The similarities and differences in modeling videos between cueing and eye movement modeling examples

Maurice Potuijt 6591566

Universiteit Utrecht 2019-2020

Supervisor 1: Tim van Marlen

Supervisor 2: Ellen Kok

Words: 5337

Abstract

Attention guiding is an instructional method often used in multimedia learning. One form of attention guiding is called cueing This method uses a cue (e.g. pointing arrow, highlight) to guide a learner's attention to a certain key visual element during a video to enhance learning. A recent development in the cueing department has been the use of Eye Movement Modeling Examples (EMME). EMME have been shown to find mixed results, much like Cueing Modeling Examples (CME). This brought to question whether these two methods are generalizable. To investigate this, an experiment was conducted with 46 Dutch middle school students learning geometry problems using the F-rule and Z-rule. The objective was to see how an EMME, CME and non-visual guided modeling example compare in learning outcome, mental effort and instructional quality. This study found no differences between the conditions. One reason this could have been was because the sample consisted of relatively high prior knowledge individuals. Findings were further discussed in terms of theoretical and practical implications.

Keywords: Attention guiding, eye movement modeling examples, cueing.

The similarities and differences in modeling videos between cueing and eye movement modeling examples

Since the rise of digital video platforms last decade, interest in effective instructional videos has substantially increased within education and training (Van Gog & Rummel, 2010). These instructional videos utilize their multimedia properties to provide example-based instruction where an instructor shows how a certain task is carried out. Example-based instructional videos are shown to be quite useful for garnering new knowledge and skills (Renkl, 2014; Van Gog & Rummel, 2010). However, the way these videos are designed affect their effectiveness (Fiorella & Mayer, 2018; Van Gog, Rummel and Renkl, 2019), warranting research to find improvements for this new digital age instructional format.

Attention Guiding

One problem instructional video face is that information can disappear before the learner has had enough time to process it. For example, when watching a video animation explaining the workings of a jet engine the instructor verbally describes certain parts, but before the learner knows where the part is located and what it does the instructor continuous explaining the next part, causing the image to change as well. This effect is the so-called *transient information effect*, which leads to a loss in students' learning (Ayres & Paas, 2007). According to cognitive load theory this imposes extraneous cognitive load which burdens the limited working memory (Sweller, Ayres & Kalyuga, 2011). A way to combat this problem is by guiding the learners' attention using cues (e.g. arrows, highlights) at key moments to help the learner select and process information (Jamet, 2014; Mayer, 2005; Tabbers, Martens & Van Merriënboer, 2004; Scheiter & Eitel, 2015). A meta-analysis on the cueing method found that cueing helps with reducing perceived invested mental effort (Xie et al., 2017). As part of Mayer's (2014) multimedia learning theory, cueing is defined as helping the learner in three aspects; selecting locations, organizing structure, and integrate relations between elements (De Koning, Tabbers, Rikers & Paas, 2009; Mayer, 2014). For example, in an eye tracking study Ozcelik, Arslan-Ari, and Cagiltay (2010) tried to teach the participants about a jet engine using example videos. The study found that the group of students who got a video cueing example, where text was highlighted in red during key moments, outperformed those who got the same video example but without cues on a retention test. The authors stated it was because of two reasons: (a) cueing guided the learner's attention to relevant information during the video, and (b) participants spend less time finding relevant information on the retention test. Furthermore, Johnson, Ozogul, and Reisslein (2014) showed that students regardless of prior knowledge benefitted from having been guided by a static arrow in algebra and graph reading problems, but lower prior knowledge students tend to benefit more.

One problem cueing faces is the often-varying effects on learning (De Koning, Tabbers, Rikers & Paas, 2010; Moreno, 2007). According to Jarodzka, Van Gog, Dorr, Scheiter, and Gerjets (2013) one reason this could be is because of a mismatch between the video's visual cue and where the learner should truly look to understand or solve a problem. Grant and Spivey (2003) experimented with guiding the learners' attention to a crucial area in Duncker's radiation problem. The problem consisted of a black dot with a black circle around it. Participants were told that the dot was a tumor and the circle was the skin of a patient. Their objective was to kill the tumor, but not the skin, using lasers. To solve the problem, participants had to use two lasers in different positions and angles converging on the tumor as to not damage the skin, but only the tumor inside it. The experiment had three groups: The control group, the tumor visually pulsing group and the skin visually pulsing group. Grant and Spivey (2003) found that the pulsing skin group outperformed the other groups, claiming it helped participants recognize this crucial information. Following this Litchfield and Ball (2011) experimented with showing an expert's eye movements when solving the same problem. It was found that the eye movement guided the learner's attention correctly and allowed the participants to solve the problem. This visualization technique could dynamically guide the learner's attention, leading to the learner synchronizing their attention to the model's gaze when solving a problem (Litchfield & Ball, 2011). Using the idea of synchronized attention guiding Van Gog, Jarodzka, Scheiter, Gerjets, and Paas (2009) developed eye movement modeling examples (EMME). In these videos the model's eye movements (shown as a moving dot) try to guide the learner through understanding a problem-solving method.

