The effect of model selection on learning and motivation when learning with video-modeling examples.

Noortje Pater (6034926)

Research Plan Master's Thesis

Course Code: 201600025

University Utrecht

Supervisor: Dr. T.V.A. van Marlen

Submission date: 08-06-2020

Word count: 7986

In the educational field, the use of video modeling examples, in which a model demonstrates and explains a task, has increased in the past few years. The observer-similarity hypothesis (MOS) argues that the effectiveness of modeling on self-efficacy and learning depends on how the observer perceives the model to be similar to them. However, research into MOS shows mixed findings which may be due to the effect of the task-appropriateness hypothesis, the perceived fit between the task and the model. Because the perception of task-appropriateness of a model could differ among learners, model selection could play an important role in learning and motivation. Therefore, it was investigated whether model selection would affect motivation and the effectiveness of video modeling examples for learning. In this study, participants (N = 73) watched two video modeling examples about troubleshooting electrical circuits. Half of the participants were allowed to choose a model while the other half was randomly assigned to one of the models. Results suggests that there is no need to take model selection into account when learning with video modeling examples. No effects were found for model selection on learning, motivation, invested mental effort and explanation quality of the models.

Key words: Video modeling examples, model-observer similarity hypothesis, task-appropriateness hypothesis, model selection, motivation

Video Modeling Examples

The past few years, video modeling examples, in which a model demonstrates and explains a task, have become a more and more frequently used strategy in the educational field. Beside the use of video modeling examples in informal learning settings (e.g., Youtube) it is also used in formal settings like flipped classrooms. According to Van Der Meij & Van Der Meij (2014), the increase of the use of video modeling examples may be due to the facts that they are easy to create, time saving, cost-efficient and considered to be an effective tool for learning.

Although the popularity of video modeling examples is growing, questions arise about effective design of these videos and interest in research on this topic is growing. For example, recent studies investigated the effect of the first or third person used in the explanation (Fiorella, Van Gog, Hoogerheide, & Mayer, 2017), the effect of presence or absence of the model (Hoogerheide, Loyens, & Van Gog, 2014) and the effect of seeing the instructor's face and gaze (Van Wermeskerken & Van Gog, 2017). Nevertheless, due to the diversity of formats of video modeling examples, it is difficult to establish general design guidelines (Bétrancourt & Benetos, 2018).

Therefore, is it important to pursue research on design characteristics of video modeling examples to gain more insight into effective design of video modeling examples for learning. One important design question concerns the effect of model characteristics, such as gender or age, on learning (Hoogerheide, Loyens, & Van Gog, 2016; Hoogerheide, Van Wermeskerken, Loyens, & Van Gog, 2016; Hoogerheide, Van Wermeskerken, Van Nassau, & Van Gog, 2018). Furthermore, no research has been conducted to investigate whether having the opportunity to choose your own model has an impact on learning and motivation. Understanding the effect of model selection on learning and motivation could shed light on future hypotheses for optimal design of video modeling examples and could be relevant for practice: at present the most common way to learn from video modeling examples is on sites such as YouTube where learners have to select their own videos.

The Opportunity of Choice

According to Deci and Ryan (2000), having the autonomy to choose plays an essential role for intrinsic motivation and learning as described in the Self-Determination Theory (SDT). SDT states that motivation to grow and change is based on the fulfilment of three psychological needs: autonomy (the need to experience control), competence (the need to be effective in dealing with the environment) and relatedness (the need to experience a sense of belonging and connection). When these psychological needs are addressed, intrinsic motivation increases: learners feel the drive to explore and experiment in their learning environment, they show interest in attractive activities and enjoy them (Deci & Ryan, 2000).

In SDT, autonomy refers to the need to experience choice and being in charge of own learning and behaviour (Deci & Ryan, 1987; Ryan & Connell, 1989; Sheldon & Elliot, 1999). Implementing autonomy into the learning environment may be done by providing learners with a choice (Zuckerman, Porac, Lathin, Smith, & Deci, 1978). This may have an positive influence on learners' intrinsic motivation, confidence in abilities and learning (Deci, Schwartz, Sheinmann, & Ryan, 1981; Tafarodi, Milne & Smith, 1999).

Learner Control in Education

Providing learners with a choice can be considered as a form of learner control which has a positive influence on how interesting learning is experienced (Fry, 1972; Kinzie & Sullivan, 1989; Lahey, 1973) and can determine the intrinsic motivation to learn (Deci & Ryan, 1985). According to Wolters (2003) and Zimmerman (2002), it is expected that learning becomes more appealing and has a positive effect on learners' motivation when learners are given (partial) control over aspects of their learning environment. This in turn is beneficial for learning outcomes.

For instance, providing learners with control over learning tasks leads to more effective, efficient and appealing learning (Corbalan, Kester, & Van Merriënboer, 2006), higher interest in learning task and higher mean performance (Cordova & Lepper, 1996). Research of Corbalan et al. (2006) investigated if a personalized selection of learning tasks with shared control leads to more effective, efficient and appealing learning. This study involved twenty-five nursing students who were randomly assigned to the shared instructional control condition (learner selects the final task out of a subset of tasks selected by an instructional agent) or full system control condition (learning task selection carried out by a system selecting learning tasks from an existing database). Results showed that participants in the shared-controlled condition showed a higher mean performance and a higher interest in the learning task in comparison with participants in the system-controlled condition.

Besides investigating learner control in context of learning tasks, research has shown that giving learners control over who gives the explanation affects learning (Baylor, Ave, & Park, 2005; Ozogul, Johnson, Atkinson, & Reisslein, 2013). For example, research has been conducted to investigate the impact of selecting an Animated Pedagogical Agent (APA). Research of Ozogul et al. (2013) involved middle-school students learning about electrical circuit analysis in a computer-based environment with an APA. Results showed that when a learner was provided with a choice of animated agent, this led to higher far transfer scores. Furthermore, they investigated the effect of gender matching and found that participants receiving the opposite gender scored significantly higher than participants receiving the same gender.

Baylor et al. (2005) further explored the effect of gender and results showed that the response of learners to system interfaces is affected by agent gender, ethnicity and realism of the agent. Students watched a presentation from an APA on coping with college life and rated the APA on factors of human-like, credibility, engaging, and facilitating learning. Half of the students could choose an APA while the other half could not. Results showed that students tended to choose an APA of the same ethnicity and were not more likely to choose male agents. But when students were randomly assigned to an APA, they rated the male APA as more useful, interesting and satisfactory than the female APA. Based on results of this study, Baylor et al. (2005) suggest that observers' gender and ethnicity play a role in learners' preference for a model and how learners respond to system interfaces. Therefore, it is important to take individual differences into account when learning with social interfaces.

Although research in the field of APAs shows that learning is affected by learners' ability to choose an APA, the question remains if these effects are also found when learning with video modeling examples using real-life models.

Model-observer Similarity Hypothesis

When it comes to learning with video modeling examples, a possible way to provide learners with a choice could be by allowing them to choose a model based on model characteristics such as age or gender (Hoogerheide, Loyens et al., 2016; Hoogerheide, Van Wermeskerken et al., 2016).

According to the model-observer similarity hypothesis (MOS), the extent to which the choice of a model contributes to the effectiveness of modeling on self-efficacy and learning depends on whether the observer perceives a model to be similar to them. In other words, when a model has characteristics that the learner can identify with and the model successfully demonstrates a task, social comparison may occur and make the learner believe they can perform the task as well (Bandura, 1994; Berger, 1977). Nevertheless, research in the field of video modeling examples shows varying results for MOS concerning model characteristics such as age and gender.

Age. Davidson and Smith (1982) investigated the difference in learners' ability to imitate modelled standards of self-reward, watching one of three models varying in age. Results revealed that observing a peer model or inferior model led to matched or higher standards of self-reward than observing an adult superior model. Comparable results were found in a recent study of Buritica, Eppinger, Schuck, Heekeren, and Li (2016) which showed that during observational learning children preferred imitating the choices of a peer model instead of an adult model. According to Schunk (1987) an explanation for these results could be that peer models are more beneficial for learners who experience difficulties with learning or have a low ability.

In contrast, other studies show the benefits of an adult model: learners respond more strongly to adults and effects on learning lasted longer (Hicks, 1965; Jakubczak & Walters, 1959). This is in line with recent research of Hoogerheide and Van Wermeskerken et al. (2016) which showed an effect of model's age on learning about troubleshooting electrical circuits: learning from adult models led to higher learning outcomes than learning from peer models. Participants rated the quality of the explanations of adult models higher than those who studied peer models, even though the video modeling examples were otherwise identical.

