Learning and Instruction 44 (2016) 22–30

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Learning from video modeling examples: Content kept equal, adults are more effective models than peers

Vincent Hoogerheide [a, b, *], Margot van Wermeskerken [b], Sofie M.M. Loyens [a, c], Tamara van Gog [a, b]

[a] Institute of Psychology, Erasmus University Rotterdam, The Netherlands
[b] Department of Education, Utrecht University, The Netherlands
[c] Roosevelt Center for Excellence in Education, University College Roosevelt, The Netherlands

ARTICLE INFO

Article history:
Received 15 August 2015
Received in revised form
26 January 2016
Accepted 9 February 2016
Available online 4 March 2016

Keywords:
Example-based learning
Observational learning
Instructional video
Model-observer similarity
Expertise

ABSTRACT

Learning from (video) modeling examples in which a model demonstrates how to perform a task is an effective instructional strategy. The model-observer similarity (MOS) hypothesis postulates that (perceived) similarity between learners and the model in terms of age or expertise moderates the effectiveness of modeling examples. Findings have been mixed, however, possibly because manipulations of MOS were often associated with differences in example content and manipulations of (perceived) expertise confounded with age. Therefore, we investigated whether similarity with the model in terms of age and putative expertise would affect cognitive and motivational aspects of learning when the example content is kept equal across conditions. Adolescents ($N = 157$) watched a short video in which a peer or adult model was introduced as having low or high expertise, followed by two video modeling examples in which the model demonstrated how to troubleshoot electrical circuit problems. Results showed no effects of putative expertise. In contrast to the MOS hypothesis, adult models were more effective and efficient to learn from than peer models.

1. Introduction

Instructional videos are rapidly gaining popularity in education. They form the backbone of massive open online courses (MOOCs) and blended courses, and support students during self-study at home or at school. Next to web lectures (e.g., Chen & Wu, 2015; Korving, Hernández, & De Groot, 2016; Traphagan, Kucsera, & Kishi, 2010) and short knowledge clips (e.g., Day, 2008), demonstration (i.e., "how-to") videos (e.g., Ayres, Marcus, Chan, & Qian, 2009; Van der Meij & Van der Meij, 2013) make up an important part of the instructional videos on offer. Such demonstration videos are also known as *video modeling examples*. Research inspired by Bandura's (1977, 1986) social learning theory has shown the effectiveness of observational learning from human models, and this dovetails nicely with findings from cognitive psychology and instructional design research (e.g., Anderson, 1993; Sweller, Ayres, & Kalyuga, 2011) that has shown the effectiveness of example-based learning (for reviews: Renkl, 2014; Sweller et al., 2011; Van Gog & Rummel, 2010).[1]

Video modeling examples in which a model demonstrates and explains how to solve a problem are effective for acquiring new skills (e.g., Braaksma, Rijlaarsdam, & Van den Bergh, 2002; Schunk, Hanson, & Cox, 1987; Schwan & Riempp, 2004; Van Gog, Verveer, & Verveer, 2014) and may enhance the confidence learners have in their own capabilities to perform the modeled task (i.e., self-efficacy and perceived competence; Bandura, 1997; Hoogerheide, Loyens, & Van Gog, 2014, Hoogerheide, Loyens, & Van Gog, 2016; Schunk & Hanson, 1985). Yet, when developing video modeling examples, several design choices have to be made that may influence their effectiveness, the most salient of which is the choice of model. The present study investigates whether similarity between the learner and the model in terms of age and (putative) expertise would affect self-efficacy and learning outcomes, as predicted by the model-observer similarity hypothesis.

* Corresponding author. Department of Education, Utrecht University, P.O. Box 80140, 3508 TC, The Netherlands.
E-mail address: v.hoogerheide@uu.nl (V. Hoogerheide).

[1] Note that examples can lose their effectiveness or may even hamper learning when students have some prior knowledge of the problem (Kalyuga, Chandler, Tuovinen, & Sweller, 2001; Kalyuga & Renkl, 2010).

http://dx.doi.org/10.1016/j.learninstruc.2016.02.004

1.1. The model-observer similarity hypothesis

The model-observer similarity (MOS) hypothesis (Bandura, 1994; Schunk, 1987; see also the similarity-attraction hypothesis, Montoya & Horton, 2013; Moreno & Flowerday, 2006; Reeves & Nass, 1996) postulates that, because modeling enables social comparison (Berger, 1977; Johnson & Lammers, 2012), the effectiveness of observational learning from (video) modeling examples depends in part on how similar to the model learners perceive themselves to be. Or in Bandura's (1994) words:

> The impact of modeling on perceived self-efficacy is strongly influenced by perceived similarity to the models. The greater the assumed similarity, the more persuasive are the models' successes and failures. If people see the models as very different from themselves their perceived self-efficacy is not much influenced by the models' behavior and the results it produces. (p.72)

Self-efficacy and the closely related construct of perceived competence are important, as they have been linked to factors such as academic motivation (Self-efficacy: Bandura, 1994; Schunk, 1991, 2001; Schwarzer, 1992; Perceived competence: Bong & Skaalvik, 2003; Harter, 1990) and learning outcomes (Self-efficacy: Bandura, 1994; Schwarzer, 1992; Perceived competence: Bong & Skaalvik, 2003; Harter, 1990; Ma & Kishor, 1997). Learners who perceive themselves as more similar to the model may also feel more attracted to the model and pay more attention to the model (Berscheid & Walster, 1969), and a high degree of similarity can help them form outcome expectations (Schunk, 1987). Similarity factors may be particularly important for novice learners whose self-efficacy and prior knowledge are still low, as they are especially prone to engaging in social comparison (Buunk, Zurriaga, Gonzalez-Roma, & Subirats, 2003). The present study focuses on MOS in terms of age and putative expertise.

1.2. Model-observer similarity in age and expertise

With regard to the age of a model, the MOS-hypothesis predicts that primary or secondary education students would benefit more from a model that is perceived as similar in age, such as a peer model, than dissimilar in age, such as an adult model. Findings have been mixed however, with some studies showing stronger effects of observing a peer model compared to an adult model (e.g., Davidson & Smith, 1982; Rodriguez Buritica, Eppinger, Schuck, Heekeren, & Shu-Chen & Wu, 2015; Schunk & Hanson, 1985; Zmyj, Aschersleben, Prinz, & Daum, 2012), some showing no differences (Robert, 1983; Strauss, 1978), and others showing stronger effects of an adult model (e.g., Hicks, 1965; Jakubczak & Walters, 1959). A possible explanation for these mixed findings may be that peer models are especially beneficial for learners who have encountered difficulties in learning or for learners of low ability (Schunk, 1987). Schunk and Hanson (1985), for instance, examined whether children who previously showed difficulties learning fractions benefited more from a peer model, a teacher model, or no model, and found that peer modeling was more conducive to both self-efficacy and learning than teacher modeling, while both models were more effective than no modeling. Another possible explanation is that age only becomes a salient cue when coupled with (perceived) expertise. That is, students may particularly imitate peer models when they believe them to be high in expertise, and age may become an informative cue especially for tasks in which peers are generally (perceived as) less of an expert than adults (Bandura, 1986; Schunk, 1987).