Eye Movement Modeling Examples

The effectiveness of EMME has found mixed results (Van Marlen, Van Wermeskerken, Jarodzka & Van Gog, 2016). For example, studies concerning university students (Jarodzka et al., 2012; Jarodzka, et al., 2013) showed that participants who had seen an EMME on how to classify fish fins and their locomotive patterns, outperformed those who only got a verbal explanation on a classification task. Adding to this, the way the eye movement is presented affects the cognitive load. When the guiding symbol was a spotlight (everything is blurred except part where the EMME wants you to look) the mental effort was increased (Jarodzka et al., 2013). Jarodzka et al. (2013) also found that the perceived mental effort of an EMME does not differ from a modeling example without visual aids. In other studies, concerning text processing, seventh graders were divided into two groups, control and EMME. It was found that the EMME group obtained better integrative text processing, verbal and graphical recall, and transfer scores (Mason, Pluchino & Tornatora, 2015a; 2015b).

In contrast, another study suggests that EMME do not benefit students in procedural problem-solving tasks and induces a higher mental effort (Van Gog et al., 2009). In the study by Van Gog et al. (2009), participants were shown a modeling example on how to solve an animated puzzle (i.e., frogs leaping). Either with or without verbal explanation and with or without the model's eye movement. When no explanation was present, it did not matter whether a participant was shown the eye movement or not. But when there was a verbal explanation, the eye movement had a negative effect on test performance. Similar results on EMME not benefitting students were found by a study by Van Marlen et al. (2016). In this study, university students were faced with calculating angles in a geometry problem. Participants saw either an EMME without a voice or a verbal modeling example (no visual aids). Besides the indifference on performance, the EMME group took longer to solve a transfer task that the non-EMME condition.

In another study Van Marlen, Van Wermeskerken, Jarodzka, and Van Gog (2018b) investigated the role of verbal ambiguity (specific or abstract sentences) and prior knowledge in EMME concerning geometry problems. Two groups of participants were gathered for the study, namely university students and Dutch secondary education students. Participants could get a condition differing in the type of modeling example (EMME vs just verbal) and ambiguity (ambiguous vs unambiguous) for a total of four conditions. In the experiment using university students, no difference was found regarding performance or response time. However, in the experiment using secondary education students the EMME group outperformed the regular modeling example group. Verbal ambiguity did not play a role in this difference. These findings indicated that EMME could have a place in modeling procedural problem-solving tasks when prior knowledge is low (Van Marlen, et al. 2018). Another study by Van Marlen Van

Wermeskerken, Jarodzka, and Van Gog (2018b) found no difference in perceived expertise of the modeler in EMME learning outcome. Adding to this that EMME's effects mirror research on cueing that learning is not always enhanced through the attention guidance (Van Marlen, 2019). The question remains whether an EMME's effectiveness depends on the attentional guidance or if simpler cueing can be just as effective.

The present study

The effects of EMME and cueing seem to be similar, but no research has been done comparing the two attention guiding methods. If EMME benefit solely from attention guidance, then another cueing method should provide similar if not the same results. Creating an EMME is a difficult process. Recordings for an EMME must be done with special hardware and software whilst a cueing modeling examples (CME), using other guiding methods (i.e. arrow or highlighting), can be created using a variety of software most people already have access to (i.e. Paint, Apple Paintbrush, etc.). The aim of the present study is to compare the effectiveness of EMME to CME. If EMME and CME don't differ in their ability to help learners solve problem-solving tasks it could provide insight to the generalization of EMME with cueing research.

To explore the comparison, an experiment was conducted with secondary education students where the visual guidance of procedural problem-solving modeling videos was manipulated. The video quality was compared as well. Inspiration for this study was drawn from two previous studies by Van Marlen et al. (2016; 2018a) exploring the effectiveness of EMME. In these studies, the F-rule and Z-rule were needed for geometric angle calculation as the procedural problem. Following the results of these studies, combined with the similar findings of EMME and CME (Ozcelik, Arslan-Ari & Cagiltay, 2010; Van Marlen, 2019; Van Marlen, et al., 2018a), it was hypothesized that secondary education students viewing the EMME or CME outperform those who don't on procedural problem-solving tasks (H1). Given that for both EMME and CME the visual cues were found to have a positive effect on learning, it was hypothesized that they do not differ in their effectiveness (H2a). However, based on the findings from Jarodzka et al. (2013), Van Gog et al. (2009), and Xie et al. (2017) it was expected that the perceived mental effort of the CME will be lower than the EMME (H2b). Lastly, it was hypothesized that perceived instructional quality does not differ across video conditions (H3) as in line with Van Marlen (2019).

Method

Participants and design

This data collection contained a total of four conditions, with the expected effect size being between medium and large (f = .25 - .40). A G*Power analysis (groups = 4)[1] indicated that the sample of this study should be between 76 participant and 180 participants. Therefore, the aim was to recruit a minimum of 30 participants for each condition.

The participants consisted of male (n = 23) and female (n = 23) Dutch middle school students between the ages 11 and 14 (M =12.48, SD = 0.62). All participants had (corrected to) normal eyesight. Participants were pseudo randomly spread over three conditions. The three conditions included: Eye movement modeling example EMME (n = 16), cueing modeling example (n = 16), and control (n = 14). Data collection was done as part of a bigger study regarding EMME. One other condition[1] was used during this data collection; however, it was not relevant for answering this study's research question. Participants did not receive any reward for participating in the experiment.

---

[1] Note that participants for the fourth condition were not told that the moving dot was representing the modeler's eye movement as part of research on the impact of social knowledge.