Gender. Concerning gender, empirical findings are also mixed. Where the study of Bussey and Bandura (1984) revealed that learners showed greater imitative behaviour after viewing video modeling examples with the same-gender model than the opposite gender model, research of Hoogerheide and colleagues did not (Hoogerheide, Loyens et al., 2016; Hoogerheide et al., 2018). In the study of Hoogerheide and Loyens et al. (2016) secondary education students watched video modeling examples explaining how to solve probability calculation problems. Results showed no effects of model's gender or observer's gender on learning and near transfer.

Although there are mixed findings in research of MOS in the field of video modeling examples, it is important to mention that participants in all studies mentioned above were assigned to a condition and were not allowed to choose a model themselves. Having the opportunity to choose your own model might affect motivation and learning (Deci et al., 1981; Tafarodi et al., 1999) and might explain the discrepancy in results of different studies on MOS. This is because participants may have individual differences in preference for model characteristics (Baylor et al., 2005). Groups of participants in different studies may differ in their preference for a model, which may have skewed results of studies which allocate participants to a random condition.

Task-appropriateness Hypothesis

Besides choosing a model based on model characteristics such as age or gender, the choice of model can also be based on the task appropriateness of the model: the perceived fit between the task and the model.

Whereas previously MOS literature shows that the task-appropriateness of a model has a strengthening effect on MOS effects (Schunk, 1987), recent research revealed that the perceived task-appropriateness of a model can have an effect on learning independently of similarity views, also known as the task-appropriateness hypothesis.

For instance, Hoogerheide and Van Wermeskerken et al. (2016) conducted a study with secondary education students who watched video modeling examples about troubleshooting electrical circuits. Results showed that students watching adult models had higher learning outcomes in comparison with students who watched peer models. In addition, the explanation quality of adult models was rated higher although both peer and adult models gave the exact same explanation. According to Hoogerheide and Van Wermeskerken et al. (2016) these findings could not be caused by individual characteristics of the models (multiple models were used per condition) and therefore hypothesized that the age of a model can influence the appropriateness of a model for a task. When students perceive an adult model as appropriate for explaining a task, students would pay more attention to the adult model, leading to better learning. According to Duit and von Rhöneck (1998) students tend to pay more attention to adult models than peer models when they perform a difficult physics task.

Furthermore, several studies found that a male model was more beneficial for learning for both male and female learners when a task was found more appropriate for males, which is in line with the task-appropriateness hypothesis. Research of Bandura, Ross, and Ross (1963) and Hicks (1965) revealed that children around the age of four watching a model displaying aggression towards a doll showed more imitative aggression if they watched a male model instead of watching a female model.

In contrast, research of Hoogerheide and Loyens et al. (2016) showed no effect of model gender on self-efficacy and learning outcomes but watching a male model led to higher perceived competence for both male and female learners. In addition, for male learners, watching the male model led to more learning enjoyment and less invested mental effort than male learners watching a female model.

However, there are two important things to note. First, this study only used one male or female model per condition, whereby these effects could be a result of individual model characteristics. Second, participants did not have the opportunity to choose. If so, it might have had an effect on intrinsic motivation and learning according to SDT (Deci & Ryan, 2000). A possible effect may then have been found for chosen models' gender on self-efficacy and learning.

To rule out the possibility that individual model characteristics might affect results, recent research of Hoogerheide et al. (2018) did make use of two adult male models and two adult female models. In this study, secondary education students watched two video modeling examples with a male of female model explaining a task perceived as more appropriate for males: troubleshooting electrical circuits. Results showed higher learning outcomes and a higher confidence for males then for females. On the other hand, results showed no effect of model's gender on learning, self-efficacy, perceived competence, invested learning effort and learning enjoyment. A possible explanation for these results could be that learners differ in who they find appropriate for the task and again participants were not allowed to choose their own model which might affect the results.

In sum, studies in the field of video modeling examples with regard to MOS or the task-appropriateness hypothesis show mixed findings. A possible explanation for this might lie in learners' individual differences in preference for models' gender or models' age and which model learners find appropriate for explaining the task. This makes it difficult to determine which model characteristics are important regarding learning from video modeling examples, and in particular when conditions are different. Recent studies did not take model selection into account and therefore it would be interesting to investigate if learning and motivation is affected when students have the opportunity to choose their own model from a selection of models. Research in the field of APA already shows that providing learners with a choice of agent leads to better learning (Ozogul et al., 2013). However, when it comes to learning with video modeling examples, the question remains how model selection affects learning and motivation.

The Present Study

The purpose of this study was to investigate the effect of model selection when learning with video modeling examples on learning and motivation. In this study, participants watched video modeling examples about troubleshooting electrical circuits. Half of the participants were allowed to choose a model (choice condition), the other half was quasi-randomly assigned to a model (no choice condition). Participants could choose from four models: two adult males (AM1, AM2) and two adult females (AF1, AF2).

The hypothesis was made that if participants would have the opportunity to choose their own model, participants would show an increase of performance and motivation. Based on research findings showing that having the opportunity to choose stimulates feelings of autonomy which fosters motivation and leads to better learning (Deci & Ryan, 2000), and research findings that show choice of model might differ among participants due to individual differences in preference for model characteristics(Baylor et al., 2005), the prediction was made that students in the choice condition will choose a model they find appropriate for explaining the task, show more motivation and will perform better on the posttest than participants in the no-choice condition.

Given the task-appropriateness hypothesis, students may choose adult male models more often because trouble-shooting electrical circuits is perceived as a task more appropriate for men. Although earlier research (Hoogerheide, Loyens et al., 2016; Hoogerheide et al., 2018) shows no effect of models' gender on learning, research in the field of APA (Ozogul et al., 2013) suggests that choosing a model with the opposite gender, leads to better learning. Therefore, participants who choose a model with the opposite gender might perform better on the posttest.

Additionally, the effect of model selection was examined on two variables. First, invested mental effort during the learning phase and posttest was measured to provide information about the perceived mental effort of the conditions (Van Gog & Paas, 2008). Second, participants were asked to evaluate the quality of the explanations of the models. Based on research findings suggesting that the quality of explanations influences to what extent participants view the model as an expert, and which in turn might affect perceived task-appropriateness perceptions of the model (Hoogerheide, Loyens et al., 2016), the prediction was made that participants in the choice condition rate the explanation quality of the model they have chosen higher than participants in the no-choice condition.

Lastly, the effect of observers' and models' gender was explored. Although previous research showed no effect of models' gender, observers' gender did seem to make a difference. Males performed better than females which indicates that observers' gender may influence the results of this study (Hoogerheide et al., 2018). Furthermore, in previous research participants did not have the opportunity to choose a male or female model (Hoogerheide, Loyens et al., 2016; Hoogerheide et al., 2018). Model selection might lead to better learning according to research in the field of APA (Ozogul et al., 2013).

Methods

Participants and Design

A priori power analysis was conducted using G*Power version 3.1.9.5 (Faul, Erdfelder, Lang, & Buchner, 2007) to test the difference between two independent group means using a two-tailed test ($d = 0.5$, $\alpha = .05$). Results showed that a sample size of 128 participants was required to achieve a power of .80. Unfortunately, due to the outbreak of COVID-19 the desired number of participants was not reached. The final sample consisted of 73 participants between the age of 13 and 30 (26 males, 47 females; age: $M = 23.07$, $SD = 3.83$). A post hoc power analysis using G*Power version 3.1.9.5 (Faul et al., 2007) to test the difference between two independent group means using a two-tailed test showed an achieved power of .56 ($d = 0.5$, $\alpha = .05$).

The design of this research was a quasi-experimental between-subjects design with two conditions: choice condition and no choice condition. To ensure both conditions contained an equal number of participants and an equal distribution of gender, participants were quasi-random assigned to the choice condition (12 males, 24 females) and no choice condition (14 males, 23 females).

Materials

Materials used for this study are based on materials used in earlier studies (Van Gog, Kester, & Paas, 2011; Hoogerheide et al., 2014; Hoogerheide, Van Wermeskerken et al., 2016; Hoogerheide et al., 2018) and were presented in an online survey software tool named Qualtrics (www.qualtrics.com).

Pretest. The pretest tested prior knowledge consisting of seven open-ended questions about troubleshooting and parallel circuits principles. For example, "What is likely to happen if you do not measure current in one parallel branch?". In this way, any prior knowledge was made apparent and differences between the conditions could be ruled out if necessary.

Introductory text. An introductory text provided information about Ohm's law and the three different ways of using the formula (R = U/I; I = U/R; U = I*R). In addition, a drawing of an electrical circuit was added with an explanation of the components and abbreviations used.

