UNIVERSITY OF CALIFORNIA
Santa Barbara

Benefits of Turning the Illustrations in a Narrated Slideshow into Cartoons: An Extension of the Positivity Principle

A Thesis submitted in partial satisfaction of the

requirements for the degree Master of Arts

in Psychological & Brain Sciences

by

Fangzheng Zhao

Committee in charge:

Professor Richard E. Mayer, Chair

Professor Mary Hegarty

Professor Vanessa Woods

March 2023

The thesis of Fangzheng Zhao is approved

X

Mary Hegarty

X

Vanessa Woods

X

Richard E. Mayer, Committee Chair

January 2023

Benefits of Turning the Illustrations in a Narrated Slideshow into Cartoons:

An Extension of the Positivity Principle

by

Fangzheng Zhao

ACKNOWLEDGEMENTS

I would like to extend my sincerest appreciation to my advisor and committee chair, Dr. Richard E. Mayer, for his invaluable guidance, encouragement, and expert feedback throughout the course of my research. I am also deeply grateful to my committee members, Dr. Mary Hegarty and Dr. Vanessa Woods, for their generous contributions of expertise and valuable insights. My gratitude also extends to my graduate advisor, Katherine Lowe, for her invaluable assistance and advice. Lastly, I would like to acknowledge the National Science Foundation for their support of our research endeavors.

ABSTRACT

Benefits of Turning the Illustrations in a Narrated Slideshow into Cartoons:

An Extension of the Positivity Principle

by

Fangzheng Zhao

What is the effect on students' learning of converting a narrated slideshow with simple line drawings (original group) into one in which the key elements are rendered as colorful cartoon-like characters (cartoon group)? We conducted two between-subjects experiments in which the narrator's voice in both groups was a computer-generated female happy voice in Experiment 1 or a real female happy voice in Experiment 2. The cartoon group scored higher on a transfer posttest than the original group in both experiments. On subsequent questionnaires, the cartoon group reported feeling more positive (i.e., happy and content) during learning than the original group (in Experiment 1 and the combined experiments); and reported that the instructor was more engaging, better at facilitating learning, and more human-like (in Experiment 2 and the combined experiments). The results are consistent with the positivity principle, which predicts better learning from lessons in which key elements induce positive emotions.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

I. Introduction

A. Objective and Rationale

Consider two students sitting in front of a computer screen: one of the students is watching a narrated slideshow about a scientific topic with emotionally neutral line drawing elements (as exemplified in the left side of Figure 1), and another student is watching the same narrated slideshow with the same narration but with the slides incorporating color drawings of cartoon-like characters that represent the main elements (as exemplified in the right side of Figure 1). Which type of lesson, original slideshow or cartoon slideshow, will lead to better learning outcomes and experiences for learners? In the present study, we explore this issue, in an effort to expand the field of emotional design and contribute to principles for effective multimedia instructional design in computer-based environments.

Emotional design of multimedia lessons involves converting the graphics in ways that create visual appeal intended to induce a positive emotional response in learners, typically through representing key elements with round shapes, warm colors, and facial expressions (Mayer & Estrella, 2014; Plass & Kaplan, 2016; Plass et al., 2014; Um et al., 2012). This form of emotional design has been shown to improve students' learning outcomes and experiences in online learning venues (Gong et al., 2017; Kumi, et al., 2012; Mayer & Estrella, 2014; Plass et al., 2014; Um et al., 2012). In the present study, we seek to expand the domain of emotional design of multimedia lessons to include representing key elements as cartoon-like characters such as familiar characters from the Disney film, *Frozen.* Our goal is to determine whether this form of emotional design can also be effective in promoting improvements in students' learning outcomes and experiences. An important

aspect of both forms of emotional design is that they focus on increasing the emotional appeal of informationally relevant aspects of the visual display (i.e., incorporating interesting features that are relevant to the theme of the lesson) rather than irrelevant aspects of the display, which would thereby constitute *seductive details* (i.e., adding interesting but irrelevant features; Harp & Mayer, 1998).

This study focuses on the idea that online multimedia instructional messages should be designed to prime affective processing (e.g., positive emotional responses as indicated by ratings of felt emotion during learning), social processing (e.g., feelings of social connection with the instructor as indicated by ratings of the instructor's characteristics), and cognitive processing (e.g., active processing that leads to improved posttest performance). A substantial body of multimedia learning research has investigated design features that prime appropriate cognitive processing during learning (Clark & Mayer, 2016; Fiorella & Mayer, 2022; Mayer, 2021). However, based on the Cognitive-Affective Model of Learning with Media (Horovitz & Mayer, 2021; Lawson et al., 2021a, 2021b, 2021; Mayer, 2021; Moreno & Mayer, 2007), the process of learning includes not only students' cognitive processing of the material but also their affective processing (i.e., feelings of positive emotion during learning) and social processing (i.e., feelings of partnership connection with the instructor during learning). These learning processes work together to affect learning outcomes. Although recently there has been an increasing number of studies about how to prime appropriate affective and social processing in online multimedia learning (Fiorella, 2022; Fiorella & Mayer, 2022; Horovitz & Mayer, 2021; Lawson & Mayer, 2021; Lawson et al., 2021; Mayer, 2021), this remains an understudied issue in comparison to studies on how to

prime appropriate cognitive processing. Therefore, it is worthy of expanded study as in the present project.

In particular, our work is guided by the *positivity principle* (Horovitz & Mayer, 2021; Lawson & Mayer, 2021; Lawson et al., 2021a, 2021b, 2021), which holds that positive emotion conveyed in online learning materials can help students to respond more positively during learning, causing them to form a better social connection with the instructor and work harder to learn the material. That is, online multimedia lessons that incorporate positive emotional features focused on the key elements of the lesson can prime appropriate affective, social, and cognitive processing in multimedia learning that led to better learning outcomes. Overall, this study is the exploratory research, aiming to expand the classic emotional design approach into a more complex cartoon version.

II. Theoretical Background

A. Emotional Design and Multimedia Learning

Research on emotional design has focused on the color (i.e., warm appealing colors), shape (i.e., rounded shapes), and anthropomorphisms (i.e., facial expressions) of key elements in a lesson (Kumi, et al., 2013; Mayer & Estrella, 2014; Plass et. al., 2014, 2020; Um et al., 2012). Brom, Starkova, and D'Mello (2018) conducted a meta-analysis of 20 studies on the effect of facial anthropomorphism and/or pleasant colors on multimedia learning. They found that adding anthropomorphisms and/or pleasant colors could induce positive emotion, reduce the perceived difficulty, improve intrinsic motivation and enjoyment, and most importantly, facilitate learning (including retention and transfer test scores). They also found that the effects of anthropomorphism and colors were larger for younger learners than for older learners.

More recently, Wong and Adesope (2021) conducted a meta-analysis based on 28 studies involving manipulations of anthropomorphic graphics, colors, both anthropomorphic graphics and color, and both shape and facial expressions. They found that all four ways of implementing emotional design were effective in inducing positive emotion, enhancing intrinsic motivation and enjoyment, reducing perceived difficulty, and most importantly, increasing learning outcomes.

The present study is based on the three elements of the emotional design described in Plass et al.'s (2020) paper: color (i.e., warm colors), shape (i.e., round shapes), and anthropomorphisms (i.e., facial expression). In this review, we examine the impact of each of these elements separately, based on foundational studies.

For emotional design based on colors, previous research has shown that saturated and warm colors, especially light warm colors like yellow and orange, can increase the learner's feelings of pleasure and excitement, which can lead to better learning outcomes (Gorn, et al., 1997; Plass et. al., 2020). However, red color was not as effective because red color tends to indicate danger, failure, or mistake in the learning context, thereby inducing negative emotions (Plass et. al., 2020; Elliot, et al., 2007). In a study by Plass et al. (2020), there was a medium-sized effect of color on emotional arousal, showing that warm colors (e.g., orange) were more effective than neutral colors (e.g., grayscale). In addition, when applying the emotional design element of color in a multimedia learning study, Um et. al. (2012) found that positive emotion was induced by warm colors especially when it was combined with round shapes. Although several studies demonstrated the positive effect of warm color on positive emotion induction and learning outcome, another study from Kumi, et al. (2012) qualified this finding by showing that a strongly affective color such as yellow did not lead to an increase in recall, while a milder color like blue was more effective in promoting learning.

Research on shape has shown that rounded shapes are better at inducing positive emotions and lead to better retention and transfer outcomes than square or sharp-edged shapes (Plass et. al., 2014, 2020; Um el. al., 2012). This advantage of rounded visual shapes can be explained as a form of baby-face bias (Berry & McArthur, 1986; Zebrowitz & Montepare, 2008; Plass et al, 2020; Miesler, Leder, & Herrmann, 2011). According to the concept of baby-face bias, round features are typically shown on baby-like human faces including large eyes and short chins, which can remind people of baby-like characteristics such as innocence, cuteness, and helplessness. This baby-face effect is demonstrated by

neuroscience research, showing that baby faces can induce higher BOLD signal change than human faces in the amygdala and fusiform face area, which indicates greater affective processing (Zebrowitz, et al., 2009). In addition, Plass et. al (2014) and Um et al. (2012) conducted studies that applied the baby-face effect to multimedia platforms. The results showed that the students learning from multimedia lessons in which the key characters had rounded shapes not only showed higher positive emotion but also performed better on comprehension and transfer tests, confirming round shape as one of the emotional design elements that can effectively enhance learning outcomes.

