# Watching a Fun Artistic Video about Simpson's Paradox

**Carla van de Sande**
Arizona State University
Tempe, Arizona, United States

**Abstract.** This study explores how university students respond to a brief artistic and humorous educational video covering a statistics topic. Major study activities included a background knowledge questionnaire, watching the video, completing a storyboard by identifying images from three scenes in the video, explaining the topic using the recalled images, and, finally, answering questions about the topic that were presented in a different context. The images shown to participants differed according to conceptual and detail accuracy. The results suggest that the details of the animation were more salient in participants' memory than the concepts. This study introduces a novel way of probing recognition of instructional content via visual memory rather than through written statements.

**Keywords:** artistic instructional content; humorous instructional videos; statistics learning; visual recall.

## 1. Introduction

It is well established that some students are better at processing images over words, and that some students prefer visual over verbal learning resources (Mayer & Massa, 2003). These resources were traditionally delivered in the form of static pictures, such as drawings or illustrations, as are often found in textbooks (Mayer, 1993). However, technology has now greatly opened up our ability to expand the repertoire of visual learning resources by allowing us to incorporate dynamic images or animations into instructional explanations. Indeed, many sites on the Internet cater to visual learning styles and preferences by delivering short instructional videos on demand. Students can easily search for, find, and watch both serious and humorous short videos on numerous topics that have been published on YouTube channels and at other locations for public use (Burgess & Green, 2018). This study explores how students attend to a brief artistic and humorous educational video on a statistics topic, namely Simpson's Paradox. In particular, we look at whether the focus is on details or concepts in recognizing scenes from the video, the facets of their explanation of

the topic after watching the video, and their ability to extend their understanding to a novel representational context.

The use of instructional videos to deliver content is a form of multimedia learning since it incorporates both pictures and words (Mayer, 2009; 2014). Research on multimedia learning has identified several principles of instructional design that support successful learning outcomes by facilitating cognitive processing, such as using cues or other signals to highlight the most important information and organizing it into a coherent representation (Betrancourt, 2005). Watching an instructor's hand draw as an instructional explanation unfolds is one way of motivating students to make sense of material (Fiorella, L. & Mayer, R. E., 2016).

However, the diagrams and scenario used in past research on instructional videos were exclusively academic in nature. The goal was to communicate the material directly, rather than see if students could learn while watching a video that was meant to be both educational and entertaining. Some instructional videos are designed to capture and sustain engagement by presenting content with an artistic flair. These attempts to meld art with instructional content resonate with recent efforts to add Art into innovations into Science, Technology, Engineering, and Mathematics (STEM) education, resulting in what has become known as the STEAM (STEM + Art) initiative at http://stemtosteam.org/. The integration of art and design in K-20 education is one of the major objectives of the STEAM movement so instructional videos that use hand drawn illustrations and humor to explain topics are worth exploring. The role of humor has previously been investigated in the context of the inclusion of cartoons in textbooks (Bryant, Brown, Silberberg, & Elliott, 1981) and on the application of humor in college courses (Garner, 2012). Particularly relevant for this study is that the use of humorous examples in instructional lectures can enhance the recall of related information and concepts (Kaplan & Pascoe, 1977), and that entertaining animations accompanied by instructional scaffolding (such as discussions and interactive quizzes and experiments) can promote understanding and improve explanations of scientific topics (Barak, Ashkar, & Dori, 2011).