Research on the MOS-hypothesis in terms of expertise has used different approaches. One line of research contrasted learning from a mastery model (i.e., a model who displays faultless performance from the start) to learning from a coping model (i.e., a model who shows performance errors that he or she corrects later on), and this has led to mixed results. For instance, in math, no differences in the effectiveness of both model types were found for low ability students who had had prior successful experiences with the task (e.g., Schunk & Hanson, 1985) or for average ability students (Schunk & Hanson, 1989). However, for low ability students without prior success with the task, coping models were more effective for learning (Schunk et al., 1987).

Another line of research has compared the effects of learning from a high expertise (e.g., expert) model to a lower expertise (e.g., advanced student) model, the latter being closer in knowledge and skill to novice learners. Contrary to the model-observer similarity hypothesis, older findings indicate that for primary school children, a more expert model was more beneficial for a wide range of measures such as learning communication skills or paired-associates relative to a low expertise model (e.g., Simon, Ditrichs, & Speckhart, 1975; Sonnenschein & Whitehurst, 1980). In line with the MOS-hypothesis, however, Braaksma et al. (2002) showed more recently that secondary education students who had weak writing skills benefitted more from being instructed to focus on weak models who explained and demonstrated how to write an argumentative text (on video) than from focusing on strong models, whereas the reversed effect was found for more competent students. Studies in higher professional education, however, showed no benefit of (advanced) peer models: written examples created by experts fostered transfer (i.e., applying the acquired knowledge to novel tasks) more than examples created by advanced peer students, possibly because experts' explanations contain a higher degree of abstraction (Boekhout, Van Gog, Van de Wiel, Gerards-Last, & Geraets, 2010; Lachner & Nückles, 2015).

Clearly, findings regarding both age and expertise have been mixed. There are two important things to note, however. First, in many of those studies, there were actual differences in how the models behaved across conditions or in other words, in the content of the examples. This applies, for instance, to studies that contrasted learning from coping models and mastery models because only coping models' behaviour contains expressions of uncertainty and/or errors (e.g., Kitsantas, Zimmerman, & Cleary, 2000; Schunk & Hanson, 1985; Zimmerman & Kitsantas, 2002), and to studies that compared high and lower expertise models because their explanations differ in quality (e.g., Lachner & Nückles, 2015; Simon et al., 1975; Sonnenschein & Whitehurst, 1980). This makes it hard to evaluate whether any differences in motivational or learning outcomes were due to (perceived) similarity or to differences in content. Some evidence indicating that perceived similarity may still influence cognitive, affective, or motivational aspects of learning when all else is equal, comes from studies with *animated* models (i.e., animated pedagogical agents) in which the content was kept equal. For instance, Rosenberg-Kima, Baylor, Plant, and Doerr (2008) found that self-efficacy was enhanced more for students who learned about engineering from a 'young and cool' agent than a 'young and uncool' and an 'older and (un) cool' agent. Liew, Tan, and Jayothisa (2013) found that for female university students, a peer-like agent was more enjoyable to learn programming skills from than an expert-like agent, although the expert-like agents were more credible and led to less anxiety during learning, possibly because people are more easily persuaded by those whom they perceive as experts (Chaiken & Maheswaran, 1994; Debono & Harnish, 1988). Lastly, Kim, Baylor, and Reed (2003) found that a mentor-like agent was as beneficial for learning compared to an expert-like agent, but was considered more motivating to interact with and learn from.

Secondly, age and expertise manipulations were often confounded. For example, Davidson and Smith (1982) investigated the relationship between model expertise and children's self-evaluation skills, and, instead of keeping model age constant across conditions, children observed a peer of equal skill, an adult of superior skill, or a child of inferior skill. Animated agent studies have also confounded age and expertise manipulations. In the studies of Kim et al. (2003) and Baylor and Kim (2004), for instance, the expert-like agent looked much older than the mentor-like agent. Some early video modeling example studies have tried to disentangle the effects of age and expertise. For instance, Sonnenschein and Whitehurst (1980) showed that observing high ('informative') and low expertise ('uninformative') peer models and high expertise adult models enhanced children's communicative skills more than watching low expertise adult models. Children did, however, evaluate the low expertise adult as more knowledgeable than the low expertise peer. Becker and Glidden (1979) found that expertise and age interacted for low ability children. Children observed a low or high expertise peer or adult model performing a motor task while displaying certain social behaviors. The behavior of high expertise models and peer models was imitated more than that of low expertise models and adult models, presumably because the social behavior of the peers was evaluated as more appropriate. Note though, that in these studies, the example content again varied across conditions, and it is therefore uncertain whether these effects were caused by differences in perceived similarity to the model.

In sum, it is unclear whether similarity to a model in terms of age and (perceived) expertise would play a role in learning when the content of the examples would be kept equal and whether age and perceived expertise of the model contribute independently to effects on motivation and learning outcomes or only in interaction. Therefore, the present study examined whether, when the content of the example is controlled for, the effectiveness of studying video modeling examples for novice students' perceptions of their own capabilities to perform the modelled task (i.e., perceived competence and self-efficacy) and learning outcomes (i.e., posttest performance) depends on whether the model is of similar or dissimilar age and whether the model is introduced as having low or high expertise.

1.3. *The present study*

We addressed the question of whether model-observer similarity in age, putative expertise, or both would affect novice secondary education students' learning (i.e., adolescents of about 15 years of age who did not have prior knowledge of the task). They studied two video modeling examples on how to solve a science problem (troubleshooting electrical circuits). The models were either peers (17 years old) or adults (42 years old) who were introduced prior to example study as being enrolled in a tutor-training (peers) or teacher-training (adults) program and as having low expertise or high expertise in science. We kept all else equal, both with respect to model characteristics (i.e., all models were Caucasian females from the same region of the country, wearing a black t-shirt and blue jeans) and the content of the videos (i.e., all models narrated the exact same text, spent an equal amount of time on the parts of the video and the video as a whole, and were trained to show the same movements and gestures). Moreover, to ensure that any effects of condition were not associated with one particular model, the two adult and the two peer models featured in both the high and low expertise conditions (i.e., half of the participants in the low expertise adult condition saw "adult model 1" the other half "adult model 2").