Finally, for emotional design based on facial expressions (i.e., anthropomorphic design), the key elements in multimedia learning materials are redesigned to be represented as characters with their own facial expressions showing positive, negative, or neutral emotions. The rationale for how anthropomorphic design affects students' emotions involves the contagion effect (Hatfield, et al., 2014; Hatfield, et al., 2014), in which the emotions expressed by others affect the emotions felt by observers. Specifically, people's emotions can be impacted by others' facial expressions through an unconscious process called EmoGBL, leading people to automatically observe and mimic expressive cues from other people or even virtual agents (Loderer et al., 2019; Scherer & Coutinho, 2013; D'Mello & Graesser, 2012; McQuiggan & Lester, 2007).

In applying anthropomorphic design to multimedia instruction, researchers have been investigating whether it is effective in facilitating learning performance. For example, a study conducted by Gong et al. (2017) found that the anthropomorphism design facilitated transfer performance but did not affect students' emotions. In contrast, in another study that compared the effect of color, anthropomorphic design, and sound, the results showed that

adding facial expression increased positive emotion and reduced mental effort but had no effect on transfer performance (Uzun & Yildirim, 2018). Finally, an eye-tracking study conducted by Park, et al. (2015) showed that anthropomorphism did not induce positive emotions or facilitate learning outcomes.

Overall, there is encouraging evidence for incorporating emotional elements in multimedia lessons under appropriate conditions, mainly when revised elements are relevant to the theme of the lesson rather than irrelevant. Some mixed results of introducing warm colors, round shapes, and facial expressions may be attributed to whether these elements are applied to relevant or irrelevant elements in the lesson. Therefore, in the present study, we focus on applying emotional design features to relevant elements in the multimedia lesson.

In light of the growing research base investigating the three major elements of emotional design (i.e., color, shape, and facial expression), the present study seeks to expand the domain of emotional design to include the introduction of cartoon-like characters that represent key elements in a multimedia lesson. Cartoon-like characters can incorporate the traditional elements of emotional design by being rendered in warm colors, with rounded shapes, and facial expressions, but also bring the additional element of being likable, appealing, and familiar. This line of research helps broaden the way that emotional design can be implemented in multimedia instruction, and thereby broaden ways to induce positive emotion and deeper cognitive processing during learning.

B. Learning with Cartoons

In the present study, we use the term, *cartoons*, to refer to simple color drawings representing the features of characters in a humorous or exaggerated way. Cartoons, as a type of informative pictorial material, may be used as an application of emotional design if

they can be shown to induce positive emotions and contain enough information to promote learning. Research on the educational role of cartoon-like characters in multimedia lessons is in its initial stages but can be informed by related research on the impact of cartoons in educational settings. The majority of previous studies explored the effect of humorous cartoons on short-term retention of the cartoon content (Schmidt & Williams, 2001; Takahashi & Inoue, 2009) or using concept cartoons as aids for learning (Levin et. al., 1982; Akamaca, et al., 2009; Wyk, 2011; Eker & Karadeniz, 2014; Gamage, 2019).

For studies on memory for humorous cartoons, participants typically watched a series of cartoon pictures presented for several seconds for each picture. Then, they were asked to either remember the caption or title of the cartoon pictures or recall their detailed content. For example, in a study by Schmidt and Williams (2001), participants were shown one captioned cartoon at a time and then asked to write a description of each cartoon and its caption. After the entire series, they were asked to recall all the captions. Participants remembered humorous cartoons (i.e., with emotional features) better than neutral cartoons (i.e., without any emotional features) or weird cartoons (i.e., without any emotional features or logical content). Another study asked participants to view a series of cartoons and then redraw them as a retention test (Takahashi & Inoue, 2009). When the memory test was unexpected, students who saw humorous cartoons (i.e., with emotional features) group performed better on the recall test (i.e., redrawing the cartoon pictures) than students who saw non-humorous cartoons (i.e., without the emotional feature). However, this difference was eliminated when the memory test was expected.

Although the foregoing studies showed that humorous cartoons (i.e., with emotional features) can improve students' retention of cartoon content, those cartoons did not involve

academic learning materials. In addition, previous studies only tested students' capability to recall the materials (i.e., retention) rather than the capability to apply what they learned from the materials to a new situation (i.e., transfer). The present study is not focused on rote memory of arbitrary details of the individual pictures, but instead, we aimed to explore the effect of incorporating cartoon-like elements in multimedia instructional materials on learning outcomes, including performance on transfer tests.

Some previous studies explored the use of cartoons as learning aids. One popular method is called *concept cartoons*, which are the combination of visual elements and printed text of dialogues on to-be-learned topics (Keogh and Naylor, 1999). Within the dialogues, misconceptions and alternative conceptions are stated by different characters in the cartoons embedded with various expressions and explanations. Since the misconceptions that are stated in the dialogues are often based on students' intuitions and experiences in daily life, it helps students to recognize their potential misconceptions and also triggers motivation to learn (Stephenson & Warwick, 2002). For example, Akamca et al. (2009) used computer-aided concept cartoons to help teach the classification of living things in a 4^{th}-grade science technology class. The concept cartoons improved students' learning outcomes, improved students' motivation to engage in class interactions, and effectively eliminated misconceptions. Another study (Wyk, 2011) showed that using cartoons as a teaching tool in economics education effectively helped students to enhance their constructive learning, cooperative learning, and collaborative learning with peers.

Cartoons have also been used as learning aids intended to enhance motivation and learning. For example, Gamage (2018) used individual cartoons to describe a specific scenario that can help students understand the meaning of the relevant second-language

words. The results showed that students were motivated by the usage of the cartoons in the lessons and put more effort into the learning tasks with cartoons. Another study (Özdoğru & McMorris, 2013) found that students liked humorous cartoons inserted in college textbooks, even though those cartoons did not noticeably affect their learning outcomes. In a study with 4th graders (Eker and Karadeniz, 2014), students in the experiment group first learned 18 cartoons about the course topic and drew a cartoon by themselves about the topic at the end of the lesson, while the control group learned the same topic without cartoons. The results showed that retention test scores for the cartoon learning group were higher than those of the control group, suggesting that inserted cartoons can effectively facilitate learning outcomes.

Furthermore, cartoon usage can also be an aid for lessons on second language learning. For example, Rosseto and Cheira-Macchia (2011) used The Simpsons cartoon series in Italian dialogue learning materials to help students learn faster based on a familiar storyline. Another study used cartoons to teach primary school students Turkish language grammar (Yaman, 2010), which found that students learning with cartoons were more motivated, were more willing to participate in the class, and achieved better learning outcomes. Although the cartoons used in these studies were not based on emotional design principles, this work encourages further exploration of the effects of incorporating cartoon-like characters as the main elements in multimedia lessons as a form of emotional design.

C. Theory and Prediction

Based on previous studies about emotional design and cartoon learning, we decided to apply emotional design to the relevant informational elements in a multimedia lesson rather than to unrelated and uninformative elements (which would constitute seductive details). We refer to this approach as *informative emotional design* because emotional design

is intended to draw attention to aspects of the graphic that are relevant for learning. Also, in light of their promising but largely untested potential, we decided to use cartoons as a way to apply emotional design in a multimedia lesson. In addition, in contrast to previous cartoon learning studies that only focused on retention performance, we also examined the effect of informative emotional design on students' affective processing (i.e., felt emotions), social processing (i.e., ratings of instructor characteristics), and deep learning outcomes (i.e., transfer as well as retention test performance).

We ground our predictions on an abbreviated version of the Cognitive-Affective Model of Learning with Media, which we summarize in Figure 2 (Horovitz & Mayer, 2021; Lawson et al., 2021; Mayer, 2021; Moreno & Mayer, 2007). According to this model, meaningful learning with media involves a cascading series of steps, involving affective, social, and cognitive processing, leading to learning outcomes. In the first step, students engage in affective processing by feeling positive emotion and enjoyment primed by positive emotional design features in the lesson. In the second step, students engage in social processing by feeling a sense of social partnership with the instructor, in line with their positive emotions primed by the positive emotional design features of the lesson. In the third step, the increase in positive emotion and rapport with the instructor leads to cognitive processing aimed at making sense of the material, which results in improved learning outcomes.

This chain of processing is exemplified in the *positivity principle*, which proposes that positive learning environments lead to a better learning experience and better learning outcomes for students than negative learning environments (Horovitz & Mayer, 2021; Lawson & Mayer, 2021; Lawson et al., 2021a, 2021b, 2021). In this study, we redesigned a

traditional narrated slideshow lesson into a cartoon version, including various personalized characters in color with facial expressions to induce positive emotions and social connections. Based on the positivity principle, as derived from the Cognitive-Affective Model of Learning with Media, we expected our cartoon-based emotional design implementation to affect students' affective, social, and cognitive processing, as detailed in the following predictions.

Hypothesis 1 (Cognitive processes and learning outcomes): Students' learning outcomes are affected by the emotional design of the lesson. The rationale is that students work harder to make sense of the material when they are attracted to the key elements in the lesson and positive emotional design causes them to have better rapport with the instructor.

1a) Learners who receive the cartoon lesson have a higher score on the retention test than those who receive the original lesson.

1b) Learners who receive the cartoon lesson have a higher score on the transfer test than those who receive the original lesson.