## 2. Research methodology

Students from the first author's online university mathematics courses were asked to volunteer for the study. A description of the study and request to participate was sent via e-mail to 84 former students, 68 of whom agreed to take part in the study. All participants were between the ages of 18 and 25. The entire study was conducted online. The sequence of activities was created using Articulate 360 and published out as a SCORM 1.2[1] and HTML5[2] to record the timing and responses for each participant. Figure 1 outlines the sequencing of the five major study activities that included completing a test of background knowledge, watching a video, filling in a storyboard, explaining the concept, and extending the concept to novel situations.
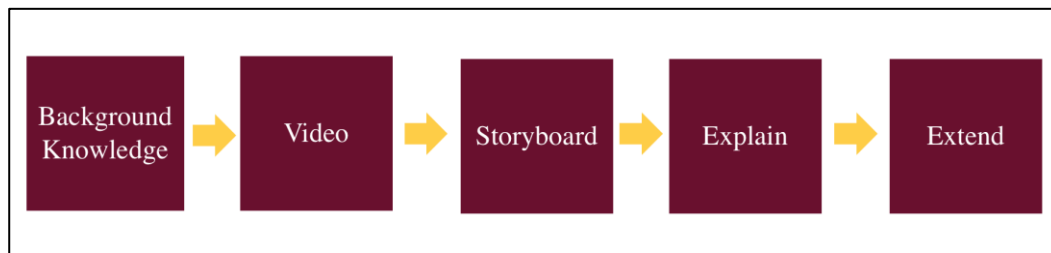
**Figure 1: Study activities**

### 2.1 Background knowledge

Background knowledge was established by asking participants to rate their knowledge of Simpson's Paradox on a scale of 1 (very low) to 5 (very high) and to indicate whether the following nine statements applied to them: "I have taken a course in statistics," "I know how to calculate percentages.", "I know what it means to have a confounding variable in an experiment.", "I know how to interpret a scatter plot.", "I know what it means for two variables to be correlated.", "I know what it means for two variables to be causally related.", "I know what it means to aggregate data across experimental groups.", "I know the definition of paradox.", and "I know how to interpret a table of values." Items were scored as 0 or 1 depending on whether a participant replied in the negative or affirmative, respectively.

### 2.2 Video

After completing the background knowledge questionnaire, students watched a brief animation (2 minutes and 40 seconds in duration) on Simpson's Paradox. The animation was adapted from a video published by MinutePhysics at http://www.minutephysics.com/, an educational YouTube channel created by Henry Reich. These short videos were designed to artistically explain physics and related topics using time-lapsed drawing.

We chose Simpson's Paradox as the topic for the study because it is relatively obscure and because it is a topic that can be explained using familiar ideas and basic calculations. Our intent was to use a topic that would increase the likelihood of participants being unfamiliar with it when they volunteered, but at the same time to have the requisite background knowledge to understand the explanation. Simpson's Paradox is a statistics phenomenon in which a trend appears in several different groups of data but disappears or reverses when these groups are combined. The video humorously describes the paradox through a comparison of recovery and survival outcomes from a group of cats and a group of humans who are being treated for a disease, and shows how two opposite conclusions can be drawn depending on how the data is partitioned. The message is that causal factors are a critical consideration in experimental design. In particular, controlled experiments need to exclude causally related factors that could otherwise influence the results, and uncontrolled experiments need to take causal factors into account.

## 2.3 Storyboard

As shown in Figure 2, the storyboard activity involved the recognition of three scenes from the video depicting the example of Simpson's Paradox, a causal explanation for one interpretation of the results, and a causal explanation for the opposite interpretation of the results. The first and last scenes from the video were provided to help participants bookend their recognition and recall efforts.
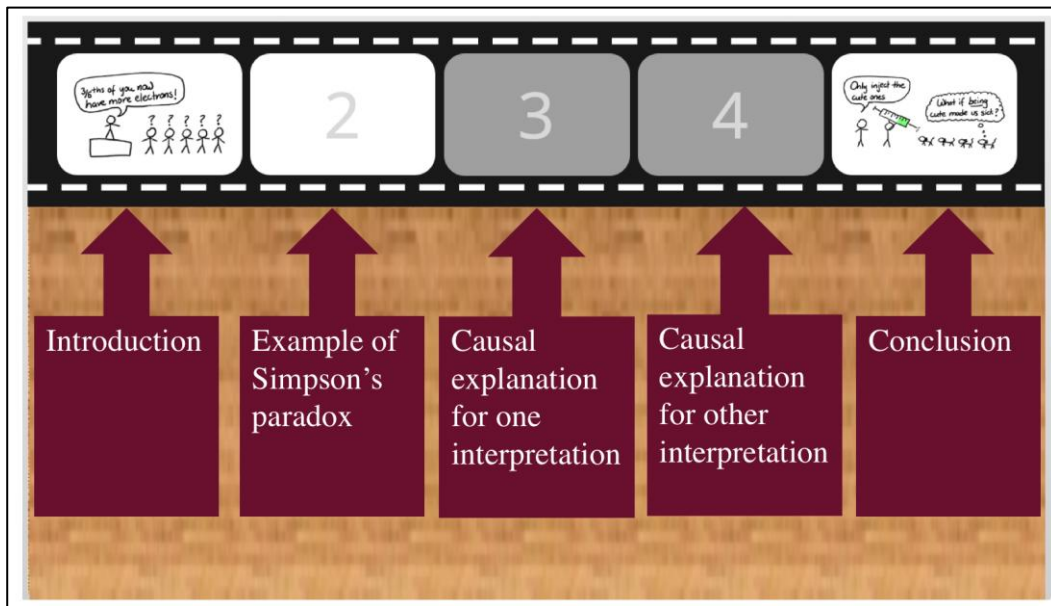