Choice Manipulation. Before watching the video modeling examples, participants in the choice condition were given the opportunity to choose one of four models: AM1, AM2, AF1 and AF2. Participants in the choice condition received video modeling examples with their selected model while participants from the no choice condition were quasi-randomly assigned to one of the four models.

Video-modeling examples. Eight video modeling examples were created for this experiment, two examples per model. Each video consists of a model standing next to a screen presenting four PowerPoint-slides (see Figure 1). In the examples the model explains how to detect and resolve an error in an electrical circuit.

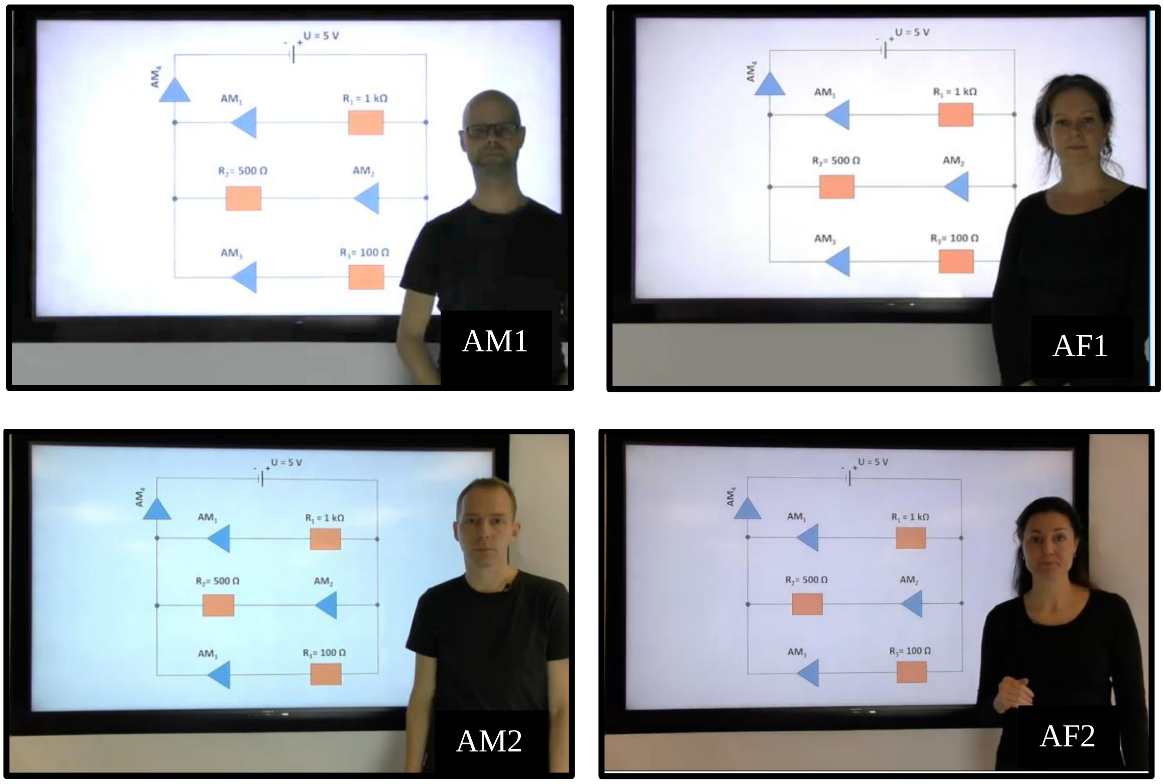

Figure 1. Adult male and adult female models.

In the first example (240s) the model started with a slide containing a drawing of three parallel branches. First, the model briefly explained the circuit and explained that the correct functioning of the circuit depends on the current in the three parallel branches and the total current and that it can be calculated using Ohm's law. Second, a slide with a smaller version of the same circuit accompanied with Ohm's law and the worked-out problem-solving steps of the calculation of the currents was presented. The model provided step-by-step information about how to calculate the current in each branch and the total current, using the voltage and resistance presented in the drawing. Next, a third a slide was presented with the measured current and faulty ammeter measurement. The model explained that the circuit is not functioning correctly because the (lower) current and resistance are not in balance. In order to let the circuit function correctly the resistance must be higher, and the model demonstrated how to calculate the actual resistance. During this explanation a fourth slide was presented with a smaller version of the same circuit with Ohm's law and the worked-out problem-solving steps. In the second example (244s) the models explained troubleshooting electrical circuits with the exact same procedure, but the example differs due to the use of different values and a different error (e.g. a parallel branch measured a higher current to which a lower resistance must be calculated).

All models were Dutch, around 40 years old and wore a black t-shirt. To make the content of the video modeling examples as identical as possible, an autocue was placed underneath the camera. The models had the opportunity to practice the entire process before recording the final version. During practice, models were instructed to make specific gestures to elements in the PowerPoint slides.

Posttest. The posttest consisted of two troubleshooting problems. The first problem was similar to the first problem in the video modeling example: a lower current indicates a higher resistance is needed and needs to be recalculated. The second problem contained two faults which are discussed in both video modeling examples (a branch with lower current and a different branch with higher current). By solving this problem, participants were reminded that the current (U) is expressed in volt (V), resistance (R) is expressed in Ohm (Ω) and power (I) is expressed in amperes (A) and they can be calculated using Ohm's law.

Motivation. Participants were asked to rate their motivation after the learning phase using the Dutch version of the Intrinsic Motivation Inventory (IMI) of Ryan & Deci (2000) measuring four clusters related to motivation: interest/enjoyment, perceived competence, perceived choice and pressure/tension (see appendix F). This questionnaire consisted of 22 statements which could be answered using a 7-point Likert scale, ranging from 1 (this does not apply to me at all) to 7 (this fully applies to me). For example, "I really enjoyed doing this activity.".

Invested mental effort. Mental effort during studying the video modeling examples and the posttest, was rated using a 9-point rating scale of Paas (1992). Participants rated their invested mental effort ranging from 1(very, very low effort) to 9 (very, very high effort).

Perceived explanation quality. The perceived explanation quality of the models was rated after studying both video modeling examples. Participants rated the explanation quality on a 9-point scale ranging from 1 (very, very bad quality) to 9 (very, very good quality) (Hoogerheide, Loyens et al., 2016).

Procedure

Via social media and e-mail an introduction text and an anonymous link was distributed to recruit participants. The experiment was conducted online and consisted of one session with an average duration of forty minutes. Participants were recommended to participate by computer. After entering the Qualtrics environment and reading a short introduction text, the experiment continued consisting of four blocks. The first block started with an informed consent form, followed by a demographic questionnaire and pre-test. The second block had a duration of two minutes where participants were provided with an introductory text which they needed to read carefully. Afterwards, half of the participants had the opportunity to choose from which model they wanted to receive lessons about troubleshooting electrical circuits. Pictures of the four models were displayed on the screen. The other half of participants were quasi-random assigned to one of the four models and moved on to block three. Block three consisted of the two video modeling examples followed by the invested mental effort scale. At the end of the block, participants were provided with the learning enjoyment and perceived explanation quality scale. Block four started with a posttest followed by the invested mental effort scale and motivation scale.

Analysis

Test performance was scored based on straightforward coding schemes of Van Gog et al., (2011). Participants could receive ten points for the pretest and eight points for the posttest. The posttest consisted of two tasks. For task one, points were given for calculating the value of the ammeters, indicating the faulty resistor and indicating its correct value. For task two, the same scoring applied with three extra points for indicating the second faulty resistor and calculating the resistance correctly. Incomplete or partially correct answers were rated with half a point.

Motivation was scored based on the coding schemes of the Dutch version of the IMI (Ryan & Deci, 2000). Scores for items 2, 9, 11, 14, 19, 21 were reversed to make sure that a higher score indicated a higher level of motivation. To calculate the total level of motivation per participant, the average of participants' scores on all 22 items was taken. Participants' score per motivation subscale was calculated by averaging participants' items scores for the items on each subscale. Items 1, 5, 8, 10, 14, 17 and 20 measured interest and enjoyment, items 4, 7, 12, 16 and 22 measured perceived competence, items 3, 11, 15, 19 and 21 measured perceived choice and items 2, 6, 9, 13 and 18 measured pressure and tension.
In addition, averages of invested mental effort during studying the video modeling examples and during the post-test and perceived explanation quality were measured[1].