Hypothesis 2 (Affective processes): Learners' felt emotions are affected by the emotional design of the lesson. The rationale is that positive emotion-inducing features in the slides cause the learner to experience stronger positive emotion.

2a) Learners who receive the cartoon lesson rate their level of positive felt emotion (i.e., happy and content) to be higher than learners who receive the original narrated slides lesson.

2b) Learners who receive the original narrated slides lesson rate their level of negative felt emotion (i.e., frustrated and bored) to be higher than learners who receive the cartoon lesson.

Hypothesis 3 (Social processes): Learners' partnership connection with the instructor is influenced by the emotional design of the lesson. The rationale is that positive emotional design causes the learner to build a stronger social partnership with the learner.

3a) Learners who receive the cartoon lesson give a higher rating of the instructor as facilitating learning than those who receive the original lesson.

3b) Learners who receive the cartoon lesson give a higher rating of the credibility of the instructor than those who receive the original lesson.

3c) Learners who receive the cartoon lesson give a higher rating of the instructor as human-like than those who receive the original lesson.

3d) Learners who receive the cartoon lesson give a higher rating of the instructor as engaging than those who receive the original lesson.

Ratings on motivational and cognitive processes are exploratory, although in line with the foregoing predictions, we expect the cartoon group to give higher ratings than the original group on their level of enjoyment, interest, motivation, and effort, and to give a lower rating on their level of difficulty. Ratings on cognitive load also are exploratory, although in line with the foregoing predictions, we expect the cartoon group to give higher ratings of germane load and lower ratings of extraneous load, as compared to the original group.

In combining Experiments 1 and 2, we expect to see the same patterns in both studies based on the above predictions, with no interactions involving the type of voice. Comparisons between the human voice and machine voice are exploratory. The classic research on this issue yields the voice principle (Mayer, 2021), which states that people learn better with an appealing human voice than with a machine voice. However, in the

present study, advances in speech synthesis technology may render machine voice as effective as the human voice in inducing positive emotion (Craig & Schroeder, 2017, 2018).

III. Experiment 1 (Machine Voice)

In this experiment, participants were asked to watch a 2.5-minute video lesson that explains how lightning forms. In a between-subject design, this experiment had two versions of a narrated slideshow: the original lesson that included line drawn slides (original group) and the emotional-designed lesson that included cartoon-style slides based on redrawing the original slides (cartoon group). Both lessons have the same narrated voices, consisting of a computer-generated female happy voice. After the lesson, participants took a postquestionnaire (involving ratings of the learners' felt happy, content, bored, or frustrated emotions during the lesson; ratings of learners' feelings of partnership connection with the instructor; and ratings of learners' cognitive and motivational processing) and a posttest (involving retention and transfer).

A. Method

a. Participants

The participants were 101 undergraduate students from a university in California. The mean age was 18.87 years (SD = 1.75); 28 students identified as males, 72 identified as females, and 1 as another gender. An additional three participants were excluded because they did not finish the study. All students indicated low prior knowledge about the topic of the lesson (i.e., lightning formation) based on a prior knowledge survey score (M = 2.88, SD = 1.78, maximum = 12). Participants were recruited from the Psychology Subject Pool, and they fulfilled a course requirement by participating in the experiment.

b. Design

In a between-subjects design, 51 participants were randomly assigned to the cartoon group and 50 were randomly assigned to the original group. For all participants, the dependent measures were retention test score, transfer test score, ratings of learners' felt emotions, ratings of learners' feeling of partnership connection with the instructor, and ratings of learners' cognitive and motivational processing.

c. Materials

The materials included a prior knowledge questionnaire, two versions of a 2.5-minute narrated slideshow about the formation of lightning, three postquestionnaires (i.e., ratings of the learners' felt emotion, and ratings of the learners' feeling of partnership connection with the instructor), two posttests (i.e., retention test and transfer test), and a brief questionnaire about demographic information[1]. All research materials were published as a Qualtrics survey and presented on Dell or iMac desktop computers in individual cubicles in a research lab.

The prior knowledge questionnaire aims to measure how much participants know about meteorology before taking the video lesson and to prevent a testing effect, in which a pretest might serve as a learning event and guide learning of the subsequent lesson (Brown, Roediger, & McDaniel, 2014; Roediger & Karpicke, 2006). The same questionnaire has been used in previous studies (Mayer & Chandler, 2001; Mayer, Heiser, & Lonn, 2001;

[1] There was also a Cognitive and Motivational Processing survey in Experiments 1 and 2, consisting of five exploratory rating items concerning diverse aspects of the learning experience, with each item tapping something different. We do not report the data for this survey in this paper because of a relatively low level of internal reliability ($\alpha = .66$) and because there was only one item for each construct.

Mayer & Johnson, 2008; Moreno & Mayer, 1999, 2000, 2002) and consists of two questions. The first question asked participants to rate how much knowledge they think they have about meteorology, on a scale of 1 (very little) to 5 (very much). The second question asked participants to check any of seven statements about their experience or understanding of meteorology-related knowledge that applied to them (i.e., "I can explain what makes the wind blow." Or "I can distinguish between cumulus and nimbus clouds."). The prior knowledge score was calculated by summing the rating of the first question and the number of items chosen in the second question together. Cronbach's alpha was $\alpha = 0.73$, which represents good internal reliability.

As the learning materials, there were two versions of a 16-slide narrated slideshow lesson with the same content about the explanation of lightning formation. The script and voice used for the narration were the same in both lessons, which involved a happy female voice generated by the online text-to-speech portal, *Voicery* (https://voicery.com). Table 1 shows the entire script, consisting of 287 words broken down into 16 segments, one for each slide. The lesson lasted 2.5 minutes.

For the original lesson, the slides were 16 line drawn illustrations depicting steps in the lightning formation and were the same as materials used in previous studies, as exemplified in the left panel of Figure 1 (Mayer & Chandler, 2001; Mayer, Heiser, & Lonn, 2001; Mayer & Johnson, 2008; Moreno & Mayer, 1999, 2000, 2002). The lesson started with cool moist air being heated, rising higher to the sky, and forming a cloud; then the lesson continued with the electrical particles being generated in the cloud and separated into positive and negative particles because of the temperature difference within the cloud; finally, the lesson ended with the formation of the flash of lightning by the interaction

between negative charges on the bottom of the cloud and the positive charges on the ground. The slides for the original lesson are shown in this link: https://osf.io/6gxa2/?view_only=a9a0be2ad2324655b3ebe248bfab0d35.

For the cartoon lesson, the slides were 16 cartoon-like illustrations redrawn in color from the original lesson, as exemplified in the right panel of Figure 1. In these cartoon-like slides, we used characters from the Disney movie *Frozen* including Elsa, Olaf, and Snowman to represent the cold air and low temperature mentioned in the lesson and increase the familiarity of the content to the student. In addition, we used water droplets with various colors (e.g., blueish color indicates cold temperature and reddish color indicates hot temperature) and different facial expressions (e.g., smiley faces mean rising and flying to the sky while frustrated faces mean freezing and falling) to represent the changes of temperature and weight (e.g., whether the water droplets were frozen and became heavier or not). Therefore, in our efforts to add emotional design to the lesson, we avoided seductive details that were irrelevant to the main content that we wanted students to learn. All the changes in the redrawn slides were aimed to provide students with more interesting and straightforward elements that were directly related to the essential information in the lesson. The slides for the cartoon lesson are shown in this link: https://osf.io/6gxa2/?view_only=a9a0be2ad2324655b3ebe248bfab0d35.

The learner emotion survey aimed to measure the learners' felt emotions while learning with the original lesson or the cartoon lesson (α = 0.82). This survey has been used in several previous studies (Lawson et al., 2021; Horovitz & Mayer, 2021). Participants were asked to rate four different felt emotions on a five-point scale from Strongly Disagree (1) to Strongly Agree (5). The items in this survey were: "This lesson made me feel happy",

"This lesson made me feel content", "This lesson made me feel bored", and "This lesson made me feel frustrated."

The Agent Persona Instrument (API) measures participants' feelings of partnership with the virtual instructor and is based on a previous study by Baylor, Ryu, and Shen (2003). Since the lessons we used in this study do not have embodied instructors, the instructor here means the narrator's voice in the videos. The goal of the API was to explore whether learners could build a partnership connection (i.e., social connection) with the instructor (i.e., the narrator with machine voices) and whether the emotional-designed lesson (i.e., cartoon lesson) could improve learners' feelings of partnership connection with the instructor as compared to the original lesson. The API survey had 23 items in total, including four subscales. The first subscale included ten items about how well the instructor could facilitate learning (e.g., "The instructor led me to think more deeply about the presentation.", $\alpha = 0.94$). The second subscale consisted of four items, which concerned the credibility of the instructor (e.g., "The instructor was knowledgeable.", $\alpha = 0.89$). The third subscale included five items about the level of human-like character of the instructor (e.g., "The instructor was human-like.", $\alpha = 0.91$). Finally, the fourth subscale contained five items concerning the engagement level of the instructor (e.g., "The instructor was motivating.", $\alpha = 0.90$).