**Figure 2: Elements of storyboard activity**

For each of the three missing scenes, participants were asked to select from four possible images, as shown in Figure 3. These images were designed to differ by conceptual and detail accuracy: an image that was taken from the video and was therefore conceptually accurate and shared the exact details, an image that was conceptually analogous with the one in the video but differed by surface features, an image that was conceptually incorrect but mirrored the details from the one in the video, and an image that had different surface features than the one in the video and was also conceptually incorrect (see Figure 4). Conceptual inaccuracy was achieved by drawing unjustifiable conclusions, and surface feature inaccuracy was achieved by interchanging the results of cats and humans. For each scene, participants could drag and drop the image they recognized from the video onto the storyboard strip before moving on to the next blank scene.
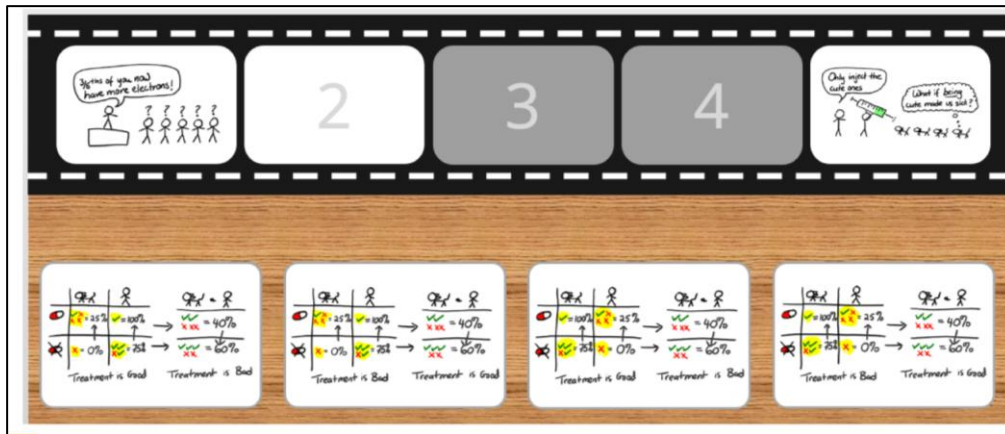
**Figure 3: Storyboard activity**

| | | Details | |
|---|---|---|---|
| | | **Accurate** | **Inaccurate** |
| **Concept** | **Accurate** |  |  |
| | **Inaccurate** |  |  |

**Figure 4: Storyboard images differed by conceptual and detail accuracy**

### 2.4 Explain

Once participants completed the storyboard activity, the storyboard strip was completely populated with their choices, and they were asked to explain Simpson's Paradox. The explanations were analyzed for whether they contained the definition, an example, and the role of causality in Simpson's Paradox. Two people independently coded the explanations with an inter-rater reliability of 93%, and all differences were resolved through discussion.

### 2.5 Extend

The last activity in the study was designed to determine whether participants could extend their understanding of Simpson's Paradox to novel situations. This type of activity is considered to measure whether learning has taken place to the extent that it can enhance related performance in another context (Perkins & Salomon, 1992). The transfer test contained two items that assessed whether participants could recognize the presence and absence of Simpson's Paradox based on an alternative representation. The video exclusively used tables of data to depict Simpson's Paradox. Therefore, the first item (Figure 5 top) asked participants to select which of four scatter plots illustrated Simpson's Paradox.

