJJAR, J. (1996). Multimedia Information and Learning. Jl. of Educational Multimedia and Hypermedia 5 (2), 129-150

Paivio, A. (1986).Mental Representations: A Dual Coding Approach, Oxford University Press,Oxford, England.

Rieber, L. P. (1990). Using computer animated graphics with science instruction with children.Journal of Educational Psychology, 82,135–140.

Rieber, L. P. (1991a). Animation, incidental learning, and continuing motivation. Journal of Educational Psychology, 83,318–328.

Rieber, L. P. (1991b). Effects of visual grouping strategies of computer-animated presentationson selective attention in science.Educational Technology, Research, and Development, 39,5–15.

Sweller, J. (1999).Instructional Design in Technical Areas, ACER, Camberwell, Australia.van Merrienboer, J. J. G. (1997).Training Complex Cognitive Skills, Educational Technology Publications, Englewood Cliffs, NJ.

Tousignant M, DesMarchais JE.(2002). Accuracy of Student Self-Assessment AbilityCompared to Their Own Performance in a Problem-Based Learning Medical Program:A Correlation Study. Advances in Health Sciences Education 7: 19–27, 2002.

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

Wittrock, M. C. (1974). Learning as a generative activity.Educ. Psychol.11: 87–95.

Yang, E.M.,Andre, T. and Greenbowe, T.J. (2003). Spatial Ability and the Impact of Visualization/ Animation on Learning Electrochemistry. International Journal of Science Education, 25(3), 329-349.