As shown in Table 2, the post-test included a retention test with one short-answer question and a transfer test with four short-answer questions. Several previous studies have also used the same posttest (Mayer & Chandler, 2001; Mayer, Heiser, & Lonn, 2001; Mayer & Johnson, 2008; Moreno & Mayer, 1999, 2000, 2002). The retention test aimed to measure students' proficiency at remembering the lesson's content, which mainly depends on the

selecting process in the cognitive theory of multimedia learning (Mayer, 2021). In the test, participants were asked to explain lightning formation by summarizing the whole process based on their memory of the content in the lesson. The grading rubric of the retention test contained 41 key ideas in the lesson; each student's score is the number of key ideas included in their answer regardless of wording. Cronbach's alpha based on the 41 items was 0.78. Two scorers independently applied the rubric, with an inter-rater reliability of $r = 0.96$. Disagreements between the two raters were resolved through consensus.

The transfer test aimed to measure students' proficiency at applying the knowledge they learned in a new situation, which required the selecting, organizing, and integrating processes to build a mental model linked with their prior knowledge, as described in the cognitive theory of multimedia learning (Mayer, 2021). The grading rubric for the transfer test included several acceptable possible answers for each question; therefore, each student's transfer test score was the total number of acceptable answers they mentioned across the four transfer questions, with a maximum score of 26. Cronbach's alpha based on the four transfer questions was 0.67. Two scorers independently applied the rubric, with an inter-rater reliability of $r = 0.81$. Disagreements between the two raters were resolved through consensus.

Finally**,** the demographic questionnaire asked participants to report their age and gender.

d. Procedure

Participants were randomly assigned to the original group or the cartoon group and were tested in groups of up to three participants per session. This study was conducted in a room in a psychology laboratory consisting of five Dell or five iMac desktop computer

stations, each in a cubicle that visually shielded each participant. The experiment was implemented on Qualtrics. After participants came to the lab, they were guided to sit in front of a computer with the initial page of the Qualtrics survey on the screen. Participants were asked to read a consent form that provided a general description of the study. After consenting to participate, participants finished the prior knowledge survey. Next, they were asked to watch a narrated slideshow lesson carefully and expect a test after the lesson. Then, participants received either the original lesson or the cartoon lesson, based on their group assignment. Both lessons took 2.5 minutes to play without a stop or rewind function available for participants. After watching the lesson, participants experienced a one-minute distractive interlude by playing the spatial game Tetris. Afterward, they successively finished the felt emotion survey, the affective and cognitive processing survey, and the API survey at their own rates, followed by the posttest, which included one retention question with a time limit of four minutes and four transfer questions with a time limit of two minutes for each question. Finally, participants received a brief demographic questionnaire, received a debriefing, and then were thanked for their participation. In this study, we obtained IRB approval and followed guidelines for the ethical treatment of human subjects.

B. Results

a. Do the Groups Differ on Basic Characteristics?

First of all, it is important to ensure that the two groups are equivalent in basic characteristics. For the age of the participants, there was no significant difference between the two groups, $t(99) = -0.18$, $p = 0.86$. For the gender of the participants, there was no significant difference in the gender ratio between the two groups, $X^2(2, N = 101) = 1.43$, $p =$

0.49. Finally, for the prior knowledge level of the participants, there was no significant difference between the two groups, $t(99) = -1.12$, $p = 0.26$. We can conclude that the two groups did not differ in basic characteristics.

b. Hypothesis 1: Are learners' learning outcomes affected by the emotional design of the lesson?

As the main focus of this study, we were interested in exploring whether the emotional-designed lesson could prime deeper cognitive processing leading to better learning outcomes than the original lesson. However, because the posttest scores showed a floor effect creating skewed distributions, we chose the Mann-Whitney U test (with alpha set at .05) to compare the groups on retention and transfer tests. The mean scores (and standard deviations) of each group on the retention and transfer test are shown in Table 3.

For the retention test, we hypothesized that the learners in the cartoon group would perform better than those in the original group. However, in contrast to hypothesis 1a, the Mann-Whitney U test did not show a significant difference between cartoon group (M = 4.80, SD = 4.21) and original group (M = 4.28, SD = 3.53), U = 1200.50, p = 0.61.

For the transfer test, hypothesis 1b was supported, showing that the cartoon group (M = 3.37, SD = 2.33) produced higher transfer test scores than the original group (M = 2.36, SD = 1.72), U = 954.00, p = 0.03, with a medium effect size of d = 0.49. In this study, we are most concerned with transfer test scores as a measure of learning outcome, because it is intended to measure the learner's ability to apply the learned knowledge to new situations, which reflects engagement in deeper processing during learning. We conclude that adding cartoon-like features improved learning outcomes, specifically transfer test performance, reflecting deeper learning of the material.

c. Hypothesis 2: Are learners' felt emotions during learning affected by the emotional design of the lesson?

To investigate whether adding emotional design to a lesson (i.e., cartoon lesson) could noticeably affect students' felt emotions, we used an independent-sample t-test to compare the ratings on the felt emotion survey between the cartoon group and the original group. The mean ratings of felt emotions with standard deviations are shown in Table 4.

First, hypothesis 2a states that learners should feel more positive (i.e., happy and content) with the cartoon lesson than with the narrated slides lesson. Consistent with hypothesis 2a, for the ratings of happy emotion, we found a significant effect favoring the cartoon group, $t(99) = -2.65$, $p = 0.02$. As shown in the top line of Table 4, learners in cartoon group ($M = 3.20$, $SD = 0.85$) reported feeling happier than those in original group ($M = 2.68$, $SD = 1.10$), with a medium effect size of $d = 0.52$. Inconsistent with hypothesis 2a, we did not find a significant effect favoring the cartoon group for ratings of content emotion, $t(99) = -2.11$, $p = 0.08$. However, as recorded in the second line of Table 4, there is a trend showing that learners in cartoon group ($M = 3.22$, $SD = 0.92$) felt more content than those in original group ($M = 2.80$, $SD = 1.05$), with a small-to-medium effect size of $d = 0.43$.

Second, hypothesis 2b posits that learners should feel more negative (i.e., bored and frustrated) with the original lesson than with the cartoon lesson. However, inconsistent with hypothesis 2b, an independent-sample t-test did not show any significant differences in the ratings of bored emotion, $t(99) = 0.89$, $p = 0.38$, or the ratings of frustrated emotion, $t(99) = -0.14$, $p = 0.89$, between the original group ($M = 3.30, 2.26$, $SD = 1.40, 1.24$, respectively) and cartoon group ($M = 3.08, 2.29$, $SD = 1.09, 1.19$, respectively). These results,

summarized in lines 3 and 4 of Table 3, indicate that learners did not feel significantly more negative when learning with the original lesson than with the cartoon lesson.

In conclusion, the evidence partially supported the second hypothesis by showing that the emotional design of the lesson affected students' felt emotions. To be more specific, the cartoon lesson made learners feel happier, whereas the original lesson did not lead learners to feel more negative during the lesson.

d. Hypothesis 3: Are learners' partnership connections with the instructor affected by the emotional design of the lesson?

The API (The Agent Persona Instrument; Baylor, Ryu, & Shen, 2003) is a survey aimed at measuring learners' partnership connections with a virtual instructor, including subscales for facilitating learning, credible, human-like, and engaging. To explore the effect of emotional design (i.e., cartoon lesson vs. original lesson) on learners' feelings of connections with their virtual instructor (i.e., the computer-synthesized voice), we used independent-sample t-tests to compare the difference of the mean ratings between two groups on each of the four subscales. The mean ratings (and standard deviations) on the four subscales for each group are reported in Table 5.

For hypothesis 3a, it predicted that students in the cartoon group should give a higher rating for facilitating learning than those in the original group. However, the hypothesis was not supported by an independent-sample t-test, $t(99) = -1.52$, $p = 0.26$, showing that the cartoon lesson ($M = 4.02$, $SD = 1.19$) was not rated to be more facilitative on learning than the original lesson ($M = 3.61$, $SD = 1.50$). According to hypothesis 3b, students in the cartoon group should rate the instructor as more credible than those in the original group. However, we did not find a significant difference on credibility ratings between the cartoon

lesson group (M = 5.11, SD = 1.00) and the original lesson (M = 4.86, SD = 1.21), t(99) = -1.14, p = 0.52. In addition, hypothesis 3c posited that students in the cartoon group should give a higher rating for the instructor seeming human-like than those in the original group. In contrast, an independent t-test failed to support hypothesis 3c, showing that there was no significant difference in the ratings between the cartoon group (M = 3.10, SD = 1.49) and the original group (M = 2.83, SD = 1.53), t(99) = -0.90, p = 0.74. Finally, hypothesis 3d stated that learners in the cartoon group should give a higher rating of the instructor being engaging than those in the original group. Inconsistent with this hypothesis, a two-tailed independent-sample t-test showed that the cartoon lesson (M = 4.11, SD = 1.35) led to higher ratings of the instructor as engaging than the original lesson (M = 3.52, SD = 1.65), t(99) = -1.97, p = 0.05, yielding a small-to-medium effect size of d = 0.40.

In conclusion, there was not consistent evidence showing that the partnership connection between learners and a virtual instructor (i.e., computer-synthesized voice) was affected by the emotional design of the lesson, even though all trends were in the predicted direction.

IV. Experiment 2 (Human Voice)

After experiment 1, we provided participants with one additional question asking for their comments about the video lesson. Based on their answer, we realized that many participants (i.e., around 50%) in both conditions provided negative comments about the narrator's voice. Therefore, it is possible that the unsatisfying voice reduced the effect of informative emotional design (i.e., cartoon lessons) on learners' affective, social, and cognitive processing. Experiment 2 served as a replication of Experiment 1, except that the voice in the lesson was an appealing human voice rather than a machine voice.