The primary research question was whether students would perform better on the posttest and show greater self-efficacy and perceived competence when they were more similar to the model in age, expertise, or both. Given that students are novices with regard to the modeled task, the MOS-hypothesis (Schunk, 1987) would predict that students' self-efficacy, perceived competence, and learning outcomes would benefit most from studying a peer model with low putative expertise. Because of the fact that prior research has produced mixed findings, often confounding age and expertise or expertise and example content, however, we are hesitant to adopt the MOS-hypothesis for the present study and rather approach this as an open question. We also measured mental effort invested during example study and the posttest to obtain more information on the cognitive efficiency of the instructional conditions (Van Gog & Paas, 2008). Effects on learning enjoyment were also explored because previous studies have shown influences on affect (e.g., Kim et al., 2003; Liew et al., 2013) and enjoyment may be an important cue for whether students would use examples during self-study (Yi & Hwang, 2003). Lastly, students evaluated the quality of the model's explanation.

2. Method

2.1. *Participants and design*

Participants were 157 Dutch secondary education students (82 male; $M^{age} = 14.99$ years, $SD = 0.64$) in their third or fourth year of pre-university education. The experiment used a 2×2 design, with Model Age (Peer vs. Adult) and Model Expertise (Low vs. High) as between-subject factors. Students were quasi-randomly (i.e., matched for gender) allocated to the Low Expertise Peer ($n = 39$, 21 males), High Expertise Peer ($n = 40$, 21 males), Low Expertise Adult ($n = 38$, 21 males), or High Expertise Adult ($n = 40$, 21 males) Model conditions. There were two adult and two peer models, featuring in both the High and Low Expertise conditions. Within each condition, half of the students received one model, the other half the other model (e.g., half of the participants in the Low Expertise Adult Model condition saw "adult model 1" the other half "adult model 2"). At the time of the experiment, students had taken basic science classes but were novices with regard to the modelled task (troubleshooting electrical circuits) as this had not yet been covered in their curriculum according to the teachers.

2.2. *Materials*

The materials for this study were based on the pen-and-paper materials on troubleshooting parallel electrical circuits from prior studies on example-based learning (e.g., Hoogerheide, Loyens, Jadi, Vrins, & Van Gog, 2015; Van Gog & Kester, 2012; Van Gog, Kester, Dirkx, et al., 2015; Van Gog, Kester, & Paas, 2011), but were presented online in the web-based Qualtrics platform (http://www.qualtrics.com).

2.2.1. *Conceptual prior knowledge test*

The prior knowledge test consisted of seven conceptual open-ended questions on troubleshooting and parallel circuits principles. This test was used as a check that students indeed had little if any prior knowledge of the principles required for troubleshooting parallel electrical circuits (e.g., relations between voltage, current, and resistance in parallel circuits) and to rule out differences among conditions in prior knowledge.

2.2.2. *Introductory text*

A short introductory text explained what the abbreviations and components in a circuit drawing stand for and described Ohm's law and the three different forms of the formula (i.e., $R = U/I$; $I = U/R$; $U = I^*R$).

2.2.3. Video modeling examples

Two video modeling examples were created by all four models (i.e., the two Peer and the two Adult Models) under both High and Low Expertise conditions. In the first example (240 s), the fault was that the measured current in one of the parallel branches was higher than one would expect, meaning that the resistance in that branch was lower; in the second example (244 s), the current was lower, meaning that resistance was higher. Each example showed the model standing to the right of a large screen displaying PowerPoint slides (see Fig. 1). Each example began with a circuit drawing containing three parallel branches that was presented on the screen; the circuit indicated how much resistance each resistor provided as well as how much voltage the power source delivered. The model explained based on this circuit drawing that the information on voltage and resistance presented in the drawing can be used to calculate what current should be measured in all three parallel branches and overall if the circuit were functioning correctly. The model then provided a step-by-step demonstration of how to calculate the current in each branch as well as the total current (sum of the currents in the branches) using Ohm's law; this explanation was supported by a slide that showed the same circuit drawing (only smaller), Ohm's law, and the worked-out problem-solving steps. The next slide presented measured current at each ammeter below the currents that should be measured if the circuit were functioning correctly and the model pointed out the discrepancy in one of the branches (i.e., either higher in example 1 or lower in example 2) and explained that this meant the resistance was lower (example 1) or higher (example 2) than indicated in the drawing and demonstrated how to calculate the actual resistance, supported by a slide displaying the measured currents, Ohm's law, and the calculation.

An autocue was placed next to the camera to ensure that all models provided the same explanation and spent the same amount of time on all parts of both videos and therefore on the videos as a whole. The models were instructed to gesture to elements in the PowerPoint slides to support their verbal explanation and had the opportunity to practice the entire process several times before the definitive recordings were created. All four models were Caucasian females and wore a black t-shirt and blue jeans.

2.2.4. Age manipulation

Two adolescents (real ages 16 and 17) served as a peer model (introduced as being 17 years old) in both Peer Model conditions, and two adults (real ages 42 and 43) as an adult model (introduced as being 42 years old) in both Adult Model conditions.

2.2.5. Expertise manipulation

A short video was created in which the models introduced themselves, to be presented prior to the examples. Each model created an introduction for both the Low and High Expertise condition (both 40 s). This video showed the model standing next to the screen (i.e., the setup of the video modeling examples), but the screen was empty. The introduction of the peer models started as follows: ‘My name is Natasja/Denise. I am 17 years old and I am enrolled in a homework tutor-training program. For a course within that program, I was instructed to create instructional videos that can be used for homework purposes.’ The introduction of the adult models started with: ‘My name is Natasja/Denise. I am 42 years old and I am enrolled in a lateral-entry teacher-training program. For a course within that program, I was instructed to create instructional videos that can be used for homework purposes.’ All introduction videos then continued: ‘Therefore, you will be shown two examples that demonstrate how to detect and solve a problem in an electrical circuit. Afterwards, you will be asked to solve similar problems yourself to see how much you have learned from my explanation.’ Next, the low vs. high expertise manipulation followed, with the Low Expertise introduction stating ‘I hope that I can explain this clearly, as I am not so proficient in physics and I am not taking