Results

First, a preliminary analysis was conducted to check whether prior knowledge differed between the two conditions by means of an independent samples *t*-test. Due to violation of the normality assumption a Mann-Whitney *U* test was conducted which indicated no significant difference on pre-test score between the choice condition (*Mean Rank* = 41.11,

[1] This study was part of a research plan which also measured perceived task-appropriateness, self-efficacy, perceived competence and learning enjoyment. These variables were not important for this study and were therefore left out of consideration.

n = 36) and no-choice condition (*Mean Rank* = 33.00, *n* = 37), *U* = 518.00, *z* = -1.68, *p* = .09, two tailed, *r* = .20. Second, it was investigated whether there was a difference between the two male and two female models by means of independent samples *t*-tests for posttest and motivation. Due to violation of the normality assumption for posttest score a Mann Whitney *U* test was conducted and revealed no significant difference between both male and female models. For motivation there was no significant difference between the two male models, but results revealed a significant difference between the two female models. Watching AF2 led to higher motivation (*M* = 3.77, *SD* = 0.90) than watching AF1 (*M* = 3.07, *SD* = 0.84), *t*(37) = -2.34, *p* = .03, two-tailed, *d* = 0.80. Since this effect can be described as large, the analysis plan for the explorative questions had to be adjusted.

As for observers' gender and models' gender the original plan was to conduct two 2 x 2 x 2 ANOVAs on posttest score and motivation, with models' gender (Male, Female), observers' gender (Male, Female) and condition (Choice, No-choice) as between subject factors. Unfortunately, the 2 x 2 x 2 ANOVA on motivation could not be performed, because the two female models had significant differential effects on motivation. Therefore, the effect of observers' gender on motivation was analysed with an independent samples *t*-test. In addition, because results for motivation based on models' gender could not be combined, the effect of different models on motivation was further explored using one-way ANOVAs within the conditions.

Then, analysis of the main research question and explorative questions was conducted. Unless otherwise indicated, the analysis consisted of independent samples *t*-tests. If the assumption for normality could not be met, a Mann Whitney U test was conducted. Table 1 shows Mean and SD scores on all variables per condition.

Table 1
Mean (SD) of test performance, invested mental effort, explanation quality and motivation per condition.

	Choice condition	No-choice condition
Pretest performance (range 0-10)	2.60 (2.17)	1.70 (1.97)
Posttest performance (rang 0-8)	4.31 (2.96)	4.31 (2.83)
Invested mental effort in learning phase (range 1-9)	4.18 (2.17)	4.11 (2.25)
Invested mental effort in Posttest (range 1-9)	4.97 (2.50)	4.57 (2.48)
Explanation quality (range 1-9)	6.50 (0.81)	6.32 (0.75)
Motivation (range 1-7)	3.71 (0.95)	3.72 (0.95)

Posttest Performance

Our first main question was whether participants attained better posttest scores when they had the opportunity to choose a model. Results indicated that the posttest performance of participants in the choice condition (*Mean Rank* = 37.07, $n = 36$) was not significantly higher than those of participants in the no-choice condition (*Mean Rank* = 36.93, $n = 37$), $U = 663.50$, $z = -.03$, $p = .50$, $r < .01$.

Motivation

Our second main question was whether the level of motivation for participants in the choice condition was higher than participants in the no-choice condition. Results show no significant difference $t(71) = 0.04$, $p = .97$, $d = 0.01$. Participants in the choice condition reported the same level of motivation ($M = 3.71$, $SD = 0.94$) as participants in the no-choice condition ($M = 3.72$, $SD = 0.95$). Furthermore, analysis per motivation subscale (interest/enjoyment, perceived competence, perceived choice and pressure/tension) showed no significant difference between the conditions (all *p*s > .40).

Choice Manipulation

Within the choice condition two binomial tests were conducted. As for the task-appropriate hypothesis, the choice for an adult male model was checked for significant deviation from what would be expected by chance (i.e., 50%). Results showed that an adult male model was not picked significantly more often than an adult female model (47%, $p = .43$). However, the option to choose a model with the same gender was picked significantly more than choosing a model with the opposite gender (69%, $p = .03$). Then, a potential difference on posttest score or motivation between participants choosing the same gender model versus participants choosing the opposite gender model was explored. Results indicated no significant difference for posttest scores between participants choosing a model with the same gender (*Mean Rank* = 18.82, $n = 25$) or a model with the opposite gender (*Mean Rank* = 17.77, $n = 11$), $U = 129.50$, $z = -.28$, $p = .78$, two-tailed, $r = .05$. Same results were found for motivation: no significant difference was found for the level of motivation between participants choosing the same gender model ($M = 3.74$, $SD = 0.88$) or participants choosing the opposite gender model ($M = 3.66$, $SD = 1.09$), $t(34) = 0.24$, $p = .81$, two-tailed, $d = 0.08$. Furthermore, analysis per motivation subscale showed no significant difference between participants choosing the same gender model or the opposite gender model (all *p*s $> .76$).

Invested Mental Effort

Results revealed no significant difference for invested mental effort during the posttest between participants in the choice condition (*Mean Rank* = 38.47, $n = 36$) and no-choice condition (*Mean Rank* = 35.57, $n = 37$), $U = 613.00$, $z = -.59$, $p = .56$, two tailed, $r = .07$. Invested mental effort during the study phase showed the same results: there was no significant difference between participants in the choice condition ($M = 4.17$, $SD = 2.17$) and no-choice condition ($M = 4.11$, $SD = 2.25$), $t(71) = -.11$, $p = .91$, two-tailed, $d < 0.01$.

Explanation Quality

The analysis on the rated explanation quality showed no significant difference between participants in the choice condition (*Mean Rank* = 38.33, $n = 36$) and no-choice condition (*Mean Rank* = 35.70, $n = 37$), $U = 618.00$, $z = -.58$, $p = .56$, two-tailed, $r = .07$.

Gender

This is an explorative analysis to investigate the effect of observers' gender and models' gender on posttest and motivation. Because there was no significant difference between the two male and two female models on the posttest score, posttest score analysis was attempted using a 2 x 2 x 2 ANOVA with models' gender (Male, Female), observers' gender (Male, Female) and condition (Choice, No-choice) as between subject factors. However, due to violation of the normality assumption this analysis could not be conducted. Subsequently a 2 x 2 ANOVA for posttest score was conducted with models' gender (Male, Female) and observers' gender (Male, Female) as between subject factors and again the assumption for normality could not be met. Therefore, analysis of the effect of observers' gender on posttest score was conducted by means of a Mann Whitney *U* test. The effect of different models on posttest score was investigated by means of one-way ANOVAs within the conditions. Results for both post-test score and motivation are shown below, divided into the effect of observers' gender and the effect of model.

Observers' Gender

First, the effect of observers' gender on posttest performance was investigated. Results showed no significant difference between males (*Mean Rank* = 38.02, $n = 26$) and females (*Mean Rank* = 36.44, $n = 47$), $U = 584.50$, $z = -.310$ (corrected for ties), $p = .76$, two-tailed, $r = .04$. Second, the interaction between motivation and observers' gender was investigated. A significant result was found, with the males ($M = 4.10$, $SD = 0.89$) reporting a higher level of motivation than the females ($M = 3.51$, $SD = 0.90$), $t(71) = -2.69$, $p = .01$, two-tailed, $d = 0.67$. This effect can be described as medium. Third, further exploration per motivation subscale revealed a main effect of observers' gender on perceived competence. Results revealed that males (*Mean Rank* = 44.42, $n = 26$), reported a higher level of perceived competence than females (*Mean Rank* = 32.89, $n = 47$), $U = 418.00$, $z = -2.23$

(corrected for ties), p = .03, two-tailed, r = .26. This effect can be described as small. Motivation subscales interest/enjoyment, perceived choice and pressure/tension showed no significant difference between the conditions (all ps > .05).

Model

As a final part of the analysis, the effect of different models on both posttest and motivation was explored within the choice condition and no-choice condition by means of one-way ANOVAs. Unfortunately, the assumption of normality was violated for posttest score in the no-choice condition. Therefore, a Kruskal Wallis ANOVA for posttest score in the no-choice condition was conducted. Table 2 shows Mean and SD scores of posttest and motivation for different models per condition.

Table 2

Mean (SD) of posttest performance and motivation for models per condition.

	Choice condition				No-choice condition			
	AM1 (n = 13)	AM2 (n = 4)	AF1 (n = 4)	AF2 (n = 15)	AM1 (n = 7)	AM2 (n = 10)	AF1 (n = 10)	AF2 (n = 10)
Posttest performance (range 0-8)	4.23 (3.19)	5.25 (3.77)	6.50 (1.91)	3.53 (2.64)	2.57 (2.70)	5.75 (1.78)	3.90 (3.57)	4.50 (2.51)
Motivation (range 1-7)	3.88 (1.12)	3.64 (0.62)	3.66 (1.17)	3.60 (0.83)	3.69 (0.81)	4.35 (0.63)	2.84 (0.60)	4.01 (0.99)

Choice condition. The ANOVAs on both posttest score and motivation showed no significant effect of model, indicating that choice of model had no influence on participants posttest score $F(3,32) = 0.71$, p = .89, η^2 = .02, or motivation $F(3,32) = 1.24$, p = .31, η2 = .10.