A. Method

a. Participants

The participants were 100 undergraduate students in a university in California, with 74 female participants, 24 male participants, 1 other gender participant, and 1 who preferred not to say. The mean age is 18.95 (SD = 1.28). Three additional participants were excluded because they did not finish the experiment. All students showed a low prior knowledge of meteorology (M = 3.35, SD = 1.65, with the highest prior knowledge score of 12). As in Experiment 1, the participants were recruited from the Psychology Subject Pool to fulfill a course requirement.

b. Design

In a between-subject design, participants were randomly assigned in two conditions, with 50 participants receiving the cartoon lesson (i.e., cartoon group) and 50 participants receiving the original narrated slides lesson (i.e., original group). The dependent measurements were the same as those in Experiment 1, including the learning outcomes (i.e., retention and transfer test scores), ratings of felt emotion, ratings of learners' partnership connections with the instructor, and ratings of learners' cognitive and motivational processing.

c. Materials

The materials and apparatus in Experiment 2 were the same as in Experiment 1, except narrator's voice in both versions of narrated slideshow lessons was changed from a synthesized happy female voice to a human happy female voice. We also added a cognitive load survey (presented after the API survey), but we do not report on the results because it is

not related to our predictions and is not part of the replication of Experiment 1. The Cronbach's alpha for the Felt Emotion survey is 0.75. For the reliability check of the posttests, Cronbach's alpha based on the 41 items of the retention test was 0.73. Two scorers independently applied the rubric, with an inter-rater reliability of r = 0.88. In addition, Cronbach's alpha based on the 19 items based on the four transfer questions was 0.67. Two scorers independently applied the rubric, with an inter-rater reliability of r = 0.81. Disagreements between the two raters were resolved through consensus.

B. Results

a. Do the Groups Differ on Basic Characteristics?

We first analyzed the basic characteristics data to ensure equivalence between the cartoon group and the original group. For the age of the participants, we did not find a significant difference between the two groups, $t(98) = 0.39$, $p = 0.70$. There was no significant difference between the groups for the prior knowledge level, $t(98) = -0.30$, $p = 0.76$. Finally, for the gender ratio, we did not find any significant difference between the groups, $X^2(3, N = 100) = 2.83$, $p = 0.42$. In conclusion, the two groups were not different in their basic characteristics.

b. Hypothesis 1: Are learners' learning outcomes affected by the emotional design of the lesson?

We were most interested in whether the emotional design of the lesson could affect students' learning outcomes in line with hypothesis 1. Therefore, we first compared the difference in retention and transfer posttest scores between the cartoon group and the original group. Since both the retention test scores and transfer test scores showed right-

skewed distributions (i.e., floor effect), we used the Mann-Whitney U test (with alpha set at 0.05) just like in Experiment 1. The mean and standard deviations of each group on the retention and transfer tests for Experiment 2 are reported in Table 3.

Hypothesis 1a is that the cartoon group would perform better on the retention test than the original group. Inconsistent with the hypothesis 1a, the Mann-Whitney U test did not show a significantly higher retention score for the cartoon group (M = 5.58, SD = 3.48) than for the original group (M = 4.48, SD = 3.59), U = 994.5, p = 0.08, with effect size d = 0.31. However, consistent with hypothesis 1b, the Mann-Whitney U tests showed that the cartoon group (M = 3.74, SD = 2.57) performed better on the transfer test than the original group (M = 2.56, SD = 2.26), U = 908.5, p = 0.02, with an effect size d = 0.49.

Since transfer test scores indicates more about the learners' deep understanding of the material and their capabilities of applying the knowledge to new situations, we were more interested in transfer performance than retention performance. In conclusion, we found that converting an original narrated slideshow lesson into a cartoon narrated slideshow lesson facilitated students' learning, as indicated specifically on transfer posttests. These finding replicates Experiment 1 and together with the transfer test results of Experiment 1 constitutes the major empirical contribution of this study.

c. Hypothesis 2: Are learners' felt emotions affected by the emotional design of the lesson?

Secondly, we were interested in whether converting from the original lesson to the cartoon lesson (i.e., informative emotional design) could significantly influence learners' felt emotions. To answer this question, we compared the difference in ratings of felt emotion between the cartoon and original groups using independent-sample t-tests. Table 4 shows the means and standard deviations of the felt emotion ratings for each group.

Inconsistent with the hypothesis 2a, the learners in the cartoon group did not feel significantly more positive (i.e., happy and content) than those in the original group. To be more specific, there was no significant effect showing the rating of felt happy emotion in the cartoon lesson group (M = 3.08, SD = 1.05) was significantly higher than in the original group (M = 2.78, SD = 0.84), $t(98) = -1.58$, $p = 0.12$, $d = 0.32$; and there was no significant difference in ratings of felt content emotion between the cartoon group (M = 3.22, SD = 1.09) and the original group (M = 3.00, SD = 1.01), $t(98) = -1.05$, $p = 0.30$, $d = 0.20$.

In addition, inconsistent with the hypothesis 2b, learners did not feel significantly more negative (i.e., bored and frustrated) while learning with original narrated slides lesson than with the cartoon lesson. To be more specific, we did not find any significant difference in the ratings of bored emotion between the cartoon group (M = 3.36, SD = 1.06) and the original group (M = 3.34, SD = 1.06), $t(98) = -0.09$, $p = 0.93$, $d = 0.02$; and we did not find a significant difference in the ratings of frustrated emotion between the cartoon group (M = 2.24, SD = 1.20) and the original group (M = 2.00, SD = 1.01), $t(98) = -1.08$, $p = 0.28$, d = 0.22.

In conclusion, in contrast to Experiment 1, the evidence did not support the second hypothesis, indicating that changing to a cartoon lesson did not significantly affect learners' felt emotions.

d. Hypothesis 3: Are learners' partnership connections with the instructor affected by the emotional design of the lesson?

To investigate whether converting the original lesson to a cartoon lesson affected learners' partnership connections with the instructor (i.e., human voice), we used independent-sample t-tests to compare the two groups on the ratings of each of the four

subscales in the API survey, including facilitating learning, credible, human-like, and engaging. The mean and the standard deviation of the ratings by condition are reported in Table 5.

Inconsistent with the hypothesis 3a, students in the cartoon group (M = 4.29, SD = 1.34) did not give a significantly higher rating of the instructor as facilitating learning than for the original group (M = 3.82, SD = 1.16), $t(98) = -1.87$, $p = 0.06$, $d = 0.38$. Moreover, inconsistent with hypothesis 3b, students in the cartoon group (M = 5.32, SD = 0.98) did not rate the instructor as significantly more credible than those in the original group (M = 5.07, SD = 1.02), $t(98) = -1.28$, $p = 0.20$, $d = 0.25$.

However, consistent with the hypothesis 3c, the rating of the instructor as human-like for the cartoon group (M = 3.87, SD = 1.53) was significantly higher than for the original lesson group (M = 3.17, SD = 1.28), $t(98) = -2.49$, $p = 0.02$, with the medium effect size $d = 0.50$. Similarly, hypothesis 3d was also confirmed ($t(98) = -3.60$, $p < 0.01$), showing that the students in the cartoon group (M = 3.76, SD = 1.50) gave a significantly higher rating of the instructor as engaging than those in the original group (M = 2.79, SD = 1.16), with a large effect size $d = 0.72$.

In conclusion, in two of four measures, the perceived relationship between learners and instructors was affected by the emotional design of the lesson. To be more specific, although ratings of the instructor's facilitative and credible level were not significantly different between the two groups, the cartoon made students feel the instructor was more human-like and engaging than the original lesson.

V. Supplementary Analysis – Combining Data from Experiment 1 and 2

As a supplemental analysis, we combined the data from Experiment 1 and Experiment 2, creating a 2 (cartoon group versus original group) x 2 (synthesized voice versus human voice) between-subjects design. To analyze the data for felt emotion, partnership connections, and cognitive and motivational processing, we used 2x2 ANOVA tests to investigate the main effect of the type of the lesson (i.e., cartoon lesson vs. original lesson), the main effect of the narrator's voice (i.e., synthesized voice vs. human voice), and interactions between lesson type and voice type. For the posttest scores, we still used the Mann-Whitney U test to compare each factor separately because of the floor effect in the distribution of the scores.

A. Hypothesis 1: Are learners' learning outcomes affected by the emotional design of the lesson?

The means and standard deviations on the posttest retention and transfer scores by conditions (i.e., cartoon lesson vs. original lesson) are shown in Table 3. According to the results of the Mann-Whitney U test, there was no significant difference in retention test scores between the cartoon group ($M = 5.19$, $SD = 3.86$) and original group ($M = 4.38$, $SD = 3.54$; $U = 4401.00$, $p = 0.11$), and no significant difference between synthesized voice ($M = 4.54$, $SD = 3.87$) and human voice ($M = 5.03$, $SD = 3.56$; $U = 4544.50$, $p = 0.22$). However, as in Experiment 1 and 2, the cartoon group ($M = 3.55$, $SD = 2.45$) achieved a significantly higher transfer score than the original group ($M = 2.46$, $SD = 2.00$), $U = 3728.00$, $p < 0.01$, with a small to medium effect size $d = 0.49$. There were no differences in transfer scores between synthesized voices ($M = 2.87$, $SD = 2.10$) and human voices ($M = 3.15$, $SD = 2.48$), $U = 4855.50$, $p = 0.63$. In conclusion, in line with hypothesis 1b, this analysis shows that the

cartoon lesson facilitated students' transfer performance, no matter whether the narrator's voice was a synthesized voice or a human voice.