Instruments

All experiments were run using Qualtrics software and used the Dutch language as all the participants were Dutch native speakers.

Pretest. To test prior knowledge a pretest was used. This pretest consisted of two open questions (i.e. The degrees of angle A = …) and three multiple-choice question (i.e. Angle C is equal to: a. 360 - angle K - angle L… e. I don't know) concerning a geometry figure.

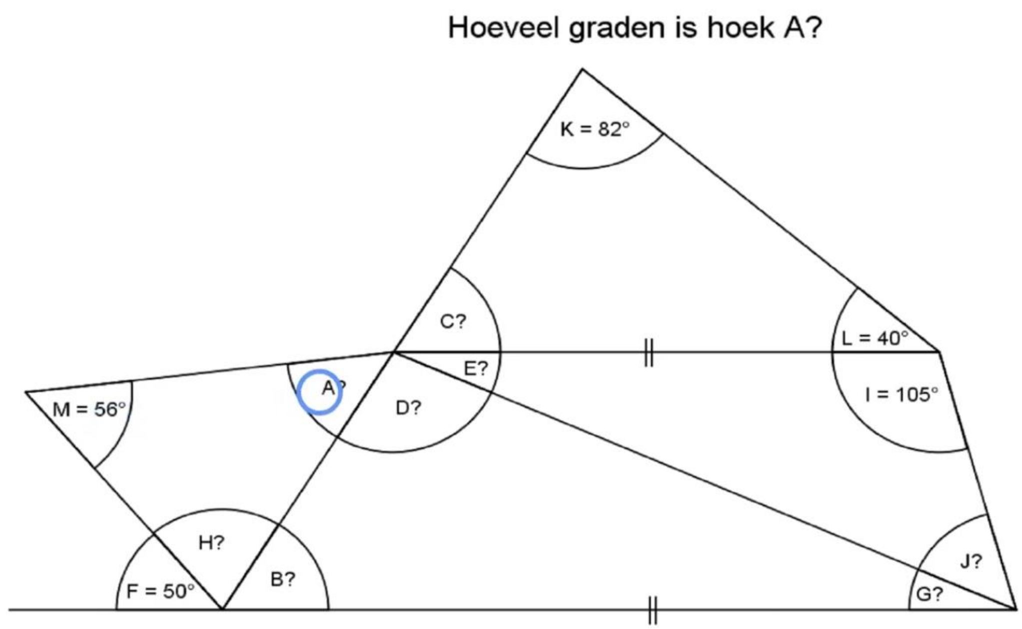

Figure 1. *An example of an EMME geometry problem used in this study. The text above asks the question "How many degrees is angle A?". The blue circle represents the eye movement from the modeler.*

Video modeling examples. In total two videos were made per condition. All videos included a modeling example for solving a geometry problem using the Z-rule and F-rule. A male model used a verbal explanation to model the problem-solving method. When an angle was calculated by the modeler the "?" was replaced by the calculated degrees, this applied for every video. The videos were manipulated in the way they visually guided the learner's attention. In the EMME video the eye movement of the speaker was recorded when providing the explanation. The eye movement was recorded using a SMI RED250 eye tracker with a recording rate of 250Hz. SMI Experiment Center 2.4.165 was used as software to process the eye movements. For the visualization of the eye movement, the software tool SMI BeGaze 3.4.52 was used to visualize the eye movement of the modeler for the video. The eye movement was shown as a 30-pixel diameter blue circle (see figure 1) as to not obscure the information behind the visual guidance. In the CME the eye movement circle was removed from the modeling example, and instead the angels, lines, and symbols were highlighted in red during key moments. These moments lasted approximately four seconds and appeared roughly around 8 times per video. During some key moments, multiple aspects could be highlighted at the same time (e.g. the "||" symbols for symmetrical lines). Key moments were identified and marked in the script before recording so the EMME and CME had the same focus points. For the control condition only the geometry problem was used providing no further visual aid for the learner.

Geometry problems. To create the figures, *Geogebra* (www.geogebra.org) was used. The problems consisted of black line drawings of triangles and parallel lines allowing the F-rule and Z-rule to be used in order to solve angles (see figure 1). Above each figure was a question concerning the degrees of a certain angle within the figure (i.e. "How many degrees is angle A?") which could only be solved by calculating four other angles. Angles were named using the alphabet starting from A. Some angles were given a value, whilst others were presented as question marks. Two geometry problems were used for the modeling video phase and two isomorphic figures (the exact same figure, but different values) for the performance phase. Additionally, four transfer problems were created which still used the F-rule and Z-rule but had different layouts from the isomorphic problems. This meant that for the transfer problems the participants had to find their own starting point in order to solve the problem, contrasting the isomorphic problems where the starting point remained the same. A physical calculator was always available to the participants whilst solving the problems.

Instructional quality. Participants were asked to rate how clear they thought the explanation was of the video on a 5-point Likert scale. This scaled ranged from 1 (*very* unclear) to 5 (*very* clear). This was asked of the participants after every modeling video.

Mental effort. To measure the difference in perceived mental effort of the testing phase, the Paas-scale (Paas & Van Merriënboer, 1993) was administered after the participant had completed the final transfer task. As part of measuring the instructional quality, the scale was administered after both modeling videos as well. This measurement did not differentiate between extraneous or intrinsic load. Participants had to select their invested mental effort on a scale from 1 (very, very low mental effort) to 9 (very, very high mental effort). The scale was presented on screen, translated to Dutch.