No-choice condition. Results of a Kruskal-Wallis ANOVA showed no significant difference between the posttest score when participants received AM1 (*Mean Rank* = 12.93), AM2 (*Mean Rank* = 23.45), AF1 (*Mean Rank* = 18.30) or AF2 (*Mean Rank* = 19.50), $H(3)$ = 4.17, p = .24, η^2 = .12. The ANOVA for motivation was statistically significant, indicating that the level of motivation was influenced by the model participants received, $F(3,33) = 7.04$, $p < .01$, η2 = .39. Post hoc analyses with Tukey's HSD (using an α of .05) revealed that participants who received AF1 (M = 2.84, SD = 0.60) showed a significantly lower score on motivation than participants who received AM1 (*M* = 4.35, *SD* = 0.63) or AF2 (*M* = 4.01, *SD* = 0.99). Effect sizes for these two comparisons can be described as large, *d* = 1.53 and *d* = 1.18.

In addition, the effect of received model on motivation subscales was further explored with one-way ANOVAs. The ANOVA for subscale interest and enjoyment was statistically significant, indicating that the level of interest and enjoyment was influenced by the model participants received, $F(3,33) = 4.27$, $p = .01$, $\eta 2 = .28$. Post hoc analyses with Tukey's HSD (using an α of .05) revealed that participants who received AM2 (*M* = 4.47, *SD* = 1.06) showed a significantly higher interest and enjoyment than participants who received AF1 (*M* = 2.46, *SD* = 1.01).

The ANOVA for subscale perceived competence was statistically significant, indicating that the level of perceived competence was influenced by the model participants received, $F(3,33) = 3.49$, $p = .03$, $\eta 2 = .24$. Post hoc analyses Tukey's HSD (using an α of .05) revealed that participants who received AM2 (*M* = 4.80, *SD* = 1.40) showed a significantly higher perceived competence than participants who received AF1 (*M* = 2.64, *SD* = 1.43).

Results of a Kruskal-Wallis ANOVA showed significant difference for perceived choice when participants received AM1 (*Mean Rank* = 16.93), AM2 (*Mean Rank* = 22.55), AF1 (*Mean Rank* = 11,20) or AF2 (*Mean Rank* = 24.70), $H(3) = 9.133$, $p = .03$, $\eta^2 = .26$.

The ANOVA for subscale pressure and tension showed no significant difference between the models ($p = .73$).

Discussion

This study examined whether model selection during learning with video modeling examples had an effect on learning and motivation. Because prior research on MOS and task-appropriateness hypothesis in the field of video modeling examples led to mixed findings and did not incorporate the opportunity to choose a model, the option to choose a model was added to see whether this would lead to different results.

Before conclusions can be drawn from the results, it is important to further explain the situation in which this study was conducted. Initially, this study was planned to be conducted in a secondary school with secondary education students. A priori power analysis revealed that a total sample size of 128 participants was required to achieve a power of .80. Unfortunately, due to the outbreak of COVID-19 this study was adapted to an online version at which everyone between the age of 13 and 30 could participate. Furthermore, the number of required participants has not been achieved leading to a lower power of .56. Therefore, all results should be interpreted with some caution.

Given that half of the participants were allowed to choose their own model which fosters feelings of autonomy, the prediction was made, in line with SDT (Deci & Ryan, 2000), that participants in the choice condition would show a higher level of motivation which in turn has a positive effect on learning. In contrast with what was expected, model selection did not significantly affect participants' test performance or motivation. Furthermore, model selection did not affect participants invested mental effort or the perceived quality of the explanations. Thus, our findings suggest that there would be no need to take model selection into account when designing video modeling examples. These findings are in line with the prior study of Baylor et al. (2005) but not with the study of Ozogul et al. (2013), who found that model selection did affect learning outcomes.

A possible explanation for this might be that model selection on its own did not foster feelings of autonomy enough to stimulate intrinsic motivation for effective learning with video modeling examples. Feelings of autonomy might be more affected if participants are allowed to make multiple choices during learning with video modeling examples. For example, in future research, participants might be provided with multiple choices for the appearance of the model. In this research, the choice of model was based on gender, but other model characteristics might be more relevant for learning (Renkl, 2017). For example, research of Baylor et al. (2005) revealed no effect of gender on model selection but students tended to choose an agent with the same ethnicity. Therefore, it could be interesting to further explore the effect of other model characteristics such as ethnicity when using model selection for learning. Nevertheless, the question remains whether participants' choice of model is most beneficial for their learning. Participants might choose a model with the same ethnicity, but a model with a different ethnicity might lead to better learning. Another way to foster feelings of autonomy could be to give participants control over the learning task. Troubleshooting electrical circuits could be divided into small learning tasks that participants can choose from or could be combined with another physics topic that is part of the curriculum. In this way, participants choosing a learning tasks experience more ownership over what they learn which is beneficial for learning (Corbalan et al., 2006; Cordova & Lepper, 1996; Deci & Ryan, 2000).

Concerning model selection in combination with task-appropriateness, it was predicted that the male models might be chosen more often. Results showed no significant difference between choosing a male or female model, but participants were more likely to choose a model with the same gender. An explanation for the result that male models were not chosen more often than female models, might lie in the male to female ratio of observers. There were twice as many women in the choice condition than men (24 females versus 12 males) and results might be biased. However, participants choosing the same gender model showed no significant difference in learning outcomes compared to participants choosing the opposite gender model. Which is in line with findings of earlier research in de the field of APA (Baylor, Shen, & Huang, 2003; Ozogul et al., 2013). Although it seems gender matching has no effect on learning, it appears that learners prefer learning from the same gender which can be seen as support for MOS: learners preferred to learn from a same gender model. However, choosing the same model did not affect the effectiveness of modeling on self-efficacy or learning which does not support MOS. These results might be due to the low power. Therefore, further research could focus on the effect of choosing the same gender model or opposite gender model on learning, preferably with a larger sample. Furthermore, it would be interesting to investigate whether learning with matched gender could have an effect on learning in the long term with multiple learning sessions. Multiple learning sessions may lead to an effect and future research may investigate whether this effect is permanent or whether an adjustment process is occurring.

As for observers' gender, results showed that male participants did not report a higher test performance than females which is not in line with research of Hoogerheide et al. (2018). This difference in results might be caused by the small group of males participating. On the other hand, a significant difference for observers' gender in the level of motivation was found, indicating that males score higher than females on motivation in general and on the subscale perceived competence. These findings are in line with previous studies at which males scored higher on perceived competence (Hoogerheide, Loyens et al., 2016; Hoogerheide et al., 2018). In addition, these findings are in line with the task-appropriate hypothesis. Concerning that troubleshooting electrical circuits is perceived as a task more appropriate for men, male participants might find themselves more suited to perform physics tasks leading to a higher score on motivation and perceived competence than females.

In addition to observers' gender, the intention was to further investigate the effect of models' gender on posttest score and motivation and if results differed between the conditions. Unfortunately, this was not possible due to the fact that the two female models led to difference in motivation of participants. Further analysis per model within conditions (Choice, No-Choice) revealed no effect of model on posttest score and motivation in the choice condition. However, participants in the no-choice condition watching adult female model one reported a significantly lower level of motivation than participants watching adult female model two or adult male model one. Different results per model might be caused by individual difference in preference for model characteristics, leading to difference in motivation to pay attention. Results in the field of APA suggests that the perceived attractiveness of the APA might affect students' attention to an APA and in turn might affect learning (Moreno & Flowerday, 2006).

The strength of this study was providing participants with a choice of model. A potential limitation of this study is that model selection was intentionally based on gender to further explore the task appropriateness hypothesis. However, results revealed a difference among others between adult female model number one and adult female model number two on motivation. Potentially there could be an effect of models' gender on model selection, but this could not be investigated. Another limitation is that the level of motivation was only measured after the learning phase. Therefore, it was not possible to check if the level of motivation before the learning phase differed among the participants and might influence the results. Lastly, due to practical constraints, this study cannot provide a comprehensive review of the use of model selection when learning with video modeling examples in secondary education. Troubleshooting electrical circuits is a subject in the curriculum of secondary education but the majority of participants were no longer in secondary school. Most participants were between 20 and 30 years old (N = 62). This makes it difficult to generalize results to secondary education students but nevertheless it gives us more insight in how participants between the age of 20 and 30 deal with model selection.