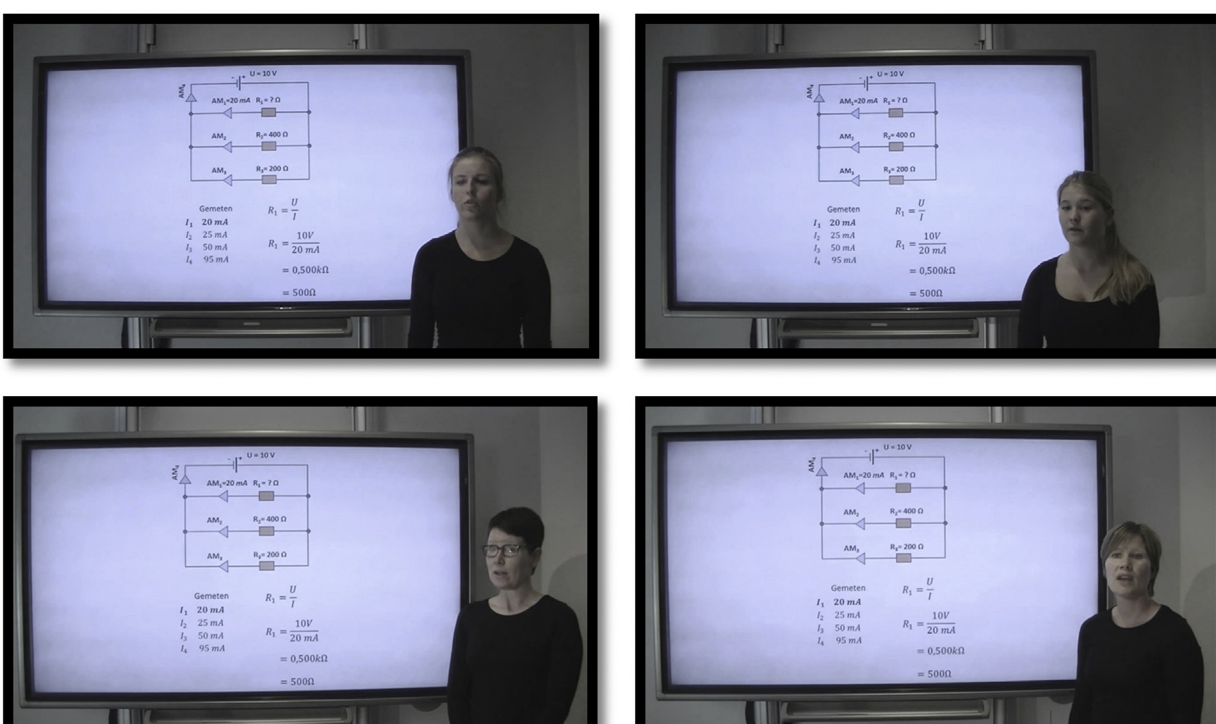

Fig. 1. Peer models (top row) and adult models (bottom row).

[peer]/did not take [adult] physics as a final examination subject in secondary education.[2] But I will do my best' while the High Expertise introduction was 'I expect that I can explain this clearly, as I am very proficient in physics and I am taking [peer]/did take [adult] physics as a final examination subject in secondary education. So I will do my best'.

2.2.6. Posttest

The posttest presented two troubleshooting problems. The first one was isomorphic to the problems used in the video modeling examples (i.e., one fault), but the second problem was slightly different in the sense that it contained both faults that had been encountered in the training. Both problems reminded students that: 'The current (U) is expressed in volt (V), resistance (R) is expressed in Ohm (Ω), and power (I) is expressed in amperes (A).'

2.2.7. Mental effort

After each video modeling example and each posttest task, invested mental effort was measured using the rating scale developed by Paas (1992), which ranges from (1) very, very low effort to (9) very, very high effort.

2.2.8. Self-efficacy and perceived competence

Self-efficacy was measured by asking students how confident they were that they had mastered the skill of detecting and solving electrical circuit problems (cf. Bandura, 2006), on a scale of 1 (not at all confident) to 9 (very, very confident). Perceived competence was measured using an adapted version of the Perceived Competence Scale for Learning of Williams and Deci (1996). Participants were asked to rate on a scale of (1) not at all true to (7) very true to what degree the following items apply to them: 'I feel able to meet the challenge of performing well in detecting and solving electrical circuit problems', 'I feel confident in my ability to detect and solve electrical circuit problems', and 'I am capable of detecting and solving electrical circuit problems'. That is, the adaptation consisted of rephrasing the questions to focus on detecting and solving electrical circuit problems and of removing the item 'I am able to achieve my goals in this course' because it did not apply to our study context.

2.2.9. Learning enjoyment

Participants were asked to give a 'school-grade' on a scale of 0 (lowest) to 10 (highest) for how enjoyable studying the video modeling examples was for them (cf. Hoogerheide et al., 2014).

2.2.10. Explanation quality

Participants were asked to rate the quality of explanations provided in the video modeling examples on a scale of (1) very, very bad quality to (9) very, very good quality.

2.3. Procedure

Prior to the experiment, participants were quasi-randomly (i.e., matched for gender) allocated to one of four conditions and, within each condition, to one of the two peer or adult models, based on a name list. This was done to ensure that all conditions and models contained an approximately equal number of students and ratio of male to female students. The experiment was run in 8 sessions of ca. 50 min. duration in a computer lab at participants' schools. Participants were told to sit at the computer that was marked with their name on a sheet of A4 paper; that sheet also contained the link to the Qualtrics questionnaire of their assigned condition. The experimenter first gave a brief plenary general introduction and instructed students how to access the Qualtrics questionnaire. This questionnaire presented 4 'blocks' of questions. Block 1 contained demographic questions and the conceptual prior knowledge test followed by self-efficacy and perceived competence ratings. Students were given 6 min to complete this block. Next, students were instructed to study the introductory text for 2 min in block 2, and the experimenter emphasized that students needed to study this information carefully to be able to comprehend the demonstration videos later on. Block 3 first presented the expertise manipulation video in which the model introduced herself, followed by the two video modeling examples (the videos were embedded in Qualtrics via YouTube). After each example, participants were asked to rate how much effort they invested in studying the example. At the end of block 3, participants rated their learning enjoyment, the perceived quality of the explanations provided in the examples, and their self-efficacy and perceived competence. Block 4 presented the posttest for which participants received 12 min; after each test problem participants were asked to rate how much effort they invested in solving it.

2.4. Data analysis

Averages were computed for effort invested during example study, effort invested in the posttest, perceived competence after the prior knowledge test, and perceived competence prior to the posttest. Test performance was scored based on straightforward coding schemes that had been developed and used in prior research (e.g., Hoogerheide et al., 2015; Van Gog et al., 2011; Van Gog et al., 2015; Van Gog & Kester, 2012). Ten points could be earned for the prior knowledge test (and partial credit was given for partially correct answers). The maximum score to be earned for the posttest was eight points. For the first task that contained only one fault, one point could be earned for calculating the correct value of all ammeters, one for indicating which resistor was faulty, and one for indicating what the faulty resistor's actual value was. For the second task that contained two faults, an extra point was granted for correctly indicating the second faulty resistor and for correctly calculating its resistance. Incomplete or partially correct answers were given half a point. Two raters scored 10% of the pretests and posttests, and Cohen's κ was run to determine if there was agreement between them (on item level). There were high levels of agreement between the two raters on the pretest scores, $\kappa = .909$, $p < .001$, and on the posttest scores, $\kappa = .922$, $p < .001$. Therefore, the remainder of the tests was scored by a single rater.