Procedure

Due to the Coronavirus lockdown all school buildings were closed at the time of data collection. Originally it was planned to conduct the experiment at the middle school in the classrooms, which was not possible anymore. To compensate, the decision was made to conduct the experiment online. Each participant got a single usable link to the experiment, where they would participate using their own device.

Participants were randomly assigned one of the conditions using the Qualtrics random distribution tool. The experiment started with an informed consent form providing the participants with information about the duration, problems to solve and data time storage. If the participant agreed and gave consent, their data would be included in the data set, if not it would be omitted whilst they could still participate in the Qualtrics. After the consent the participants were presented with an audio check. In the audio check the model spoke a sentence asking if the participant could hear her, after which participants could adjust their audio settings and listen again to make sure they could hear the upcoming videos as well. After the audio check the participants were asked to partake in the pretest consisting of 5 questions. After the pretest participants got to view a modeling video using either EMME, CME, or only audio followed by a question to rate instructional quality, their mental effort when viewing the video and an isomorphic problem. When they were done with the first video, they got to view another geometry problem modeling video followed by the same steps. Lastly, the participants got four transfer questions with each transfer question including a question to rate mental effort when they were done solving. Participants were thanked after filling out the final Paas-scale, after which they could close the Qualtrics.

The data were collected via Qualtrics, after which they were transferred to YoDa for 7-year storage (see appendix 1).

Data analysis

Prior knowledge. One point was awarded for each correctly answered question on the pretest resulting in a maximum score of 5 and a minimum score of 0. These scores were then translated to proportional scores by dividing their obtained score by the maximum score (5).

Performance. For the isomorphic problems and transfer problems, one point could be obtained for correctly answering the angle question above the figure. This led to a maximum score of 2 for the isomorphic problems and a maximum score of 4 for the transfer problems. Both scores were then translated to a proportional score by dividing the obtained score by the maximum possible score.

Instructional quality. The score of instructional quality was calculated by combining the participants' rating of both videos and dividing by two.

Mental effort. The mental effort scale for the videos administered after each video were combined and divided by two to create a single variable. For the perceived mental effort of the testing phase the score was determined by adding the value of every transfer Paas-Scale and dividing by 4.

Analysis. Participants with a time exceeding 7800 seconds (= 130 minutes) were excluded from the analysis because there was a high chance they either got distracted during the experiment or partook in parts, instead of as a whole. This meant their data were unreliable at best. Furthermore, a participant with the age of 30 was removed, as this was most likely the teacher of one of the classes and was therefore not part of the sample group. This left a total of 16 participants for both EMME and CME, and 14 for the control condition. Normality of the data was checked using SPSS version 25. If normality was assumed, the data were analyzed with one-way ANOVAs using the dependent variables: Isomorphic score, transfer score, mental effort, and instructional quality. The independent variables were the manipulated conditions: EMME, CME, and control. Effect sizes were defined as small ($f > .10$), medium ($f > .25$), or large ($f >.40$) as stated by Cohen (1988). Furthermore, a Bayesian analysis was conducted with JASP 0.8.6.0 (www.jasp-stats.org). A Bayesian analysis has the advantage that it shows how more likely an alternative hypothesis is compared to a null hypothesis instead of simply rejecting it (Wagenmakers et al., 2018). For instance, a Bayesian factor of 7.00 means that the null hypothesis is seven times more likely than the alternative hypothesis and a factor of 0.70 means the alternative hypothesis is seven times more likely than the null hypothesis. When interpreting these numbers, a factor of 1 is called no evidence, 1 to 3 is called anecdotal evidence and a factor of 3 to 10 is called substantial evidence for the null hypothesis (Wagenmakers, Wetzels, Borsboom & Van der Maas, 2011).

Table 1.

Mean (SD) and minimum and maximum of proportion correct pretest isomorphic and transfer, perceived instructional quality and mental effort of the eye movement modeling example, cueing modeling example and control condition.

	EMME (*n* = 16) *M* (SD) | *Min. – Max.*	CME (*n* = 16) *M* (SD) | *Min. – Max.*	Control (*n* = 14) *M* (SD) | *Min. – Max.*
Proportion Correct Pretest	0.54 (0.23) | 0.20 – 0.80	0.65 (0.24) | 0.20 – 1.00	0.74 (0.18) | 0.40 – 1.00
Proportion correct Isomorphic	0.59 (0.38) | 0.00 – 1.00	0.78 (0.36) | 0.00 -1.00	0.71 (0.38) | 0.00 – 1.00
Proportion Correct Transfer	0.47 (0.31) | 0.00 – 1.00	0.67 (0.39) | 0.00 – 1.00	0.61 (0.25) | 0.25 – 1.00
Video Mental Effort	2.75 (1.22) | 1.50 – 5.50	2.56 (1.75) | 1.00 – 7.00	3.04 (1.17) | 1.50 – 5.00
Transfer Mental Effort	4.76 (1.91) | 1.25 – 8.50	4.05 (1.51) | 2.00 – 7.00	3.98 (1.59) | 1.25 – 6.75
Instructional Quality	4.31 (0.48) | 3.50 – 5.00	4.28 (0.55) | 3.50 – 5.00	4.25 (0.43) | 3.50 – 5.00

Results

The pretest scores, isomorphic scores, transfer scores, instructional ratings, and mental effort ratings are shown in table 1. The average duration amongst the conditions for completing the experiment was 44 minutes. Most data were analyzed with non-parametric Kruskal-Wallis *H* tests due to violation of normality after being checked in SPSS (version 25). The transfer mental effort score however, was analyzed using a one-way ANOVA due to being normally distributed. Bayesian analysis were conducted for all variables and subsequently interpreted according to the definitions of Wagenmakers et al. (2011).