There were two examples that were correct, and two that were not. The second item (Figure 5 bottom) asked participants to select which of four scatter plots could be used to draw an unambiguous conclusion (i.e., the absence of Simpson's Paradox). In this case, there was only one correct response. In both cases, the choices were presented simultaneously in a random order across the bottom of the screen.  A score of ¼ was awarded for each correct identification (either as an example or as a non-example) so that scores on each item ranged from 0 to 1.
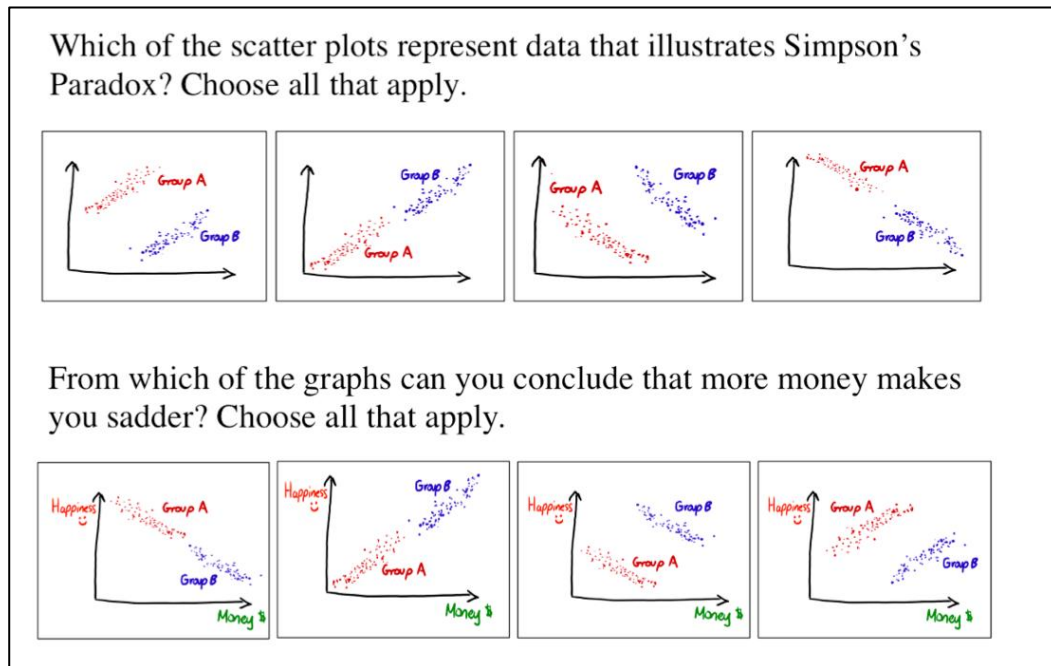


**Figure 5: Extension activity**

## 3. Results

The results of this study are presented according to each study activity. The results for the background knowledge activity are based on the responses to the questionnaire on prior knowledge of Simpson's Paradox and related concepts. The results for the storyboard activity are based on the selections of storyboard images according to conceptual and detail accuracy for each of the three scenes that were shown to participants. The results of the explain activity are based on elements that were present in participants' retelling of Simpson's Paradox after completing the storyboard activity. Finally, the results from the extend activity are based on the responses to items that either depicted or failed to depict Simpson's Paradox in a context that differed from the one presented in the video.

### 3.1 Background knowledge

First, none of the participants reported having knowledge of Simpson's Paradox. Therefore, this topic was unfamiliar to all of the participants prior to watching the video. At the same time, participants reported an understanding of the basic information and calculations that are necessary to understand the explanation of

Simpson's Paradox and answer the items in the Extend activity. Table 1 shows the percentage of students who responded "Yes" to the background knowledge items. Items 1-7 involved familiarity with major ideas that were used in the video as part of the explanation of Simpson's Paradox, and items 8-9 involved knowledge of skills that were requisite to answering the extension questions after watching the video.