One student in the High Expertise Adult Model condition was removed from all the analyses due to non-compliance with the instructions, leaving 156 participants. One participant from the Low Expertise Adult Model condition had to leave early and was therefore excluded from the posttest and invested mental effort in the posttest analyses. Three participants had one missing value on invested mental effort in the posttest, which were replaced by the series mean.

3. Results

Two types of preliminary analyses were conducted. First, it was investigated whether there was a difference between the two peer models and the two adult models by means of independent samples t-tests on all outcome variables. Because there were no significant differences, we proceeded analysing the data at condition level. Second, we checked whether participants' prior knowledge was indeed low (which it was, as can be seen in Table 1) and did not

[2] In the Netherlands, after some general years with a common curriculum, secondary education students can chose 'profiles' of subjects in which they will take the final examination, and not all of those include physics.

Table 1
Mean (SD) of test performance, invested mental effort, self-efficacy, perceived competence, explanation quality, and learning enjoyment per condition.

	Peer Model		Adult Model	
	Low putative expertise	High putative expertise	Low putative expertise	High putative expertise
Performance Pretest *(range 0–10)*	1.59 (1.16)	1.61 (1.38)	1.79 (1.11)	1.49 (0.98)
Performance Posttest *(range 0-8)*	2.90 (2.17)	3.50 (2.49)	3.97 (2.47)	4.29 (2.43)
Mental Effort Study Phase *(range: 1–9)*	3.96 (2.09)	4.10 (1.53)	3.55 (1.72)	3.31 (1.40)
Mental Effort Posttest *(range 1–9)*	4.81 (2.01)	4.24 (1.72)	4.41 (1.97)	4.32 (1.51)
Self-efficacy Pretest *(range 1–9)*	3.18 (1.73)	3.58 (1.68)	3.68 (1.65)	3.85 (2.05)
Self-efficacy Posttest *(range 1–9)*	5.10 (1.93)	5.63 (1.17)	5.55 (1.39)	5.92 (1.36)
Perceived Competence Pretest *(range 1–7)*	2.56 (1.30)	3.03 (1.08)	3.18 (1.16)	2.97 (1.39)
Perceived Competence Posttest *(range 1–7)*	4.08 (1.47)	4.48 (0.99)	4.24 (1.34)	4.51 (1.32)
Explanation Quality *(range 1–9)*	5.62 (1.60)	6.20 (1.29)	6.34 (1.17)	6.59 (1.14)
Learning Enjoyment *(range 0–10)*	4.00 (2.24)	4.08 (2.31)	4.45 (2.33)	4.56 (2.28)

differ among conditions. Indeed, a 2 × 2 ANOVA on the prior knowledge test scores with Model Age (Peer, Adult) and Model Expertise (Low, High) as between-subject factors revealed no main or interaction effects (all Fs < 1).

Test performance, invested mental effort, learning enjoyment, and explanation quality results were analysed using 2 × 2 ANOVAs, with Model Age (Peer, Adult) and Model Expertise (Low, High) as between-subject factors. The self-efficacy and perceived competence results were analysed using repeated measures ANOVAs with Test Moment (Before and After Example Study) as within-subjects factor and Model Age (Peer, Adult) and Model Expertise (Low, High) as between-subjects factors. Mean (and SD) scores on all variables are shown in Table 1.

3.1. Posttest performance

There was a main effect of Model Age on posttest performance, $F(1,151) = 5.92$, $p = .016$, $\eta_p^2 = .038$, indicating that participants who had observed an adult model ($M = 4.14$, $SD = 2.44$) outperformed those who had observed a peer model ($M = 3.20$, $SD = 2.34$). There was no main effect of Model Expertise, $F(1,151) = 1.45$, $p = .231$, $\eta_p^2 = .009$, nor an interaction effect, $F < 1$.[3]

3.2. Mental effort

We found a main effect of Model Age on mental effort invested during example study, $F(1,152) = 4.84$, $p = .029$, $\eta_p^2 = .031$, indicating that participants who observed an adult model ($M = 3.43$, $SD = 1.56$) invested less effort than those who observed a peer model ($M = 4.03$, $SD = 1.82$). There was no main effect of Model Expertise, nor an interaction effect, both Fs < 1. With regard to mental effort invested in completing the posttest tasks, there were no significant main or interaction effects (Model Age: $F < 1$; Model Expertise: $F(1,151) = 1.27$, $p = .263$, $\eta_p^2 = .008$; interaction: $F < 1$).

3.3. Self-efficacy and perceived competence

The analysis of self-efficacy showed a main effect of Test Moment, $F(1,152) = 233.53$, $p < .001$, $\eta_p^2 = .606$, indicating that self-efficacy improved from before ($M = 3.57$, $SD = 1.78$) to after example study ($M = 5.55$, $SD = 1.50$). Other than that, there were no significant main effects (Model Age: $F(1,152) = 2.77$, $p = .098$, $\eta_p^2 = .018$; Model Expertise: $F(1,152) = 2.51$, $p = .115$, $\eta_p^2 = .016$) or interaction effects (all Fs < 1).

A similar pattern was found for students' perceptions of their own competence. There was a main effect of Test Moment, $F(1,152) = 244.36$, $p < .001$, $\eta_p^2 = .617$, indicating that perceived competence improved from before ($M = 2.94$, $SD = 1.25$) to after example study ($M = 4.33$, $SD = 1.29$). Other than that, there were neither main effects (Model Age: $F(1,152) = 1.06$, $p = .304$, $\eta_p^2 = .007$; Model Expertise: $F(1,152) = 1.74$, $p = .189$, $\eta_p^2 = .011$), nor interaction effects (Test Moment * Model Expertise, $F(1,152) = 1.30$, $p = .257$, $\eta_p^2 = .008$; Test Moment * Model Age, $F(1,152) = 1.09$, $p = .299$, $\eta_p^2 = .007$; Model Expertise * Model Age, $F(1,152) = 1.24$, $p = .267$, $\eta_p^2 = .008$; Test Moment * Model Expertise * Model Age, $F(1,152) = 2.37$, $p = .126$, $\eta_p^2 = .015$).

3.4. Learning enjoyment

There were no significant effects on learning enjoyment (Model Age: $F(1,152) = 1.63$, $p = .204$, $\eta_p^2 = .011$; Model Expertise: $F < 1$; interaction: $F < 1$).