As mentioned before, future research could further investigate the effect of other model characteristics rather than gender on motivation and learning when selecting a model. Furthermore, it would be interesting to further explore the effect of learning with matched gender with a larger sample size. In addition, investigating whether multiple learning sessions with model selection affects motivation and learning might reveal interesting results. Finally, it is strongly recommended to replicate this study with the intended target group: secondary education students. In this way, results are more in line with previous studies on learning with video modeling examples and results can be generalized for secondary education students who show an increase of using video modeling examples.

In sum, a first step has been made to investigate the effect model selection when learning with video modeling example. The use of model selection gained attention in the field of APA (Ozogul et al., 2013) but might also be important when learning with video modeling examples, especially to further investigate the task appropriateness hypothesis. In addition, this study contributes to an understanding of what design guidelines are effective for learning with video modeling examples especially since the use of video modeling examples has increased over the past few years.

Model selection does not seem to be an important factor when learning with video modeling examples which you would not expect based on the task-appropriateness hypothesis. However, future research is recommended to further investigate whether these findings are generalizable in secondary education and if selecting a model based on other model characteristics would emerge different results.

References

Bandurra, A. (1994). Self-efficacy. In V.S. Ramachaudran (ed.), *Vol. 4. Encyclopedia of human behaviour* (pp. 71-81). New York: Academic Press.

Berger, S. M. (1977). Social comparison, modeling, and perseverance. In J. M. Suls & R. L. Miller (Eds.), Social comparison processes: Theoretical and empirical perspectives (pp. 209–234). Washington, DC: Hemisphere.

Bandura, A., Ross, D., & Ross, S. A. (1963). Vicarious reinforcement and imitative learning. *Journal of Abnormal and Social Psychology*, *67*(6), 601–607. https://doi.org/10.1037/h0045550

Baylor, A. L. D. P., Ave, L., & Park, I. (2005). *The Impact of Pedagogical Agent Image on Affective Outcomes.* International Conference on Intelligent User Interfaces, San Diego, CA. Retrieved from http://amybaylor.com/Articles/IUI2005.pdf

Baylor, A., Shen, E., & Huang, X. (2003). *Which Pedagogical Agent do Learners Choose? The effect of Gender and Ethnicity.* In E-learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (pp. 1507–1510). Retrieved from https://www.learntechlib.org/primary/p/12159/

Bétrancourt, M., & Benetos, K. (2018). Why and when does instructional video facilitate learning? A commentary to the special issue "developments and trends in learning with instructional video". *Computers in Human Behavior, 89*. https://doi.org/10.1016/j.chb.2018.08.035

Buritica, R., Eppinger, J. M., Schuck, B., Heekeren, N. W., Hauke, R., & Li, S. C. (2016). Electrophysiological correlates of observational learning in children. *Developmental Science*, *19*(5), 699–709. https://doi.org/10.1111/desc.12317

Bussey, K., & Bandura, A. (1984). Influence of gender constancy and social power on sex-linked modeling. *Journal of Personality and Social Psychology*, *47*(6), 1292–1302. https://doi.org/10.1037/0022-3514.47.6.1292

Corbalan, G., Kester, L., & Van Merriënboer, J. J. G. (2006). Towards a personalized task selection model with shared instructional control. *Instructional Science*, *34*(5), 399–422. https://doi.org/10.1007/s11251-005-5774-2

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, *88*(4), 715–730. https://doi.org/10.1037//0022-0663.88.4.715

Davidson, E. S., & Smith, W. P. (1982). Imitation , Social Comparison , and Self-Reward. *Child Development*, *53*(4), 928–932.

Deci, E. L., & Ryan, R. M. (1987). The support of autnomy and the control of behaviour. *Jorunal of Personality and Social Pyschology*, *53*(6), 1024–1037.

Deci, E. L., & Ryan, R. M. (2000). The " What " and " Why " of Goal Pursuits : Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, *11*(4), 37–41. https://doi.org/10.1207/S15327965PLI1104

Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An Instrument to Assess Adults' Orientations Toward Control Versus Autonomy With Children: Reflections on Intrinsic Motivation and Perceived Competence. *Journal of Educational Psychology*, *73*(5), 642–650.

Duit, R., & Von Rhöneck, C. (1998). Learning and understanding key concepts of electricity. In a. Tiberghien, E. L. Jossem, & J. Barajos (Eds.), Connecting research in physics education with teacher education. International Commision on Physics Education.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fiorella, L., van Gog, T., Hoogerheide, V., & Mayer, R. E. (2017). It's all a matter of perspective: Viewing first-person video modeling examples promotes learning of an assembly task. *Journal of Educational Psychology*, Vol. 109, pp. 653–665. https://doi.org/10.1037/edu0000161

Fry, J. P. (1972). Interactive relationship between inquisitiveness and student control of instruction. *Journal of Educational Psychology*, *63*(5), 459–465. https://doi.org/10.1037/h0033237

Hicks, D. J. (1965). Imitation and retention of film-mediated aggressive peer and adult models. *Journal of Personality and Social Psychology*, Vol. 2, pp. 97–100. https://doi.org/10.1037/h0022075

Hoogerheide, V., Loyens, S. M. M., & Van Gog, T. (2014). Comparing the effects of worked examples and modeling examples on learning. *Computers in Human Behavior*, *41*, 80–91. https://doi.org/10.1016/j.chb.2014.09.013

Hoogerheide, V., Loyens, S. M. M., & Van Gog, T. (2016). Learning from video modeling examples: does gender matter? *Instructional Science*, *44*(1), 69–86. https://doi.org/10.1007/s11251-015-9360-y

Hoogerheide, V., Van Wermeskerken, M., Loyens, S. M. M., & Van Gog, T. (2016). Learning from video modeling examples: Content kept equal, adults are more effective models than peers. *Learning and Instruction*, *44*, 22–30. https://doi.org/10.1016/j.learninstruc.2016.02.004

Hoogerheide, V., Van Wermeskerken, M., Van Nassau, H., & Van Gog, T. (2018). Model-observer similarity and task-appropriateness in learning from video modeling examples: Do model and student gender affect test performance, self-efficacy, and perceived competence? *Computers in Human Behavior*, *89*, 457–464. https://doi.org/10.1016/j.chb.2017.11.012

Jakubczak, L. F., & Walters, R. H. (1959). Suggestibility as dependency behavior. *The Journal of Abnormal and Social Psychology*, Vol. 59, pp. 102–107. https://doi.org/10.1037/h0040627

Kinzie, M. B., & Sullivan, H. J. (1989). Continuing motivation, learner control, and CAI. *Educational Technology Research and Development*, *37*(2), 5–14. https://doi.org/10.1007/BF02298286

Lahey, G. F. (1973). *Post Lesson Remediation and Student Control of Branching in Computer Based Training.*

Moreno, R., & Flowerday, T. (2006). Students' choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology*, *31*(2), 186–207. https://doi.org/10.1016/j.cedpsych.2005.05.002

Ozogul, G., Johnson, A. M., Atkinson, R. K., & Reisslein, M. (2013). Investigating the impact of pedagogical agent gender matching and learner choice on learning outcomes and perceptions. *Computers and Education*, *67*, 36–50. https://doi.org/10.1016/j.compedu.2013.02.006

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, Vol. 84, pp. 429–434. https://doi.org/10.1037/0022-0663.84.4.429

Renkl, A. (2017). Instruction based on examples. In R. E. Mayer, & P. A. Alexander (Eds.), Handbook of research on learning and instruction (2nd ed.). New York, NY: Routledge (in press).

Ryan, R.M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Ryan, R. M., & Connell, J. P. (1989). Perceived Locus of Causality and Internalization: Examining Reasons for Acting in Two Domains. *Journal of Personality and Social Psychology*, *57*(5), 749–761. https://doi.org/10.1037/0022-3514.57.5.749

Schunk, D. H. (1987). Peer Models and Children's Behavioral Change. *Review of Educational Research*, *57*(2), 149–174. https://doi.org/10.3102/00346543057002149

Sheldon, K. M., & Elliot, A. J. (1999). Goal Striving, Need satisfaction, and Longitudinal Well-being: The Self-Concordance Model. *Journal of Personality and Social Psychology*, *76*(7), 482–497. Retrieved from papers2://publication/uuid/6925367A-4F11-4698-A182-CE7C28C63B89

Tafarodi, R. W., Milne, A. B., & Smith, A. J. (1999). The confidence of choice: Evidence for an augmentation effect on self-perceived performance. *Personality and Social Psychology Bulletin, 25*, 1405-1516.