**Table 1: Percentage of participants who responded affirmatively**

| Item | "Yes" |
|---|---|
| 1. I have taken a course in statistics. | 29% |
| 2. I know what it means for two variables to be causally related. | 68% |
| 3. I know what it means for two variables to be correlated. | 79% |
| 4. I know what it means to have a confounding variable in an experiment. | 50% |
| 5. I know what it means to aggregate data across experimental groups. | 35% |
| 6. I know the definition of a paradox. | 76% |
| 7. I know how to calculate percentages. | 97% |
| 8. I know how to interpret a scatter plot. | 91% |
| 9. I know how to interpret a table of values. | 91% |

### 3.2 Storyboard

Figure 6 shows the percentage of responses for each scene by detail and concept accuracy. First, the majority of participants were able to recognize the correct image for each of the three scenes, and very few participants chose an image that was completely incorrect (with respect to both concept and detail). The most interesting finding was that participants favored images that had correct details and incorrect concepts over those that had incorrect details and correct concepts, and this was true for all three scenes.
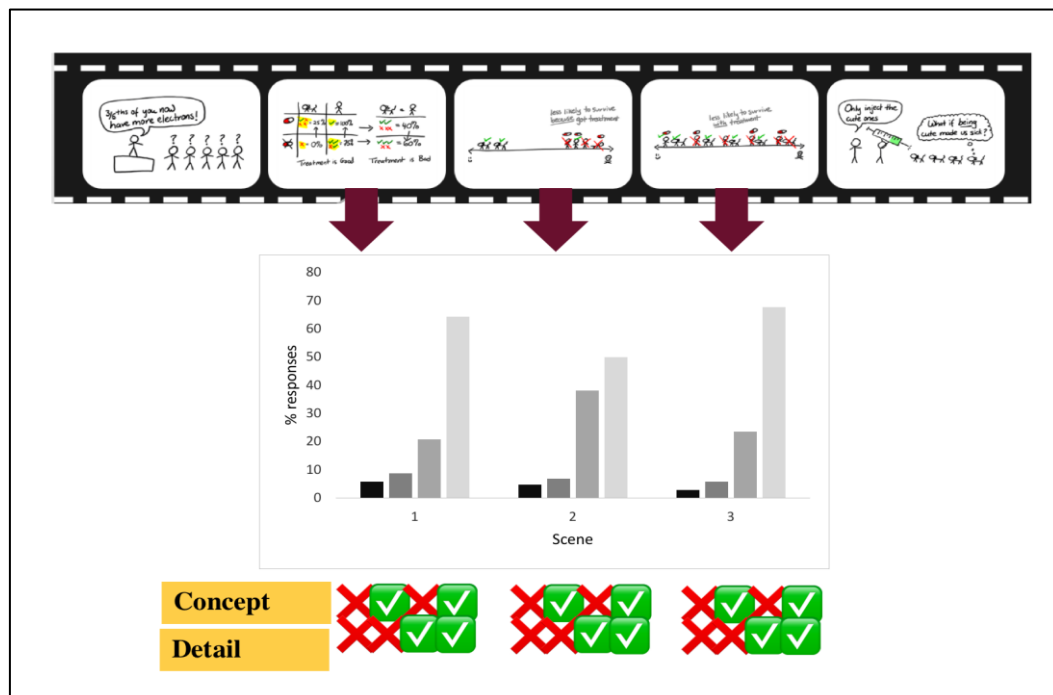


**Figure 6: Images identified by participants for each scene according to accuracy of concept and detail**

### 3.3 Explain

The majority of participants (68%) included a definition in their explanation, whereas many fewer (21%) used an example or discussed the role of causality (18%) in creating Simpson's Paradox. However, when these were present, they were very detailed and closely mirrored the explanation from the video. Table 2 shows a sample portion of participants' explanations from each of these categories.