3.5. Explanation quality

Despite the fact that the models provided the exact same explanations, there was a main effect of Model Age on students' ratings of the quality of explanations provided in the examples, $F(1,152) = 7.06$, $p = .009$, $\eta_p^2 = .044$, indicating that students who had observed an adult model rated the explanations as being of higher quality ($M = 6.47$, $SD = 1.15$) than students who had observed a peer model ($M = 5.91$, $SD = 1.47$). Moreover, there was also a main effect of Model Expertise, $F(1,152) = 3.92$, $p = .049$, $\eta_p^2 = .025$, showing that students who observed low expertise models rated the explanations as being of lower quality ($M = 5.97$, $SD = 1.44$) than the explanations of high expertise models ($M = 6.39$, $SD = 1.22$). There was no interaction effect, $F < 1$.[4]

4. Discussion

This experiment examined whether similarity to a model in terms of age and (putative) expertise would affect secondary education students' learning from video modeling examples. Because prior research on the model-observer similarity (MOS) hypothesis

[3] Upon a reviewer's request we explored whether there were differences among conditions in the degree to which participants correctly solved each problem as a whole. Chi-square tests showed that the number of students who managed to correctly solve the problem in its entirety (i.e., max score) did not differ among conditions (problem 1: $X^2(1, N = 156) = 0.01$, $p = .92$; problem 2: $X^2(1, N = 156) = 1.71$, $p = .19$).

[4] Upon a reviewer's request we re-ran all analyses with students' gender and pretest scores as covariates. The ANCOVA led to the same outcomes as the ANOVA with regard to all dependent variables. Pretest scores were a significant predictor of posttest scores, effort invested during example study and the posttest, and self-efficacy and perceived competence (as one might expect); gender was not a significant predictor of any of the dependent variables.

led to mixed findings, and many studies confounded expertise manipulations with example content and age, we took care to keep the example content equal across conditions and to disentangle the age and perceived expertise factors. Moreover, we used two models in each condition to rule out that effects would be caused by incidental model characteristics.

Given that students were adolescents and novices with regard to the modeled task, the MOS-hypothesis (Schunk, 1987) would predict that students' self-efficacy, perceived competence, and learning outcomes would benefit most from studying a peer model with low (putative) expertise. Our results do not support this hypothesis. On the contrary, with regard to model age, we found the opposite of what the MOS-hypothesis would predict: learners who studied *adult* models invested less effort and attained better learning outcomes than those who studied peer models. Thus, an adult model was more effective to learn from and more efficient in the sense that higher test performance was achieved with less effort investment in example study (see Van Gog & Paas, 2008, for a discussion of efficiency in terms of the relation between mental effort and performance). Students also rated the adult models' explanations as being of higher quality. Note that these findings are quite remarkable, given that the content of the examples was exactly the same. They cannot be explained through increased self-efficacy or perceived competence, however, as all students' ratings increased after example study, but did not differ among conditions.

A possible explanation for the finding that adult models were more effective might lie in perceived age-appropriateness of the modeled task. It has been proposed (Bandura, 1986; Schunk, 1987; Zmyj & Seehagen, 2013) that adult models may be more beneficial than peer models for behaviours that are viewed as more appropriate for adults and in which adults are considered to be more of an expert. The tasks demonstrated by the model were in the domain of physics, and research has shown that students typically struggle with learning physics skills and often continue to experience difficulties after extensive instruction (Duit & Von Rhöneck, 1998; Fredette & Lockhead, 1980; McDermott & Shaffer, 1992; Shipstone, 1984). As such, they might have attributed more expertise to the adult models, which would explain why students who had observed adult models found the model's explanations to be of higher quality than those who observed peer models, even though —again— the peer and adult models provided the exact same explanations. This would also explain why we did not find effects on self-efficacy; while MOS-effects predominantly occur via enhanced self-efficacy, task-appropriateness effects may not (Schunk, 1987).

But how might task-appropriateness explain better learning outcomes? Although this is a question for future research to address definitively, the answer might lie in students' attention allocation during example study. Bandura (1977, 1986) postulated that paying attention to a model is an important prerequisite for being able to emulate the modeled behavior later on, and that, in addition to MOS, model characteristics can affect how much attention is paid to a model. Peer models might lead to focusing more on task-irrelevant aspects of the video such as the model's appearance rather than aspects of the video that contribute to building a cognitive schema as a result of increased interest in and attraction to the model due to higher levels of perceived similarity (Berscheid & Walster, 1969). Moreno and Flowerday (2006) made a similar argument in the animated pedagogical agent literature to explain why a similar-ethnicity agent hampered learning; they suggested that students may have focused more on how the agent's appearance and behaviour represented them.

Next to peer models having 'negative' effects on attention, adults might have beneficial effects: students may find it easier and more natural to pay attention to adult models. Adolescents are used to learning from adults and to adults being more knowledgeable and therefore giving higher quality explanations when it comes to complex subjects such as physics, and the idea that this may enhance students' attention to what the model is saying resonates with findings from research on group interaction. It is well-established that group members are more influenced by those perceived as more knowledgeable (Bottger, 1984; Littlepage, Schmidt, Whisler, & Frost, 1995; Ridgeway, 1987), and individuals who observe group interactions on video have been shown to pay more visual attention to group members perceived as more knowledgeable (Cheng, Tracy, Kingstone, Foulsham, & Henrich, 2013). Although these findings cannot be directly translated to learning from modeling examples, as in group settings, there are always multiple individuals that an observer or group member may pay attention to, these findings do suggest that attention processes are the key to the effect of adult models in our study. The hypothesis that students attribute more expertise to adult models and therefore pay more attention to them during example study also resonates well with the communication maxim's of Grice (1975), in which it is stated that it is a social rule in conversations to pay more attention to those expected to be more knowledgeable. Higher attention levels may prevent students' minds from wandering; especially for complex tasks, learners often have difficulties building an accurate cognitive schema of the task at hand because their mind wanders easily (Smallwood, Fishman, & Schooler, 2007; Smallwood & Schooler, 2015; Szpunar, Moulton, & Schacter, 2013). Of course, these attention explanations are tentative in nature and need to be examined in future research.

The model expertise manipulation, consisting of a brief introductory video in which the model stated she is (or is not) very proficient in physics and is (not) taking/has (not) taken physics classes, seems to have been effective in the sense that students evaluated the explanations provided by the low expertise models as being of lower quality (even though they were exactly the same as in the high expertise conditions). However, this manipulation did not result in differences in self-efficacy, perceived competence, mental effort, enjoyment, or posttest performance. It is possible that this manipulation was too subtle to affect attention processes; in contrast to age-related cues that are automatically processed upon seeing another person and were continuously available during example study, cues regarding expertise were not. Another possibility is that the expertise manipulation did not result in students perceiving the low expertise models as more similar to themselves than the high expertise models, but instead regarded them as lower in competence (i.e., having no credibility at all, due to the model's uncertainty about her own ability to explain the task), which could have affected students' willingness to listen to the explanation. Although they indeed rated the quality of the explanation as lower than students in the high expertise condition, there were no significant effects of model expertise on learning outcome. As such, it seems unlikely that the explanations provided by the low expertise models were discarded by students.