Prior knowledge

First, prior knowledge was checked to make sure there were no differences between the conditions. Results indicated no difference in prior knowledge between the groups, χ^2 (2) = 5.38, p = .07, BF_1 = 1.6. From the Bayesian factor it can be interpreted that there is anecdotal evidence for the null hypothesis of there being no difference.

Learning outcome

Secondly, to check whether the type of video viewed affected learning outcome the scores on the test were analyzed. Performance on the isomorphic tests did not differ between conditions, χ^2 (2) = 2.56, p = .28, BF_1 = 3.0 neither did the transfer tests χ^2 (2) = 3.25, p = .20 BF_1 = 2.0. Both Bayesian factors provide anecdotal evidence that there is no difference.

Mental effort

Thirdly, to test the hypothesis that CME would have lower mental effort the mental effort ratings were analyzed. Mental effort after viewing a video did not differ between conditions, χ^2 (2) = 2.45, p = .29, BF_1 = 4.6, nor did the mental effort after the transfer tests, F (2, 43) = 1.04, p = .362, r = .04, BF_1 = 3.0. The Bayesian factor concludes that the video mental effort provides substantial evidence for finding no difference and the transfer mental effort providing anecdotal evidence.

Instructional quality

Lastly, the perceived instructional quality was analyzed to test the hypothesis that quality does not differ across conditions. The analysis showed that there is no difference between the conditions regarding instructional quality, χ^2 (2) = 0.10, p = .95, BF_1 = 5.9, with the Bayesian factor providing substantial evidence.

Discussion

The aim of the current study was to compare how EMME and CME affect learning, perceived instructional quality and mental effort. This study tried to extend current research by taking two attention guiding methods (Sweller, Ayres & Kalyuga, 2011), namely EMME (Van Gog, 2009) and CME (Mayer, 2014), and compare them to one another. The experiment was focused on procedural problem-solving tasks where EMME seem to mirror the findings of CME research where learning is not always enhanced through attention guidance (Van Marlen, 2019). Based on previous research (Ozcelik, Arslan-Ari & Cagiltay, 2010; Van Marlen, 2019; Van Marlen et al. 2018a) it was hypothesized that EMME and CME would not differ in learning outcome, but both would outperform the non-guided modeling example condition. Furthermore, it was hypothesized that CME would have a lower perceived mental effort than EMME (Jarodzka et al. 2013; Van Gog et al. 2009; Xie et al, 2017), and that none of the video conditions would differ in instructional quality (Van Marlen, 2019).

Contrary to the first hypothesis, the EMME and CME did not outperform the control condition on either the isomorphic task or transfer task. When taking the Bayesian factors into account (BF_1 = 3.0, and BF_1 = 2.0 respectively) it can be interpreted that these results provide anecdotal evidence that condition does not matter. This finding is in line with the findings of De Koning et al. (2010), Moreno (2007), and Van Marlen et al. (2016) that cueing methods often find varying results. One reason for this finding could be the relatively high prior knowledge of the sample group. More than half of the pretest questions were answered correctly across all groups, meaning the participants had at least some, if not high prior knowledge on how to solve geometry problems concerning triangles using the F-rule and Z-rule. This could have influenced the scores as Richter, Scheiter, and Eitel (2016) found that prior knowledge moderates learning outcomes when using signaling in multimedia materials.

As found by Van Marlen et al. (2018b), lower prior knowledge for solving geometry problems lead to an effect of EMME on learning outcome but not when prior knowledge was high. In this study Van Marlen et al. (2018b) compared two groups; namely university students (experiment 1) and secondary education students (experiment 2). The second experiment consisted of low prior knowledge individuals who got about 40% correct on the pretest. The sample used had the same characteristics as the sample used in this experiment; the same age and educational year, and educational level. However, when comparing the group of Van Marlen et al (2018b) to the group of this experiment a discrepancy can be found. In this experiment an average of 69% correct can be found on the pretest, which is considerably more than the sample used by Van Marlen et al. (2018b). The findings of this study thus lend credit to both Van Marlen et al. (2018b) and Johnson, Ozogul, and Reisslein (2014) that lower prior knowledge students benefit more from attention guiding.

As expected, none of the conditions differed in instructional quality or mental effort. The Bayesian analysis also concludes that there is substantial evidence that none of these conditions differ regarding instructional quality. This finding lends credit to the finding of Van Marlen (2018a) that EMME and non-attention guiding modeling examples are perceived to be similar in instructional quality regarding geometry problems. In another study by Van Marlen et al. (2018b) a small effect was found on instructional quality when comparing EMME to ME. However, this effect was also found to have a significant interaction with ambiguity. In this experiment ambiguity was not a factor, thus it could be that no effect of condition was found due to the modeling examples being unambiguous.