**Table 2: Sample responses containing a definition, example, and role of causality**

|  | Definition | Example | Causality |
|---|---|---|---|
| Sample | "Simpson's Paradox is the paradox that a person can come to different conclusions from the same set of data depending on how the data is divided up. " | "For example, data may show that survival rates decline after a certain treatment for specific groups but when data is aggregated, the survival rate seems to have increased. " | "This may be due to issues in the environment from which subjects are gathered. Some subjects may have certain traits that make them more or less likely to respond to treatment in a certain way. The ideal in statistics is to minimize the effect of Simpson's paradox by either controlling the kind of subjects used or by altering treatment to minimize the effect of confounding variables." |

### 3.4 Extend

The participants who responded that they were unfamiliar with interpreting scatter plots in the background knowledge activity were excluded from these analyses. The first extension item required participants to identify illustrations of Simpson's Paradox from a scatter plot instead of from a table of values as was done in the video. One choice displayed a scenario in which two groups had data with a negative trend but, if aggregated, would have a positive trend. One choice displayed the opposite scenario in which the two groups had data with a positive trend but, if aggregated, would have a negative trend. The remaining two choices were non-examples of Simpson's Paradox in which the trend of the aggregated data was either negative or positive and matched the trend of each of the two groups.

**Table 3: Percentage of correct and incorrect identifications of examples and non-examples in first extension activity item**

| Identifications | Percentage |
|---|---|
| Identified both of the examples and both of the non-examples correctly | 41% |
| Failed to identify one of the two examples but correctly identified the two non-examples | 38% |
| Failed to identify both of the examples and instead identified one or both of the non-examples as examples | 18% |
| Identified both of the examples correctly but also identified both of the non-examples as examples | 3% |

As depicted in Table 3, the majority of participants (79%) were able to correctly identify both (41%) or at least one (38%) of the two scatter plots illustrating examples of Simpson's Paradox and distinguish these from non-examples. At the same time, however, 18% of the participants failed to apply the definition of Simpson's Paradox in this context and reversed the scenarios identifying Simpson's Paradox with those that did not. Finally, 3% of participants did not at all distinguish between examples and non-examples of the paradox in a scatter plot representation.

The second extension item required participants to identify whether the data as represented in a scatter plot would support a definite conclusion as opposed to illustrating Simpson's Paradox. In this case, there was only one correct response. As shown in Table 4, 35% of the participants correctly identified the scatter plot from which Simpson's Paradox was absent. However, 44% of the participants did not detect the presence of Simpson's Paradox in at least one additional scenario. Of these, the most common error was to select the scatter plot in which the individual groups displayed the trend that would support the conclusion but aggregating the data over both groups would not. Finally, 20% of the participants failed to identify the scatter plot from which the conclusion could be unambiguously drawn and instead chose one or more of the other scenarios in which the group trends differed from the trend that would be present if the groups were aggregated.

**Table 4: Percentage of correct and incorrect identifications of examples and non-examples in second extension activity item**

| Identifications | Percentage |
|---|---|
| Identified the one example and all three non-examples | 35% |
| Identified the one example but also identified at least one non-example | 44% |
| Failed to identify the example and identified at least one non-example | 20% |

## 4. Discussion

In addition to exploring learning from brief entertaining animations such as those commonly found online, this study introduces a novel way of probing recognition of instructional content. Instead of asking participants to identify whether or not phrases from the narrative were present in the instructional content, we asked participants to select images from the video. Although visual memory when tested on its own is very reliable (Wolfe, 1998), we were asking participants to listen to an explanation while attending to the animation. Our results suggest that the details of the animation were more salient in their memory than the concepts since they were more likely to select scenes that mirrored the scenes in the video but were conceptually incorrect than vice-versa. This opens up questions about whether the revision of details in instructional videos to be humorous (e.g., replacing one group of people with a group of cats) distracts from the coherence of the core content and violates the Coherence Principle of multimedia instructional design (Mayer, 1999). Other research has found that interspersing or proceeding instructional animations with on topic but conceptually irrelevant seductive details can harm student understanding (Mayer, Heiser, & Lonn, 2001; Rey, 2012) but not whether humorous details that are intrinsically relevant to the concept explained by the video have a beneficial or detrimental effect on student learning. Future studies should compare retention and transfer from instructional videos in which the content is designed to be humorous versus not.