B. Hypothesis 2: Are learners' felt emotions affected by the emotional design of the lesson?

Table 4 shows the means (and standard deviations) of the felt emotion ratings by condition (i.e., cartoon lesson vs. original lesson). For the felt happy ratings, there was a significant main effect emotional design, showing that the cartoon group (M = 3.14, SD = 0.95) gave a significantly higher rating of felt happy emotion than the original group (M = 2.73, SD = 0.97), $F(1, 197) = 9.00$, $p < 0.01$, with a medium effect size $d = 0.43$. There was no significant main effect of voice type, $F(1, 197) = 0.00$, $p = 0.95$, nor interaction between lesson design type and voice type, $F(1, 197) = 0.63$, $p = 0.43$. For the felt content ratings, there was a significant main effect of emotional design, $F(1, 197) = 4.87$, $p = 0.03$, showing that the cartoon lesson (M = 3.22, SD = 1.01) led learners to feel more content than the original lesson (M = 2.90, SD = 1.03), with the small to medium effect size $d = 0.31$. There were no main effect of voice type, $F(1, 197) = 0.50$, $p = 0.48$, nor interaction, $F(1, 197) = 0.46$, $p = 0.50$. For the felt bored ratings, a 2x2 ANOVA did not show any significant effect for emotional design, $F(1, 197) = 0.38$, $p = 0.54$; voice ,$F(1, 197) = 0.96$, $p = 0.33$; or interaction, $F(1, 197) = 0.54$, $p = 0.46$. For felt frustrated ratings, a 2 x 2 ANOVA did not show any significant effects of emotional design, $F(1, 197) = 0.70$, $p = 0.41$; voice, $F(1, 197) = 0.91$, $p = 0.34$; or interaction, $F(1, 197) = 0.39$, $p = 0.53$. We conclude that in line with hypothesis 2a, the cartoon lesson caused learners to experience higher levels of positive emotion during learning, no matter whether the narrator's voice was a synthesized voice or a human voice.

C. Hypothesis 3: Are learners' partnership connections with the instructor affected by the emotional design of the lesson?

Table 5 shows the means and standard deviations of the ratings on the four subscales in the API survey for the cartoon and original groups. For the ratings of how well the instructor facilitates learning, there was a significant main effect of emotional design, showing that the cartoon lesson ($M = 4.15$, $SD = 1.27$) made learners consider the instructor to be more facilitative of learning than the original lesson ($M = 3.71$, $SD = 1.34$), $F(1, 197) = 5.72$, $p = 0.02$, with a small to medium effect size $d = 0.34$. There was no significant main effect of voice type, $F(1, 197) = 1.70$, $p = 0.19$, nor interaction, $F(1, 197) = 0.03$, $p = 0.87$.

For ratings of how human-like the instructor seemed, there was a significant main effect of emotional design, showing that the ratings of human-like in the cartoon group ($M = 3.48$, $SD = 1.55$) were significantly higher than those in original lesson group ($M = 3.00$, $SD = 1.41$), $F(1, 197) = 5.55$, $p = 0.02$, with the small to medium effect size $d = 0.32$. As expected, there was also a main effect of voice type, indicating that the human voice ($M = 3.52$, $SD = 1.44$) sounded more human-like than the synthesized voice ($M = 2.96$, $SD = 1.51$), $F(1, 197) = 7.27$, $p = 0.01$, with the small to medium effect size $d = 0.38$. There was no interaction between emotional design and voice, $F(1, 197) = 1.09$, $p = 0.30$.

For ratings of how engaging the instructor seemed, a 2x2 ANOVA indicated a main effect of emotional design, $F(1, 197) = 14.92$, $p < 0.01$, with a medium effect size of $d = 0.53$. It showed that the cartoon lesson ($M = 3.93$, $SD = 1.43$) made students feel the instructor was more engaging than the original lesson ($M = 3.16$, $SD = 1.47$). In addition, there was also a main effect of voice, showing that the human voice ($M = 3.27$, $SD = 1.42$) was less engaging than the synthesized voice ($M = 3.82$, $SD = 1.52$), $F(1, 197) = 7.26$, $p <$

0.01, with a small to medium effect size $d = 0.37$. There was no significant interaction, $F(1, 197) = 0.88$, $p = 0.35$.

Finally, for the ratings of how credible the instructor seemed, a 2x2 ANOVA did not find any significant effect for lesson type $F(1, 197) = 2.91$, $p = 0.09$; voice type, $F(1, 197) = 2.01$, $p = 0.16$; or interaction, $F(1, 197) = 0.00$, $p = 0.99$.

In conclusion, in line with hypothesis 3, there is evidence from three of the four API scales that the cartoon lesson leads to stronger partnership connections between students and instructors than the original lesson. The lack of interactions indicates that no matter whether the narrator's voice was generated from the computer or recorded by an actual human, the cartoon lesson could help students build a stronger partnership connection with instructors.

VI. General Discussion

A. Empirical Contributions

In line with hypothesis 1, the results of Experiment 1, Experiment 2, and the combination of the data from Experiment 1 and 2 show that the cartoon group scored significantly higher than the original group on the transfer test but not the retention test. Since the transfer test measured a deeper understanding of the learning materials, we were more interested in this result. This is the major empirical contribution of the study.

In line with hypothesis 2, Experiment 1 and the data combined from Experiment 1 and Experiment 2, showed that learners in the cartoon group tended to feel more positive (i.e., happier and more content) than those in the original group, while the original lesson group did not make learners feel more negative than the cartoon lesson group. There were no interactions between the lesson design types (i.e., cartoon lesson vs. original narrated

slides lesson) and the voice types (i.e., synthesized voice vs. human voice), suggesting that the patterns with the synthesized voice group were not different from those with the human voice group. Overall, we conclude that the cartoon lesson leads to more positive emotions for students than the original lesson (based on Experiment 1 and the combined data).

In line with the third hypothesis, the results of Experiment 2 and the combined data from Experiments 1 and 2 showed that learners in the cartoon group rated the instructor higher than those in the original group on facilitating learning, being human-like, and being engaging, but not on being more credible. Moreover, the combined data from Experiment 1 and 2 also showed differences in the ratings of three of four subscales between the two conditions. There were no interactions between the lesson type and voice type for all the subscales, indicating that the effects of emotional design were not different for human or synthesized voices.

Concerning the role of voice, we did not find consistent evidence that students learned better with a human voice than with a synthesized voice. The positive effects of emotional design were not moderated by the type of instructor's voice.

B. Theoretical Implications

This study supports the *positivity principle*, which states that people learn better when the key elements in a lesson are designed to elicit positive affect (Horovitz & Mayer, 2021; Lawson & Mayer, in press; Lawson et al., 2021a, 2021b, in press). Specifically, this study provides supporting evidence for the Cognitive-Affective Model of Learning with Media by showing how emotional design can affect steps in the model. In the first link, positive emotional design in the form of a cartoon lesson impacts the learner's affective processing (as indicated by ratings of felt positive emotion during learning). In the second

link, positive emotional design in the form of a cartoon lesson impacts the learner's social processing (as indicated by ratings of the instructor's characteristics). In the final link, positive emotional design in the form of a cartoon lesson causes learners to engage in deeper cognitive processing resulting in better learning outcomes (as indicated by transfer test scores).

In addition, the lack of evidence for the *voice principle*–the idea that people learn better from a human voice than a machine voice–suggests the possibility that current text-to-speech synthesizers have advanced enough to no longer create negative social cues in a lesson (Craig & Schroeder, 2017, 2018; Mayer, 2021).

C. Practical Implications

Based on the positivity principle, this study suggests the need to consider the role of emotional design in multimedia lessons. Specifically, based on this study, we recommend that the key visual elements in multimedia lessons should be rendered in ways that induce positive emotion in learners. When the goal is to prime students' deeper processing of the learning materials, it might be helpful to design a cartoon lesson by making the visual elements not only include the traditional features of the emotional design (i.e., warm color, rounded shapes, and personalized characters with facial expressions) but also relate to the theme of the lesson. Rendering relevant visual elements as cartoon-like characters based on traditional emotional design principles can be considered as *informative emotional design*, that is, emotional design that is applied to instructional relevant content. This approach is a practical tool for facilitative learning in multimedia learning environments.

A secondary practical implication is that a synthesized voice designed to express positive emotion can be just as effective as a human voice.

D. Limitations and Future Directions

There are several limitations to this study. First, the cognitive and motivational processing and cognitive load surveys might not be as reliable. Some of the items might be confusing for students and make them misunderstand the meanings of the questions, which might influence the results of the ratings. Therefore, researchers should find other more valid and reliable surveys to measure students' motivations and cognitive load for future studies.

Second, the posttest cannot be considered as a delayed test since we only had three surveys with 32 items in total between the slideshow lesson and the posttest. Such a short interval was not sufficient to measure the long-term effects on learning. Therefore, the future study should make the interval between the learning session and posttest to be longer, such as several days or several weeks, to help us better understand the effects of the cartoon lesson (i.e., informative emotional design) in a longer term.