Van Der Meij, H., & Van Der Meij, J. (2014). A comparison of paper-based and video tutorials for software learning. *Computers and Education*, *78*, 150–159. https://doi.org/10.1016/j.compedu.2014.06.003

Van Gog, T., Kester, L., & Paas, F. (2011). Effects of worked examples, example-problem, and problem-example pairs on novices' learning. *Contemporary Educational Psychology*, *36*(3), 212–218. https://doi.org/10.1016/j.cedpsych.2010.10.004

Van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist*, *43*(1), 16–26. https://doi.org/10.1080/00461520701756248

Van Wermeskerken, M., & van Gog, T. (2017). Seeing the instructor's face and gaze in demonstration video examples affects attention allocation but not learning. *Computers and Education*, *113*. https://doi.org/10.1016/j.compedu.2017.05.013

Wolters, C. A. (2003). Conceptualizing the Role and Influence of Student- Teacher Relationships on Children ' s Social and Cognitive Development. *Educational Psychologist*, *38*(4), 207–234. https://doi.org/10.1207/S15326985EP3804

Zimmerman, B. J. (2002). Achieving academic excellence: A self-regulatory perspective. In *The pursuit of excellence through education*.

Zuckerman, M., Porac, J., Lathin, D., Smith, R., & Deci, E. L. (1978). Zuckerman_ImportanceofSD_PersonlaityandSocialPsychology_(1978).pdf. *Personality and Social Psychology Bulletin*, *4*(3), 443–446.

Appendix A: FETC form

Section 1: Basic Study Information

1. Name student:

Noortje Pater

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

Tim van Marlen

3. Title of the thesis (plan):

The effect of model selection on learning and motivation when learning with video-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.

The data collection will take place at 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 (I think this is a reverser question, I continue with question 7)

7. Where is the data stored?

The collection data will be gathered in an online survey tool named Qualtrics (www.qualtrics.com). In addition, data will be stored in a personal folder in Your Data (YoDa). According to the laws of storing research data from Utrecht University, the retention period for raw research data is at least 10 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. For the purpose of this study we reive participants student number and gender. If student numbers are not available we will receive the names of the participants. This data will be used for quasi-random assignment of participants to the conditions in order to ensure both conditions contained an equal number of participants and an equal distribution of gender.

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

The research leaders (Noortje Pater and Jonneke van Oijen) as well as the supervisor (Tim van Marlen) have the key to permit re-identification.

Section 3: Participants

11. What age group is included in your study?

This study includes participants between the ages of 13-30 years.

12. Will be participants that are recruited be > 16 years? Yes/~~No~~
~~13.~~ Will participants be mentally competent (wilsbekwam in Dutch)? Yes/~~No~~
14. Does the participant population contain vulnerable persons? (e.g., incapacitated, children, mentally challenged, traumatized, pregnant) ~~Yes~~/No

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 into the effects of characteristics of video modeling examples for this age group. The study contains a physic instruction about troubleshooting electrical circuits which is part of the curriculum of the participants. When using older participants, we would not be able to make predictions about the targeted age group, and the older participants probably already have prior knowledge which influences the results.

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

In general, this research has hardly any risks for the participants. A possible risk of the study is that students experience stress of insecurity during the study because they want to do well. They can also get upset if they don't get the model they prefe or experience time pressure. Finally, it must be mentioned that we take (valuable) time from the participants, which can be identified as a risk for the participant.

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

To minimize risk (or burden), participants and parents receive all needed information about the experiment so they know what they can expect. In addition, the research leaders will be present throughout the investigation and will try to absorb any risk and try to stick tot he time limit.

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

About fifty 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?

Participants do not receive any form of compensation for their effort.

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

This research provides more knowledge in the field of learning with video modeling examples. Learning with video modeling examples is also increasingly being applied in (secondary) education, so students can also benefit from this in the future. This study contributes to more clarity into what makes instructional videos effective with witch design guidelines. In my opinion, the burden on the participants is minimal and possible potential scientific contribution maximal. In addition, the physic topic treated in the experiment is in line with the curriculum of the participants.

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.)

The numbers of participants is 128. A g-power analysis is conducted with a moderate effect size d = 0.25, df = 1, number of groups: 2, number of covariates: 1. The current power of this analysis is 0.8.

It is difficult to achieve 128 participants fort his research. Therefore, we use a general directive and try to recruit a minimum of 30 participants per condition, which means a total sample size of 60 participants.

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

The participants fort his research are acquired by means of a recruitment letter that is sent to secondary education schools and an information letter sent to parents (see appendix C and D of the research plan).

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

Ultimately, the school decides whether they want to participate in the research, assuming that the students (participants) will participate in this experiment.

The schools we approach can decide until the first of April if they want to participate in this experiment. If so, parents of the participants receive an passive consent form, a week in advance of the experiment.

The participants are informed prior to the investigation and subsequently have the option to decide not to participate in this experiment. Participant need to fill in an active consent form in advance of 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.

Prior tot he experiment, participants and parents receive all needed information about the experiment. Subsequently parents receive a passive consent form. Participants fill in a active consent in the online survey tool qualtrics. This informed consent forms can be found as Appendix D and appendix E in the research plan.

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. Class participation is determined by the school, but students can stop participating at any time during the experiment.

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?

Yes, participants can contact the course coördinator. E-mail: edu.acma.thesis@uu.nl

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

Yes, participant can contact the complaints officer. E-mail: klachtenfunctionaris-fetcsocwet@uu.nl

Section 4: Data management

29. Who has access to the data and who will be responsible for managing (access to) the data?

The researches of this experiment (Noortje Pater and Jonneke van Oijen) and their supervisor (Tim van Marlen).

30. What type of data will you collect or create? Please provide a description of the instruments.

Data will be collected with an online survey tool named qualtrics. We will collect the following data; age, gender, model preference, pre test and post test performance, motivation, explanation quality of the model, invested mental effort and learning

enjoyment.

31. Will you be exchanging (personal) data with organizations/research partners outside the UU?

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

32. If so, will a data processing agreement be made up?

Yes / No
If yes: Please attach the agreement.
If no: Please explain.

n.v.t.

33. Where will the data be stored and for how long?

The data will be stored in Your Data (YoDa) for at least 10 years.

34. Will the data potentially be used for other purposes than the master's thesis? (e.g., publication, reporting back to participants, etc.)

No.

35. Will the data potentially be used for other purposes than the master's thesis? (e.g., publication, reporting back to participants, etc.)

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

Appendix B: Recruitment letter I

Lesgeven in het Corona-tijdperk

Beste heer, mevrouw,

Opeens zien scholen zich genoodzaakt om het onderwijs voor hun leerlingen digitaal aan te bieden. Een trend die al langer speelt, maar nu moeten scholen plotseling gedwongen omschakelen. Vaak zonder dat er al lesmateriaal digitaal beschikbaar is. Graag brengen wij daarom deze online les over parallelle elektrische schakelingen bij u onder de aandacht.

In deze les wordt er gebruik gemaakt van videovoorbeelden: een video waarin iemand iets uitlegt en voordoet. Ondanks dat dit een effectieve manier blijkt te zijn om te leren, is er weinig bekend over 'wat werkt' in deze video's. Voor onze master in onderwijswetenschappen aan de Universiteit Utrecht doen wij, Jonneke van Oijen en Noortje Pater, hier daarom onderzoek naar.

Digitale les Natuurkunde
Deze les maakt deel uit van het curriculum Natuurkunde voor HAVO en VWO klassen. Meestal komt dit in klassen 3 of 4 HAVO/VWO aan bod, maar kan ook in 2VWO al aangeboden worden. De les is eenvoudig via onderstaande link uit te zetten aan uw leerlingen als volwaardige les, of als huiswerk, en biedt zo uw organisatie een verlichting van de werkdruk.

Deelname via link
Wij bieden dit lesmateriaal aan, als onderdeel van een onderzoeksprogramma naar het leren van videovoorbeelden. Meer informatie over dit onderzoek en de wijze waarop wij met de data omgaan, leest u in de bijlage. De link die uw leerlingen kunnen aanklikken is:
https://tinyurl.com/lerenvanvideovoorbeelden

Wij raden u aan om uw leerlingen in de Chrome browser te laten werken, omdat de site daar het beste in draait. Deelname aan dit onderzoek is geheel vrijwillig en duurt ongeveer 50 minuten. Onder de deelnemers wordt een bol.com bon verloot met een waarde van 20 euro.

Vragen?
Wij vertrouwen erop dat u met de brief in de bijlage een goed idee krijgt over wat dit onderzoek inhoudt. Heeft u vragen? Mail ons dan gerust op onderstaande e-mailadressen. Wij hopen op uw medewerking.

Met vriendelijke groet,
Jonneke van Oijen en Noortje Pater

j.vanoijen@students.uu.nl / n.pater@students.uu.nl

Appendix C: recruitment letter II

GEZOCHT: deelnemers voor wetenschappelijk onderzoek.