Interestingly, the mental effort experienced when watching instructional videos was thought to be lower for the CME group as opposed to the erratic and wobbly movements of an EMME or the unguided approach of the control. The results of the experiment contradict this hypothesis, with the Bayesian analysis providing substantial evidence that there is no difference. This confirms the finding of Jarodzka et al (2013) that mental effort does not differ between EMME and a non-visual modeling example but, contradicts the meta-analysis findings of Xie et al. (2017) that cueing helps reduce cognitive load. One reason this could be was the due to the sample groups consisting of generally high prior knowledge, meaning they experienced less cognitive load overall because the subject was not new to them. Adding to this that the effect found by Xie et al. (2017) was small-medium (d = .11), it could mean that the prior knowledge compensated for the added benefit of cueing. This was a factor that Xie et al. (2017) did not consider when conducting their meta-analysis, even though they mention that intrinsic cognitive load is an interactive element between learning material and learner prior knowledge (Sweller, Ayres & Kalyuga, 2011).

Besides these speculations for finding no difference between the conditions, the sample size and power could also be a factor as there is little power due to the size. This study aimed to have about 30 participants per group, but due to the Covid-19 outbreak prior arrangements to test the groups were put on halt and less individuals were able to participate. This led to the experiment being conducted not in a classroom, but individually, at home on a laptop. This also meant the participants were in a less controlled environment and might have been exposed to some forms of distractions when following the online experiment or have gotten help from another person. However, when comparing the time spend on the experiment (44 minutes) to a similar study (Van Marlen, 2018a) where students spend a maximum of 50 minutes, the time spent on the experiment do not seem that different. The influence of distractions can only be speculated, as no data were gathered concerning these factors. This makes the data a bit more unreliable than it otherwise could have been.

However, random distribution and looking at the pretest outcomes suggest some, if not most, of the groups had high prior knowledge anyway. When comparing these outcomes to other EMME experiments (e.g. Van Marlen et al., 2016; 2018b) they are not that strange of a result and make the experiments' findings more reliable. Lastly, the Bayesian analysis conducted consider all possible statistical origins where these findings could have come out of. With these analyses providing anecdotal evidence that there might be no difference between these attention guiding methods, lending credit to Van Marlen (2019) saying EMME mirrors cueing research in that learning is not always enhanced when using these methods and finding mixed results (De Koning, Tabbers, Rikers & Paas, 2010; Moreno, 2007; Van Marlen et al., 2016).

For future research, it could be interesting to use eye tracking data when comparing EMME to CME. As found by Van Marlen et al. (2018b) in an ambiguous modeling example EMME tend to guide a learner's attention more often and faster to the point of interest. Using this method for comparing EMME and CME could provide more insight into the generalizability of EMME and CME to one another. It could also be that geometry problems are not visually complex enough for the learner to benefit from the attention guiding. For this reason, it could be interesting to see how EMME and CME compare when engaging in more visually cluttered procedural problem-solving tasks. Furthermore, it is recommended for future research to take prior knowledge into account whenever studies are conducted that use a form of attention guiding and subsequently measure cognitive load. It has been shown that prior knowledge is a moderating factor in learning outcome (Richter et al., 2016; Van Marlen et al., 2018b) and influences cognitive load (Sweller, Ayres & Paas, 2011).

As for practical implications, it seems that no matter what method you use to guide a learner's attention during geometry problems, the effect could be the same. For math teachers trying to teach their students geometry they could just stick to their old methods. But, if they notice that pointing with a pointer does not seem to work, they could always try an EMME or CME. When doing this however, a CME would be recommended as these seem to be easier to make using readily available tools as opposed to an expansive eye tracker, whilst having the same learning outcomes.

To conclude, the current study found no support that EMME or CME are any different from one another in either learning outcome, perceived instructional quality, or mental effort. This result could have been influenced by the high prior knowledge of which this sample group possessed. This indicated that research concerning attention guiding methods should keep prior knowledge in mind as a strong mediating factor. This is good news for math teacher trying to create videos for solving geometry problems, as they would not need either expensive eye tracking equipment or editing skills to create a valid educational video.

References

Ayres, P., & Paas, F. (2007). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology, 21*, 695–700. doi:10.1002/acp.1343.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a Framework for Attention Cueing in Instructional Animations: Guidelines for Research and Design. *Educational Psychology Review, 21*(2), 113–140. doi.org/10.1007/s10648-009-9098-7

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: seeing is understanding? *Learning and Instruction, 20*, 111-122. doi.org/10.1016/j.learninstruc.2009.02.010.