Note that we took care to use two different models in each condition, to decrease the likelihood that effects of model age or putative expertise would result from specific characteristics of the particular model in a condition. Nevertheless, we cannot definitively rule out that model characteristics could have affected our results. Because there were no significant differences between both peer models and between both adult models on any of the outcome measures, however, it seems unlikely that specific model characteristics would have a strong influence on the findings. Another limitation of this study is that we only used one type of learning task, so future research should investigate whether these results hold with different types of tasks and tasks from other domains. Moreover, we only asked students to rate the quality of the model's explanations, but not the model's expertise, the perceived similarity to the model, or the appropriateness of the task for the model. Such information would have been helpful in determining whether students actually viewed the low perceived expertise models as similar, and whether the effect of model age would indeed be due to attributions of task appropriateness or expertise. An interesting avenue for future research would be to compare the effects of learning from video modeling examples with peer models to adult models for students learning a task that they view as more appropriate for their own age. If students' views of the age appropriateness of troubleshooting electrical circuit problems indeed caused adults to be more effective models than peers, then peers can be expected to be more effective models than adults in this case. Considering that views of task-appropriateness and expertise may be different for students of different ages, it would be interesting to investigate whether these findings generalize to different age groups. Moreover, it would be interesting to investigate whether these findings extend to other kinds of instructional video, such as short knowledge clips (e.g., Day, 2008) or web lectures (e.g., Korving et al., 2016 Traphagan et al., 2010).

Despite these limitations, our findings are of interest for educational practice. With video modeling examples being increasingly used in online learning environments because they have become easier to create, instructional designers and educational practitioners may want to design and use video modeling examples with an adult model rather than a peer model when the skill to be learned is viewed as more appropriate for adults because they are perceived as more of an expert.

Acknowledgements

This research was funded by Kennisnet. The authors would like to thank Vincent van Dam, Arjan Bijleveld, Rinus van der Hoek, and Peter Roseboom for facilitating this study. We also thank Natasja Hoogerheide, Denise Hoogerheide, Lisa Minderman, Celine Vermeulen, Kirsten Versijde, and Susan Ravensbergen for their help with materials and conducting the study.

References

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale: Erlbaum.
Ayres, P., Marcus, N., Chan, C., & Qian, N. (2009). Learning hand manipulative tasks: when instructional animations are superior to equivalent static representations. *Computers in Human Behavior, 25*, 348–353. http://dx.doi.org/10.1016/j.chb.2008.12.013.
Bandura, A. (1977). *Social learning theory.* Englewood Cliffs: Prentice Hall.
Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs: Prentice Hall.
Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (pp. 71–81). New York: Academic Press (Reprinted in H. Friedman [Ed.], Encyclopedia of mental health. San Diego: Academic Press, 1998).
Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.
Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares, & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Greenwich, CT: Information Age Publishing.
Baylor, A. L., & Kim, Y. (2004). Pedagogical agent design: the impact of agent realism, gender, ethnicity, and instructional role. In J. C. Lester, R. M. Vicari, & F. Paraguacu (Eds.), *Intelligent tutoring systems* (pp. 592–603). Berlin: Springer.
Becker, S., & Glidden, L. (1979). Imitation in EMR boys: model competency and age. *American Journal of Mental Deficiency, 83*, 360–366.
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.
Berscheid, E., & Walster, E. H. (1969). *Interpersonal attraction*. Reading, MA: Addison-Wesley.
Boekhout, P., Van Gog, T., Van de Wiel, M. W. J., Gerards-Last, D., & Geraets, J. (2010). Example-based learning: effects of model expertise in relation to student expertise. *British Journal of Educational Psychology, 80*, 557–566. http://dx.doi.org/10.1348/000709910X497130.
Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: how different are they really? *Educational Psychology Review, 15*, 1–40. http://dx.doi.org/10.1023/A:1021302408382.
Bottger, P. C. (1984). Expertise and air time as basis of actual and perceived influence in problem solving groups. *Journal of Applied Psychology, 69*, 214–221. http://dx.doi.org/10.1037/0021-9010.69.2.214.
Braaksma, M. A. H., Rijlaarsdam, G., & Van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology, 94*, 405–415. http://dx.doi.org/10.1037/0022-0663.94.2.405.
Buritica, J. M. R., Eppinger, B., Schuck, N. W., Heekeren, H. R., & Li, S. C. (2015). Electrophysiological correlates of observational learning in children. *Developmental Science*. Advance online publication http://dx.doi.org/10.1111/desc.12317.
Buunk, B. P., Zurriaga, R., Gonzalez-Roma, V., & Subirats, M. (2003). Engaging in upward and downward comparisons as a determinant of relative deprivation at work: a longitudinal study. *Journal of Vocational Behavior, 62*, 370–388. http://dx.doi.org/10.1016/S0001-8791(02)00015-5.
Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology, 66*, 460–473. http://dx.doi.org/10.1037/0022-3514.66.3.460.
Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology, 104*, 103–125. http://dx.doi.org/10.1037/a0030398.
Chen, C., & Wu, C. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education, 80*, 108–121.
Davidson, E. S., & Smith, W. P. (1982). Imitation, social comparison, and self-reward. *Child Development, 53*, 928–932. http://dx.doi.org/10.2307/1129130.
Day, J. (2008). *Investigating learning with web lectures*. Georgia Institute of Technology (Doctoral dissertation). Available from:.
Debono, K. G., & Harnish, R. J. (1988). Source expertise, source attractiveness, and the processing of persuasive information – a functional-approach. *Journal of Personality and Social Psychology, 55*, 541–546. http://dx.doi.org/10.1037/0022-3514.55.4.541.
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 Commission on Physics Education.
Fredette, N. H., & Lockhead, J. (1980). Students' conceptions of simple circuits. *The Physics Teacher, 18*, 194–198. http://dx.doi.org/10.1119/1.2340470.
Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41–58). New York: Academic Press.
Harter, S. (1990). Causes, correlates, and the functional role of global self-worth: a life-span perspective. In R. J. Sternberg, & J. Kolligian (Eds.), *Competence considered* (pp. 67–97). New Haven, CT: Yale University Press.
Hicks, D. J. (1965). Imitation and retention of film-mediated aggressive peer and adult models. *Journal of Personality and Social Psychology, 2*, 97–100. http://dx.doi.org/10.1037/h0022075.
Hoogerheide, V., Loyens, S. M. M., Jadi, F., Vrins, A., & Van Gog, T. (2015). Testing the model-observer similarity hypothesis with text-based worked examples. Educational Psychology. *Advance online publication*. http://dx.doi.org/10.1080/01443410.2015.1109609.
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. http://dx.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*, 69–86. http://dx.doi.org/10.1007/s11251-015-9360-y.
Jakubczak, L. F., & Walters, R. H. (1959). Suggestibility as dependency behavior. *Journal of Abnormal and Social Psychology, 59*, 102–107.
Johnson, C. S., & Lammers, J. (2012). The powerful disregard social comparison information. *Journal of Experimental Social Psychology, 48*, 329–334. http://dx.doi.org/10.1016/j.jesp.2011.10.010.
Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology, 93*, 579–588. http://dx.doi.org/10.1037/0022-0663.93.3.579.
Kalyuga, S., & Renkl, A. (2010). Expertise reversal effect and its instructional implications: Introduction to the special issue. *Instructional Science, 38*, 209–215. http://dx.doi.org/10.1007/s11251-009-9102-0.
Kim, Y., Baylor, A. L., & Reed, G. (2003). The impact of image and voice with pedagogical agents. In *Paper presented at the E-learn (World conference on E-learning in corporate, government, healthcare, & higher education), Phoenix, Arizona*.
Kitsantas, A., Zimmerman, B. J., & Cleary, T. (2000). The role of observation and emulation in the development of athletic self-regulation. *Journal of Educational Psychology, 91*, 241–250.
Korving, H., Hernández, M., & De Groot, E. (2016). Look at me and pay attention! A study on the relation between visibility and attention in weblectures. *Computers & Education, 94*, 151–161. http://dx.doi.org/10.1016/j.compedu.2015.11.011.
Lachner, A., & Nückles, M. (2015). Bothered by abstractness or engaged by cohesion? Experts' explanations enhance novices' deep-learning. *Journal of Experimental Psychology: Applied*. http://dx.doi.org/10.1037/xap0000038.
Liew, T., Tan, S., & Jayothisa, C. (2013). The effects of peer-like and expert-like pedagogical agents on learners' agent perceptions, task-related attitudes, and learning achievement. *Educational Technology & Society, 16*, 275–286.
Littlepage, G. E., Schmidt, G. W., Whisler, E. W., & Frost, A. G. (1995). An