Ben jij of ken jij iemand die ons verder kan helpen?

Voor onze master thesis zijn wij, Jonneke en Noortje, opzoek naar deelnemers tussen de 13 en 30 jaar oud!

Op dit moment doen wij onderzoek naar het leren met behulp van videovoorbeelden: een video waarin iemand iets uitlegt of voordoet. Tegenwoordig zijn dit soort video's overal te vinden, denk bijvoorbeeld aan een make-up tutorial op YouTube of een korte video van Khan Academy over wiskundige principes.

Eén vraag houdt ons bezig: maakt het uit *wie* de uitleg geeft in zo'n instructievideo? Daar kan jij bij helpen!

Hoe kun je deelnemen?
Deelname kan online en werkt het beste via Google Chrome.
Via onderstaande link kan je via mobiel, tablet of computer deelnemen aan dit onderzoek:
https://tinyurl.com/lerenvanvideovoorbeelden

De instructies volgen vanzelf. Deelname aan dit onderzoek is geheel vrijwillig en duurt ongeveer 50 minuten.

Wat kun je ervoor terug verwachten?
Als je besluit deel te nemen, zijn wij hier natuurlijk heel dankbaar voor. Daarnaast wordt er onder de deelnemers een bol.com bon verloot ter waarde van 20 euro!

Wil je meer weten of heb je vragen? Mail dan naar j.vanoijen@students.uu.nl of n.pater@students.uu.nl

Appendix D: Informed consent (age < 16)

This active consent form is displayed in an online survey tool named Qualtrics (www.qualtrics.com)

Onderzoek naar leren van videovoorbeelden

Beste ouder(s)/verzorger(s),

Tegenwoordig kijkt iedereen wel eens filmpjes waarin iemand iets uitlegt en voordoet, bijvoorbeeld een filmpje op YouTube waarin iemand laat zien hoe je een lekke fietsband moet plakken of eentje waarin wordt uitgelegd hoe je een moeilijke wiskunde som kan oplossen. Het is maar goed ook dat we dit soort 'videovoorbeelden' bekijken, want het blijkt een hele effectieve manier te zijn om iets te leren. Maar ondanks dat dit soort videovoorbeelden heel veel bekeken worden en worden gebruikt in onderwijs, is er weinig bekend over 'wat werkt' in deze video's. Voor onze master in onderwijswetenschappen aan de Universiteit Utrecht doen wij, Jonneke van Oijen en Noortje Pater, hier daarom onderzoek naar.

1. Wat houdt het onderzoek in?
Het onderzoek bestaat uit één online sessie van ongeveer 50 minuten.
Uw kind zal gedurende deze sessie een vragenlijst invullen en hij/zij zal een videovoorbeeld bekijken en hier vragen over maken.
De stof die wordt behandeld in de videovoorbeelden maakt deel uit van het curriculum van het vak natuurkunde op HAVO/VWO niveau.

2. Privacy en vertrouwelijkheid
Alle gegevens worden vertrouwelijk behandeld en anoniem verwerkt en opgeslagen. We rapporteren alleen resultaten over groepen respondenten, niet over individuele respondenten.
Persoonsgegevens worden alleen bewaard zo lang dit nodig is voor het doel waarvoor ze verzameld zijn. Overige gegevens worden minimaal 10 jaar veilig bewaard en worden mogelijk gebruikt voor vervolgonderzoek of toekomstig onderzoek met een ander doel.

3. Deelname
Deelname aan dit onderzoek is geheel vrijwillig en de participant kan op elk moment stoppen zonder opgave van redenen en zonder consequenties. Mocht u vragen hebben dan kunt u contact opnemen via de mail: n.pater@students.uu.nl of j.vanoijen@students.uu.nl.

4. Toestemming
Voor deelname van uw kind hebben wij uw toestemming nodig.
Ik heb de informatie gelezen en:

- ○ Mijn kind doet wel mee aan het onderzoek
- ○ Mijn kind doet niet mee aan het onderzoek

Appendix E: Informed consent (age > 16)

This active consent form is displayed in an online survey tool named Qualtrics (www.qualtrics.com)

In deze les ga je aan de hand van videovoorbeelden leren hoe je fouten in elektrische schakelingen kunt vinden en oplossen.
Voor dit onderzoek is het belangrijk om te kijken hoe goed jij leert aan de hand van deze videovoorbeelden. In deze les gaan wij dus kijken hoe goed je het doet.

Deze les is onderdeel van een wetenschappelijk onderzoek. Voor jouw deelname en het gebruik van de onderzoeksgegevens is jouw toestemming nodig. Hieronder staat een toestemmingsverklaring. Lees deze verklaring door en teken vervolgens deze verklaring onderaan de pagina. Zodra deze verklaring is ingevuld mag je op de knop **'ga verder'** drukken om te beginnen.

Toestemmingsverklaring deelname onderzoek

Ik heb uitleg over het onderzoek gekregen en ik heb mijn vragen over het onderzoek kunnen stellen en ben tevreden over de antwoorden.
Ik heb goed kunnen nadenken over deelname aan het onderzoek. Ik weet dat meedoen helemaal vrijwillig is. Ik weet dat ik op elk moment kan beslissen te stoppen met het onderzoek. Daarvoor hoef ik geen reden te geven.

Ik vind het goed dat de gegevens voor wetenschappelijk onderzoek worden gebruikt. Ik weet dat alles minimaal 10 jaar veilig wordt bewaard.

Ik heb deze informatie gelezen en:

- ○ Ik doe wel mee aan het onderzoek
- ○ Ik doe niet mee aan het onderzoek

Appendix F: Intrinsic Motivation Inventory (IMI)

Vernieuwenderwijs

Intrinsieke Motivatie Inventaris (IMI)

Hieronder volgen 22 stellingen over een activiteit. Je docent zal uitleggen over welke activiteit of opdracht de vragen gaan. Omcirkel per stelling in hoeverre je jezelf in die stelling herkent. Er is geen goed of fout antwoord.

Naam: ________________________ **Leeftijd:** ___________ **Jongen** ☐ **Meisje** ☐

Klas: ____________ **Docent:** _______________________

	Pas helemaal niet bij mij						**Past helemaal bij mij**
1. Ik vond deze activiteit erg leuk om te doen.	1	2	3	4	5	6	7
2. Ik voelde mij niet nerveus terwijl ik bezig was met de activiteit.	1	2	3	4	5	6	7
3. Ik had voor mijn gevoel de keuze om de activiteit wel of niet te doen.	1	2	3	4	5	6	7
4. Ik denk dat ik best wel goed ben in deze activiteit.	1	2	3	4	5	6	7
5. Ik vond deze activiteit erg interessant.	1	2	3	4	5	6	7
6. Ik voelde mij gespannen tijdens de activiteit.	1	2	3	4	5	6	7
7. Ik denk dat ik best wel goed ben in deze activiteit, vergeleken met andere leerlingen.	1	2	3	4	5	6	7
8. Het was een leuke activiteit om te doen	1	2	3	4	5	6	7
9. Ik was ontspannen tijdens de activiteit.	1	2	3	4	5	6	7
10. Ik vond het erg leuk om deze acititeit te doen.	1	2	3	4	5	6	7
11. Ik had niet echt een keuze om de activiteit wel of niet te doen.	1	2	3	4	5	6	7
12. Ik ben tevreden met hoe ik het heb gedaan bij deze activiteit.	1	2	3	4	5	6	7
13. Ik was nerveus tijdens de activiteit.	1	2	3	4	5	6	7
14. Ik vond de activiteit erg saai.	1	2	3	4	5	6	7

Intrinsic Motivation Inventory (IMI) door Ryan & Desi (2000).
Vertaling door Wessel Peeters (Vernieuwenderwijs).

15. Ik heb het gevoel dat ik deed wat ik wilde doen terwijl bezig was met de activiteit.	1	2	3	4	5	6	7
16. Ik voelde mij competent* bij deze activiteit.	1	2	3	4	5	6	7
17. Ik vond de activiteit erg interessant.	1	2	3	4	5	6	7
18. Ik ervaarde druk tijdens de activiteit.	1	2	3	4	5	6	7
19. Ik heb het gevoel dat ik de activiteit moest doen.	1	2	3	4	5	6	7
20. Ik zou de activiteit als 'erg leuk' omschrijven.	1	2	3	4	5	6	7
21. Ik deed de activiteit omdat ik geen keuze had.	1	2	3	4	5	6	7
22. Nadat ik enige tijd bezig was met deze activiteit, voelde ik mij best competent.*	1	2	3	4	5	6	7

** Competent: in hoeverre je goed bent in iets.*