Fiorella, L., & Mayer, R. E. (2018). What works and doesn't work with instructional video. *Computers in Human Behavior, 89*, 465–470. doi.org/10.1016/j.chb.2018.07.015

Grant, E. R., & Spivey, M. J. (2003). Eye Movements and Problem Solving. *Psychological Science, 14*(5), 462–466. doi.org/10.1111/1467-9280.02454

Jamet, E. (2014). An eye-tracking study of cueing effects in multimedia learning. *Computers in Human Behavior, 32*, 47–53. doi.org/10.1016/j.chb.2013.11.013

Jarodzka, H., Balslev, T., Holmqvist, K., Nyström, M., Scheiter, K., Gerjets, P., & Eika, B. (2012). Conveying clinical reasoning based on visual observation via eye-movement modeling examples. *Instructional Science, 40*(5), 813–827. doi.org/10.1007/s11251-012-9218-5

Jarodzka, H., Van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see: Guiding students' attention via a Model's eye movements fosters learning. *Learning and Instruction, 25*, 62–70. doi.org/10.1016/j.learninstruc.2012.11.004

Johnson, A., Ozogul, G., & Reisslein, M. (2014). Supporting multimedia learning with visual signalling and animated pedagogical agent: moderating effects of prior knowledge. *Journal of Computer Assisted Learning, 31*(2), 97–115. doi.org/10.1111/jcal.12078

Litchfield, D., & Ball, L. J. (2011). Rapid communication: Using another's gaze as an explicit aid to insight problem solving. *Quarterly Journal of Experimental Psychology, 64*(4), 649–656. doi.org/10.1080/17470218.2011.558628

Mason, L., Pluchino, P., & Tornatora, M. C. (2015a). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology, 41*, 172–187. doi.org/10.1016/j.cedpsych.2015.01.004

Mason, L., Pluchino, P., & Tornatora, M. C. (2015b). Using eye-tracking technology as an indirect instruction tool to improve text and picture processing and learning. *British Journal of Educational Technology, 47*(6), 1083–1095. doi.org/10.1111/bjet.12271

Mayer, R. E., & Mayer, R. E. (Eds.). (2005*). The Cambridge handbook of multimedia learning*. Cambridge, United Kingdom: Cambridge university press.

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43-71). New York: Cambridge University Press.

Moreno, R. (2007). Optimizing learning from animations by minimizing cognitive load: cognitive and affective consequences of signalling and segmenting methods. *Applied Cognitive Psychology, 21*, 765-782. doi.org/10.1002/acp.1348.

Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior, 26*(1), 110–117. doi.org/10.1016/j.chb.2009.09.001

Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1993). The Efficiency of Instructional Conditions: An Approach to Combine Mental Effort and Performance Measures. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 35*(4), 737–743. doi.org/10.1177/001872089303500412

Tabbers, H. K., Martens, R. L., & Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology, 74*(1), 71–81. doi.org/10.1348/000709904322848824

Renkl, A. (2013). Toward an Instructionally Oriented Theory of Example-Based Learning. *Cognitive Science, 38*(1), 1–37. doi.org/10.1111/cogs.12086

Richter, J., Scheiter, K., & Eitel, A. (2016). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educational Research Review, 17*, 19–36. https://doi.org/10.1016/j.edurev.2015.12.003

Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction, 36*, 11–26. doi.org/10.1016/j.learninstruc.2014.11.002

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York, NY: Springer

Van Gog, T., & Rummel, N. (2010). Example-Based Learning: Integrating Cognitive and Social-Cognitive Research Perspectives. *Educational Psychology Review, 22*(2), 155–174. doi.org/10.1007/s10648-010-9134-7

Van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Computers in Human Behavior, 25*, 785-

791. doi.org/10.1016/j.chb.2009.02.007

Van Gog, T., Rummel, N., & Renkl, A. (2019). Learning How to Solve Problems by Studying Examples. In J. Dunlosky & K. Rawson (Eds.), *The Cambridge Handbook of Cognition and Education* (Cambridge Handbooks in Psychology, pp. 183-208). Cambridge: Cambridge University Press. doi.org/10.1017/9781108235631.009

Van Marlen, T. (2019). *Looking Through the Teacher's Eyes: Effects of Eye Movement Modeling Examples on Learning to Solve Procedural Problems* (Doctoral dissertation, University Of Utrecht, Utrecht, Netherlands). Retrieved from https://dspace.library.uu.nl/handle/1874/380365

Van Marlen, T., Van Wermeskerken, M., Jarodzka, H., & Van Gog, T. (2016). Showing a model's eye movements in examples does not imrpove learning of problem-solving tasks. *Computers in Human Behavior, 65*, 448-459. doi.org/10.1016/j.chb.2016.08.041

Van Marlen, T., Van Wermeskerken, M., Jarodzka, H., & Van Gog, T. (2018a). *Does the Purported Exerptise of a Model Influence the Effectiveness of Eye Movement Modeling Examples?* Manuscript submitted for publication.

Van Marlen, T., Van Wermeskerken, M., Jarodzka, H., & Van Gog, T. (2018b). Effectiveness of Eye Movement Modeling Examples in problem solving: The role of verbal ambiguity and prior knowledge. *Learning and Instruction, 58*, 274-283. doi.org/10.1016/j.learninstruc.2018.07.005

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., et al. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review, 25*, 35–37. doi.org/10.3758/s13423-017-1343-3.

Wagenmakers, E.J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology, 100*(3), 426–432. https://doi.org/10.1037/a0022790

Xie H, Wang F, Hao Y, Chen J, An J, Wang Y, et al. (2017) The more total cognitive load is reduced by cues, the better retention and transfer of multimedia learning: A meta-analysis and two meta-regression analyses. *PLoS ONE, 12*(8), 1-20. doi.org/10.1371/journal.pone.0183884

Appendix

Appendix 1. FETC Form

Section 1: Basic Study Information

1. Name student:

Maurice Potuijt

2. Name(s) of the supervisor(s):

Tim van Marlen

3. Title of the thesis (plan):

The similarities and differences in modeling videos between cueing and eye movement modeling examples

4. Does the study concern a multi-center project, e.g. a collaboration with other organizations, universities, a GGZ mental health care institution, or a university medical center?