Third, in this study, we only compared the difference between the cartoon lesson (i.e., with informative emotional design) and the original narrated slides lesson (i.e., without informative emotional design). It is still not clear whether the effects we found in this study were caused by traditional emotional design features (i.e., warm colors, rounded shapes, and anthropomorphism) or by applying these features by converting key visual elements into cartoon-like characters with facial expressions. Therefore, for future studies, it is necessary to replicate the study and add another condition of traditional emotional design not applied to key visual elements as the comparison.

In addition, future work is needed to ensure the generality of the current findings. In the present study, most of the participants were young female college students, so future

studies should broaden the participant base. Moreover, the lesson was relatively short (i.e., 2.5 minutes) and focused on an introductory-level science topic. Future studies can further explore the effect of cartoon learning with longer lessons about different topics such as history. Finally, future studies might want to further differentiate the design elements of a cartoon lesson in order to investigate which specific features lead to strong effects.

VII. Conclusion

Taken together, in these two experiments, we found that compared with the original narrated slides, using informative emotional design through cartoon features in the lesson could not only induce positive emotions in students and feelings of social partnership with the instructor but also effectively facilitate deeper learning as indicated by their transfer learning performance. This work has implications for theory and practice, as noted above.

References

1. Akamca, G. Ö., Ellez, A. M., & Hamurcu, H. (2009). Effects of computer aided concept cartoons on learning outcomes. *Procedia-Social and Behavioral Sciences, 1*(1), 296-301.

2. Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces: the impact of age-related craniofacial changes on social perception. *Psychological bulletin, 100*(1), 3-18.

3. Brom, C., Starkova, T., & D'Mello, S. K. (2018). How effective is emotional design? A meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning. *Educational Research Review, 25,* 100-119.

4. Clark, R. C., & Mayer, R. E. (2016). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning* (4th ed.). Pfeiffer/John Wiley & Sons.

5. Craig, S. D., & Schroeder, N. L. (2017). Reconsidering the voice effect when learning from a virtual human. *Computers & Education, 114,* 193-205.

6. Craig, S. D., & Schroeder, N. L. (2018). Text-to-speech software and learning: Investigating the relevancy of the voice effect. *Journal of Educational Computing Research, 57,* 1534-1548.

7. D'Mello, S. K., & Graesser, A. C. (2012). AutoTutor and Affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems, 2*, 22–39.

8. Eker, C., & Karadeniz, O. (2014). The effects of educational practice with cartoons on learning outcomes. *International Journal of Humanities and Social Science, 4*(14), 223-234.

9. Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: The effect of red on performance attainment. *Journal of Experimental Psychology: General, 136*, 154–168.

10. Fiorella, L. (2022). The embodiment principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3nd ed; 286-295). New York: Cambridge University Press.

11. Fiorella, L., & Mayer, R. E. (2022). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3nd ed; 277-285). New York: Cambridge University Press.

12. Gamage, S. (2019). Cartoons as an authentic supplementary teaching tool in nglish as a second language classrooms. *Advances in Language and Literary Studies, 10*(1), 107-116.

13. Gong, S., Shangguan, C., Zhai, K., & Guo, Y. (2017). The effects of emotional design on multimedia learning. *Acta Psychologica Sinica, 49*(6), 771-782.

14. Gorn, G. J., Chattopadhyay, A., Yi, T., & Dahl, D. W. (1997). Effects of color as an executional cue in advertising: They're in the shade. *Management science, 43*(10), 1387-1400.

15. Hatfield, E., Bensman, L., Thornton, P. D., & Rapson, R. L. (2014). New perspectives on emotional contagion: A review of classic and recent research on facial mimicry and contagion. *Interpersona: An International Journal on Personal Relationships, 8*(2).

16. Hatfield, E., Carpenter, M., & Rapson, R. L. (2014). Emotional contagion as a precursor to collective emotions. *Collective emotions: Perspectives from psychology, philosophy, and sociology*, 108-122.

17. Horovitz, T., & Mayer, R. E. (2021). Learning with human and virtual instructors who display happy or bored emotions in video lectures. *Computers in Human Behavior, 119*, 106724.

18. Johnson, C. I., & Priest, H. A. (2022). 34. The Feedback Principle in Multimedia Learning. *The Cambridge handbook of multimedia learning*, pp. 403-417.

19. Keogh, B., & Naylor, S. (1999). Concept cartoons, teaching and learning in science: an evaluation. *International journal of science education*, *21*(4), 431-446.

20. Kumi, R., Conway, C. M., Limayem, M., & Goyal, S. (2012). Research article learning in color: how color and affect influence learning outcomes. *IEEE transactions on professional communication, 56*(1), 2-15.

21. Lawson, A. P., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (2021a). Do learners recognize and relate to the emotions displayed by virtual instructors? *International Journal of Artificial Intelligence in Education, 31,* 134-153.

22. Lawson, A., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (2021b). The positivity principle: Do positive instructors improve learning from instruction video lectures? *Educational Technology Research and Development, 69*, 3101-3129.

23. Lawson, A. P., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (2021). Recognizing the emotional state of human and virtual instructors. *Computers in Human Behavior, 114,* 106554.

24. Lawson, A. P., & Mayer, R. E. (2021). The power of voice to convey emotion in multimedia instructional messages. *International Journal of Artificial Intelligence in Education*.

25. Leppink, J., Paas, F., Van der Vleuten, C. P., Van Gog, T., & Van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior research methods, 45*(4), 1058-1072.

26. Levin, J. R., McCormick, C. B., Miller, G. E., Berry, J. K., & Pressley, M. (1982). Mnemonic versus nonmnemonic vocabulary-learning strategies for children. *American Educational Research Journal, 19*(1), 121-136.

27. Loderer, K., Pekrun, R., & Plass, J. L. (2020). Emotional foundations of game-based learning. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of game-based learning* (pp. 111–151). The MIT Press.

28. Mayer, R. E. (2021). *Multimedia learning* (3rd ed). New York: Cambridge University Press.

29. Mayer, R. E. (2020). Searching for the role of emotions in e-learning. *Learning and Instruction, 70,* 101213.

30. Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction, 33*, 12-18.

31. McQuiggan, S. W., & Lester, J. C. (2007). Modeling and evaluating empathy in embodied companion agents. *International Journal of Human-Computer Studies, 65*, 348–360.

32. Miesler, L. l., Leder, H., & Herrmann, A. (2011). Isn't it Cute: An evolutionary perspective of Baby-Schema effects in visual product designs. *International Journal of Design, 5*(3), 17–30

33. Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*, 309-326.

34. Naylor, S., & Keogh, B. (2013). Concept cartoons: what have we learnt?. *Journal of Turkish Science Education, 10*(1), 3-11.

35. Özdoğru, A. A., & McMorris, R. F. (2013). Humorous cartoons in college textbooks: Student perceptions and learning. *Humor, 26*(1), 135-154.

36. Park, B., Knörzer, L., Plass, J. L., & Brünken, R. (2015). Emotional design and positive emotions in multimedia learning: An eye-tracking study on the use of anthropomorphisms. *Computers & Education, 86*, 30-42.

37. Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction, 29*, 128-140.

38. Plass, J. L., Homer, B. D., MacNamara, A., Ober, T., Rose, M. C., Pawar, S., ... & Olsen, A. (2020). Emotional design for digital games for learning: The effect of expression, color, shape, and dimensionality on the affective quality of game characters. *Learning and instruction, 70*, 101194.

39. Plass, J. L., & Kaplan, U. (2016). Emotional design in digital media for learning. In S. Y. Tettegah & M. P. McCreery (Eds.), *Emotions, Technology, and Learning* (pp. 131–161). London: Academic Press.

40. Rossetto, M., & Chiera-Macchia, A. (2011). 'Visual learning is the best learning – it lets you be creative while learning' Exploring ways to begin guided writing in second language learning through the use of comics. *Babel, 45*(2/3), 35-40.

41. Scherer, K. R., & Coutinho, E. (2013). How music creates emotion: A multifactorial process approach. In T. Cochrane, B. Fantini, & K. R. Scherer (Eds.). *The emotional power of music: Multidisciplinary perspectives on musical arousal, expression, and social control* (pp. 121–145). Oxford, UK: Oxford University Press.

42. Schmidt, S. R., & Williams, A. R. (2001). Memory for humorous cartoons. *Memory & Cognition, 29*(2), 305-311.

43. Stephenson, P., & Warwick, P. (2002). Using concept cartoons to support progression in students' understanding of light. *Physics education, 37*(2), 135-141.

44. Takahashi, M., & Inoue, T. (2009). The effects of humor on memory for non-sensical pictures. *Acta Psychologica, 132*(1), 80-84.

45. Um, E. "R.", Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). Emotional design in multimedia learning. *Journal of Educational Psychology, 104*(2), 485–498.

46. Uzun, A. M., & Yıldırım, Z. (2018). Exploring the effect of using different levels of emotional design features in multimedia science learning. *Computers & Education, 119,* 112-128.

47. Wong, R. M., & Adesope, O. O. (2021). Meta-analysis of emotional designs in multimedia learning: A replication and extension study. *Educational Psychology Review, 33*(2), 357-385.