### 4.1 Limitations

The impact of this study is limited by the choice of the materials, the design of the study activities, the sample, and the experimental setting. First, we only explored a single video about a single topic. Therefore, it is unclear which property of the video led to the effects that we observed. However, we deliberately chose the video because it is representative of a genre of online instructional videos that employ humor and fast-paced hand-drawn animations to combine entertainment with educational merit. Second, the study activities were limited to short term recognition, retention, and transfer. The effects of the animations on retention and understanding might vary after students are given time to either absorb or forget the material. Third, the participants in our sample were college students who all had no previous knowledge of the topic. Other studies have found that prior knowledge can play a role in learning outcomes in that learners with low versus high prior knowledge process instructional materials in different ways (Fiorella & Mayer, 2016). Fourth, participants performed all of the study activities within an online environment. In particular, participants' explanations may have differed if they were delivered verbally to an actual person. Additional work is needed to determine how infusing artistically inspired instructional videos with humor affects learning and motivation.

## 5. Conclusion

Contrary to what many students may think (Stodolsky, Salk, & Glaessner, 1991), STEM content does not have to be dry and boring. Instead, it can be presented in a whimsical manner that may engage students in ways that differ from a

straightforward presentation of the material. This study paves the way for research on the impact of instruction and curricular materials that exemplify the STEAM initiative at the university level. The role of art in STEM instructional design is worth exploring further as more and more videos and visual materials are developed for use both in and outside of formal educational settings. Perhaps making art part of the picture can provide innovative ways to make learning STEM topics more entertaining and engaging for students at all levels.

## References

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. doi.org/10.1016/j.compedu.2010.10.025

Betrancourt, M. (2005). *The animation and interactivity principles in multimedia learning.* In R. E. Mayer (Ed.), The Cambridge handbook of multimedia learning (pp. 287–296). New York: Cambridge University Press. doi.org/10.1017%2Fcbo9780511816819.019

Bryant, J., Brown D., Silberberg, A. R., & Elliott, S. M. (1981). Effects of humorous illustrations in college textbooks. *Human Communication Research, 8(1),* 43–57. doi.org/10.1111%2Fj.1468-2958.1981.tb00655.x

Burgess, J., & Green, J. (2018). *YouTube: Online video and participatory culture*. John Wiley & Sons. doi.org/10.7146%2Fmediekultur.v25i47.2132

Fiorella, L. & Mayer, R. E. (2016). Effects of observing the instructor draw diagrams on learning from multimedia messages. *Journal of Educational Psychology, 108*(4), 528 – 546. doi.org/10.1037%2Fedu0000065

Garner, R. L. (2006). Humor in pedagogy: How ha-ha can lead to aha! *College Teaching, 54(1),* 177-180. doi.org/10.3200%2Fctch.54.1.177-180

Kaplan, R. M., & Pascoe, G. C. (1977). Humorous lectures and humorous examples: Some effects upon comprehension and retention. *Journal of Educational Psychology, 69*(1), 61–65. doi.org/10.1037%2F0022-0663.69.1.61

Mayer, R. E. (1993). Illustrations that instruct. In R. Glaser (Ed.), *Advances in instructional psychology,* Vol. 4, pp. 253 – 284). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Mayer, R. E. (1999). Research-based principles for the design of instructional messages: The case of multimedia explanations. *Document Design, 1,* 7–20. doi.org/10.1075%2Fdd.1.1.02may

Mayer, R. E. (2009). Multimedia learning (2nd ed.). New York, NY: Cambridge University Press.

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), The Cambridge handbook ofmultimedia learning (2nd ed., pp. 43–71). New York, NY: Cambridge University Press. doi.org/10.1017%2Fcbo9781139547369.005

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(1),* 187-198. doi.org/10.1037%2F0022-0663.93.1.187

Mayer, R. E. & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology, 95*(4), 833 – 846. doi.org/10.1037%2F0022-0663.95.4.833

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. Contribution to the International Encyclopedia of Education (2nd ed.). Oxford, England: Pergamon Press.

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review, 7(3),* 216–237. doi.org/10.1016%2Fj.edurev.2012.05.003

Stodolsky, S. S., Salk, S., & Glaessner, B. (1991). Student views about learning math and social studies. *American educational research journal*, *28*(1), 89-116. doi.org/10.3102%2F00028312028001089

Wolfe, J. M. (1998). Visual memory: What do you know about what you saw? *Current Biology, 8*(9), R303 – R 304. doi.org/10.1016%2Fs0960-9822%2898%2970192-7