~~Yes~~ / No
If yes: Explain.

5. Where will the study (data collection) be conducted? If this is abroad, please note that you have to be sure of the local ethical codes of conducts and permissions.

This study will be conducted on secondary education schools in the Netherlands.

Section 2: Study Details I

6. Will you collect data?

Yes / ~~No~~

Yes → Continue to question 11
No → Continue to question 7

7. Where is the data stored?

The data will be first stored Qualtric and then on YODA. It will be saved on YODA for 7 years

8. Is the data publicly available?

~~Yes /~~ No
If yes: Where?

9. Can participants be identified by the student? (e.g., does the data contain (indirectly retrievable) personal information, video, or audio data?)

Yes / ~~No~~
If yes: Explain. Name and gender will be asked to distribute participants to conditions.

10. If the data is pseudonymized, who has the key to permit re-identification?

Tim van Marlen, Maurice Potuijt, and Kirsten van der Lecq will have the key to the information

Section 3: Participants

11. What age group is included in your study?

11 – 14 year old students

12. Will be participants that are recruited be > 16 years? Yes/~~No~~

13. Will participants be mentally competent (wilsbekwaam in Dutch)? Yes/~~No~~
14. Does the participant population contain vulnerable persons?
    (e.g., incapacitated, children, mentally challenged, traumatized, ~~Yes~~/No

    pregnant)

15. If you answered ‘Yes’ to any of the three questions above: Please provide reasons to justify why this particular groups of participant is included in your study.

> This study is a follow-up study. It is based on studies by Van Marlen et al. (2016; 2018). These studies found a positive learning effect of EMME with secondary education students, focusing on (yet) unknown mathematical for them. Older student might have too much mathematical prior knowledge, which could make the data unreliable.

16. What possible risk could participating hold for your participants?

> There are low risks for the participants in this experiment. Possible risks are that this experiment could cause a learner to have stress about performing in the experiments. It takes the participants one math class in their regular school schedule, which is about 45 minutes of their time.

17. What measures are implemented to minimize risks (or burden) for the participants?

> The experiment will be conducted on the school of the students. Therefore, student do not have to put effort in travelling. The experiment will take place during a normal school hour, which means that the participants do not have to spend extra time in school for this experiment. Also, the subject being taught fit within the curriculum. Participation is voluntary and if participants do not want to participate they can quit at any moment.

18. What time investment and effort will be requested from participants?

> The time investment of the participants will approximately be 45 minutes.

19. Will be participants be reimbursed for their efforts? If yes, how? (financial reimbursement, travelling expenses, otherwise). What is the amount? Will this compensation depend on certain conditions, such as the completion of the study?

No the participants will not be reimbursed for their efforts.

20. How does the burden on the participants compare to the study's potential scientific or practical contribution?

There is a small burden on the participants. Participation in this experiment only requires 45 minutes of their time. Modelling examples are increasingly being used for informal and formal learning and EMME is a form of a modelling example. Which could potentially be used more often in the future. However, not a lot of research is yet conducted on the effectiveness of EMME for learning and on instructional guidelines for modelling examples. Therefore, research is necessary to get more insights on this topic.

21. What is the number of participants? Provide a power analysis and/or motivation for the number of participants. The current convention is a power of 0.80. If the study deviates from this convention, the FERB would like you to justify why this is necessary. (Note, you want to include enough participants to be able to answer your research questions adequately, but you do not want to include too many participants and unnecessarily burden participants.)

This study contains four conditions and the expected effect size is between medium and large (.25-.40). The power analysis with G*Power indicated that the sample of this study should be between 76 participant and 180 participants (Groups= 4, β=0.80, α= 0.05, F= .25- .40). Therefore, we aim to recruit a minimum of 30 participants for each condition.

22. How will the participants be recruited? Explain and attach the information letter to this document.

School will be approached via mail with information about the experiment. Asking school to participate in this experiment. Also, an information letter for the parents is attached with information about participation.

23. How much time will prospective participants have to decide as to whether they will indeed participate in the study?

The participants have two weeks to decide whether they want to participate in the experiment.

24. Please explain the consent procedures. Note, active consent of participants (or their parents) is in principle mandatory. Enclose the consent letters as attachments. You can use the consent forms on Blackboard.

Two weeks before the experiments students and parents are being informed about the experiments. A passive consent is asked form the parents, meaning that they actively have to state if they do not want their children to participate. The students have to give active consent to participate in the experiment. This consent is asked at the start of the experiment. Participation is voluntary and if participants do not want to participate they can quit at any moment.

25. Are the participants fully free to participate and terminate their participation whenever they want and without stating their grounds for doing so? Explain.

Yes participants are fully free to choose to participate in the study.

26. Will the participants be in a dependent relationship with the researcher?

~~Yes~~ / No
If yes: Explain.

27. Is there an independent contact person or a general email address of a complaint officer whom the participant can contact?

Anouschka van Leeuwen (A.vanLeeuwen@uu.n)
Tim van Marlen (T.V.A.vanMarlen@uu.nl)