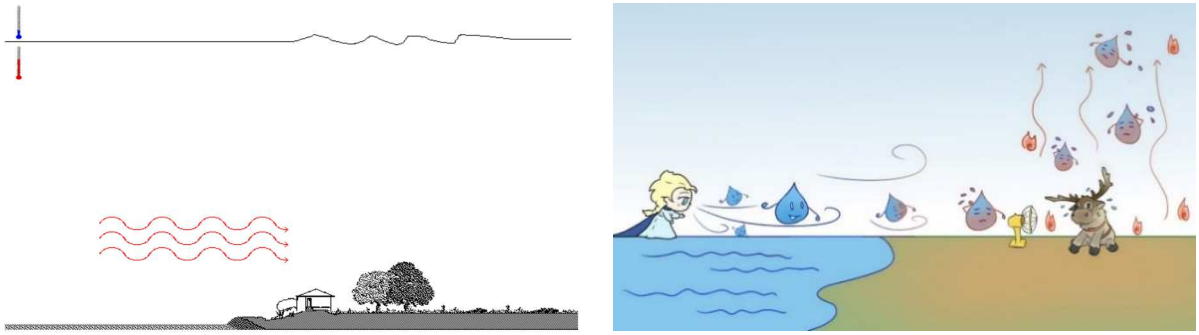
48. Wyk, M. M. V. (2011). The use of cartoons as a teaching tool to enhance student learning in economics education. *Journal of Social Sciences, 26*(2), 117-130.

49. Yaman, H. (2010). Cartoons as a Teaching Tool: A Research on Turkish Language Grammar Teaching. *Educational Sciences: Theory and Practice, 10*(2), 1231-1242.

50. Zebrowitz, L. A., Luevano, V. X., Bronstad, P. M., & Aharon, I. (2009). Neural activation to babyfaced men matches activation to babies. *Social Neuroscience, 4*(1), 1-10.

51. Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and personality psychology compass, 2*(3), 1497-1517.

Figure 1. Example Slides from the Original and Cartoon Lessons



"Cool moist air moves over a warmer surface and becomes heated."

Note. The example of the original narrated slideshow is on the left side of Figure 1, and the example of the cartoon slideshow is on the right side of Figure 1.

Figure 2. Abbreviated Version of the Cognitive-Affective Model of Learning with Media

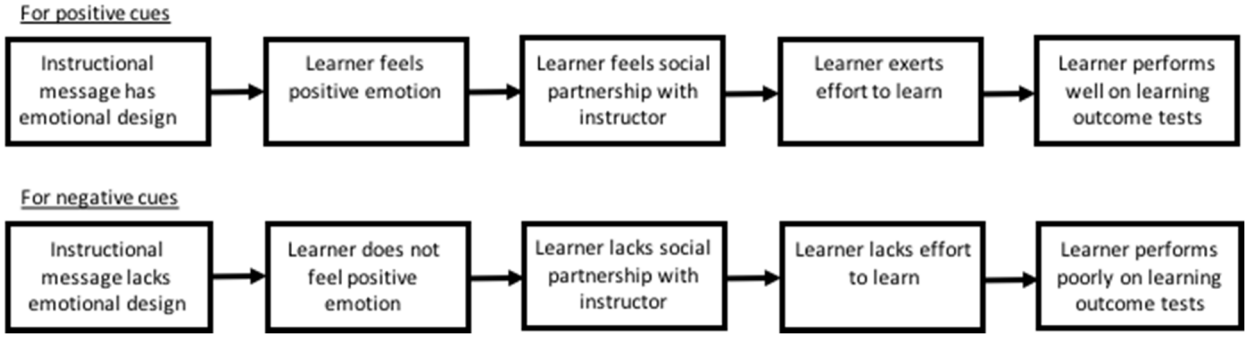


Table 1
Script of Lighting Lesson Used in Experiments 1 and 2

Slide	Script
1	Cool moist air moves over a warmer surface and becomes heated.
2	Warmed moist air near the earth's surface rises rapidly.
3	As the air in this updraft cools, water vapor condenses into water droplets and forms a cloud.
4	The cloud's top extends above the freezing level, so the upper portion of the cloud is composed of tiny ice crystals.
5	Eventually, the water droplets and ice crystals become too large to be suspended by the updrafts.
6	As raindrops and ice crystals fall through the cloud, they drag some of the air in the cloud downward, producing downdrafts.
7	When downdrafts strike the ground, they spread out in all directions, producing the gusts of cool wind people feel just before the start of the rain.
8	Within the cloud, the rising and falling air currents cause electrical charges to build.
9	The charge results from the collision of the cloud's rising water droplets against heavier, falling pieces of ice.
10	The negatively charged particles fall to the middle of the cloud, and most of the positively charged particles rise to the top.
11	A stepped leader of negative charges moves downward in a series of steps. It nears the ground.
12	A positively charged leader travels up from such objects as trees and buildings.
13	The two leaders generally meet about 165-feet above the ground.
14	Negatively charged particles then rush from the cloud to the ground along the path created by the leaders. It is not very bright.
15	As the leader stroke nears the ground, it induces an opposite charge, so positively charged particles from the ground rush upward along the same path.
16	This upward motion of the current is the return stroke. It produces the bright light that people notice as a flash of lightning.

Table 2
Test Items Used in Experiments 1 and 2

Retention Test :	Please write down an explanation of how lightning works.
Transfer Test:	What could you do to decrease the intensity of lightning?
	Suppose you see clouds in the sky, but no lightning. Why not?
	What does air temperature have to do with lightning?
	What causes lightning?

Table 3

Means and Standard Deviations of Posttest Scores by Group in Experiments 1, 2, and 1&2

Group	Retention test score *M*	*SD*	Transfer test score *M*	*SD*
EXPERIMENT 1 (Synthesized Voice)				
Cartoon Lesson	4.80	4.21	**3.37**	2.33
Narrated Slides Lesson	4.28	3.53	2.36*	1.72
EXPERIMENT 2 (Human Voice)				
Cartoon Lesson	5.58	3.48	**3.74**	2.57
Narrated Slides Lesson	4.48	3.59	2.56*	2.26
EXPERIMENT 1&2 (Supplemental Analysis)				
Cartoon Lesson	5.19	3.86	**3.55**	2.45
Narrated Slides Lesson	4.38	3.54	2.46*	2.00

Note. Asterisk(*) represents significant difference from the **bolded** condition.

Table 4
Means and Standard Deviations of Learners' Felt Emotion Ratings for Four Emotions by Group in Experiments 1, 2, and 1&2

	Happy		Content		Frustrated		Bored	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
EXPERIMENT 1 (Synthesized Voice)								
Cartoon Lesson	**3.20**	0.85	3.22	0.92	3.08	1.09	2.29	1.19
Narrated Slides Lesson	2.68*	1.10	2.80	1.05	3.30	1.40	2.26	1.24
EXPERIMENT 2 (Human Voice)								
Cartoon Lesson	3.08	1.05	3.22	1.09	3.36	1.06	2.24	1.20
Narrated Slides Lesson	2.78	0.84	3.00	1.01	3.34	1.06	2.00	1.01
EXPERIMENT 1&2 (Supplemental Analysis)								
Cartoon Lesson	**3.14**	0.95	**3.22**	1.01	3.22	1.08	2.27	1.19
Narrated Slides Lesson	2.73*	0.97	2.90*	1.03	3.32	1.24	2.13	1.13

Note. Asterisk(*) represents a significant difference from the **bolded** condition.

Table 5
Means and Standard Deviations on Four Subscales of the API by Group in Experiments 1, 2, and 1&2

	Facilitative Learning		Credible		Human-like		Engaging	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
EXPERIMENT 1 (Synthesized Voice)								
Cartoon Lesson	4.02	1.19	5.11	1.00	3.10	1.49	**4.11**	1.35
Narrated Slides Lesson	3.61	1.50	4.86	1.21	2.83	1.53	3.52*	1.65
EXPERIMENT 2 (Human Voice)								
Cartoon Lesson	4.29	1.34	5.32	0.98	**3.87**	1.53	**3.76**	1.50
Narrated Slides Lesson	3.82	1.16	5.07	1.02	3.17*	1.28	2.79*	1.16
EXPERIMENT 1&2 (Supplemental Analysis)								
Cartoon Lesson	**4.15**	1.27	5.21	0.99	**3.48**	1.55	**3.93**	1.43
Narrated Slides Lesson	3.71*	1.34	4.86	1.12	3.00*	1.41	3.16*	1.47

Note. Asterisk(*) represents a significant difference from the bolded condition.

Table 6
Mean Ratings and Standard Deviations on Cognitive and motivational Processing by Group in Experiments 1, 2, and 1&2

Group	Motivated		Difficult		Effort to Understand		Enjoyment		More Lesson	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
EXPERIMENT 1 (Synthesized Voice)										
Cartoon Lesson	3.12	1.22	2.63	1.02	3.41	0.90	**3.47**	0.90	2.82	1.05
Narrated Slides Lesson	2.84	1.33	2.52	1.22	3.08	1.19	2.90*	1.15	2.50	1.20
EXPERIMENT 2 (Human Voice)										
Cartoon Lesson	3.38	1.28	2.37	1.00	3.14	1.14	3.50	0.89	3.10	1.09
Narrated Slides Lesson	3.08	1.23	2.51	0.95	3.34	1.00	3.38	0.88	3.04	1.21
EXPERIMENT 1&2 (Supplemental Analysis)										
Cartoon Lesson	3.25	1.25	2.54	1.03	3.28	1.03	**3.49**	0.89	2.96	1.08
Narrated Slides Lesson	2.96	1.28	2.53	1.08	3.21	1.01	3.14*	1.05	2.77	1.23

Note. Asterisk(*) represents a significant difference from the **bolded** condition.