input–process–output analysis of influence and performance in problem-solving groups. *Journal of Personality and Social Psychology, 69*, 877–889. http://dx.doi.org/10.1037/0022-3514.69.5.877.

Ma, X., & Kishor, N. (1997). Attitude toward self, social factors, and achievement in mathematics: a meta-analytic view. *Educational Psychology Review, 9*, 89–120. http://dx.doi.org/10.1023/A:1024785812050.

McDermott, L., & Shaffer, P. (1992). Research as a guide for curriculum development: an example from introductory electricity. *American journal of physics, 60*, 994–1013. http://dx.doi.org/10.1119/1.17003.

Montoya, R. M., & Horton, R. S. (2013). A meta-analytic investigation of the processes underlying the similarity-attraction effect. *Journal of Personal and Social Relationships, 30*, 64–94. http://dx.doi.org/10.1177/0265407512452989.

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*, 186–207. http://dx.doi.org/10.1016/j.cedpsych.2005.05.002.

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

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge, MA: Cambridge University Press.

Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cognitive Science, 38*, 1–37. http://dx.doi.org/10.1111/cogs.12086.

Ridgeway, C. L. (1987). Nonverbal behavior, dominance, and the basis of status in task groups. *American Sociological Review, 52*, 683–694. http://dx.doi.org/10.2307/2095603.

Robert, M. (1983). Observational learning of conservation: Its independence from social influence. *British Journal of Psychology, 74*, 1–10. http://dx.doi.org/10.1111/j.2044-8295.1983.tb01838.x.

Rosenberg-Kima, R. B., Baylor, A. L., Plant, E. A., & Doerr, C. E. (2008). Interface agents as social models for female students: the effects of agent visual presence and appearance on female students' attitudes and beliefs. *Computers in Human Behavior, 24*, 2741–2756. http://dx.doi.org/10.1016/j.chb.2008.03.017.

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research, 57*, 149–174.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*, 207–231.

Schunk, D. H. (2001). Social cognitive theory and self-regulated learning. In B. J. Zimmerman, & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 125–151). Mahwah, NJ: Erlbaum.

Schunk, D. H., & Hanson, A. R. (1985). Peer models: influence on children's self-efficacy and achievement. *Journal of Educational psychology, 77*, 313–322. http://dx.doi.org/10.1037//0022-0663.77.3.313.

Schunk, D. H., & Hanson, A. R. (1989). Self-modeling and children's cognitive skill learning. *Journal of Educational Psychology, 81*, 155–163. http://dx.doi.org/10.1037//0022-0663.81.2.155.

Schunk, D. H., Hanson, A. R., & Cox, P. D. (1987). Peer-model attributes and children's achievement behaviors. *Journal of Educational Psychology, 79*, 54–61. http://dx.doi.org/10.1037/0022-0663.79.1.54.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction, 14*, 293–305. http://dx.doi.org/10.1016/j.learninstruc.2004.06.005.

Schwarzer, R. (1992). *Self-efficacy: Thought control of action*. Washington, DC: Hemisphere.

Shipstone, D. M. (1984). A study of children's understanding of electricity in simple DC circuits. *European Journal of Science Education, 6*, 185–198. http://dx.doi.org/10.1080/0140528840060208.

Simon, S., Ditrichs, R., & Speckhart, L. (1975). Studies in observational paired-associate learning: informational, social, and individual difference variables. *Journal of Experimental Child Psychology, 20*, 81–104.

Smallwood, J., Fishman, D. J., & Schooler, J. W. (2007). Counting the cost of an absent mind: mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin & Review, 14*, 230–236. http://dx.doi.org/10.3758/BF03194057.

Smallwood, J., & Schooler, J. W. (2015). The Science of mind wandering: empirically navigating the stream of consciousness. *Annual Review of Psychology, 66*, 487–518. http://dx.doi.org/10.1146/annurev-psych-010814-015331.

Sonnenschein, S., & Whitehurst, G. J. (1980). The development of communication: when a bad model makes a good teacher. *Journal of Experimental Child Psychology, 3*, 371–390. http://dx.doi.org/10.1016/0022-0965(80)90101-0.

Strauss, S. S. (1978). The influence of peer and adult models upon object oriented children's performance of a task. *Dissertation Abstracts International, 39*, 3484.

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

Szpunar, K. K., Moulton, S. T., & Schacter, D. L. (2013). Mind wandering and education: from the classroom to online learning. *Frontiers in Psychology, 4*, 1–7. http://dx.doi.org/10.3389/fpsyg.2013.00495.

Traphagan, T., Kucsera, J. V., & Kishi, K. (2010). Impact of class lecture webcasting on attendance and learning. *Educational Technology Research and Development, 58*, 19–37. http://dx.doi.org/10.1007/s11423-009-9128-7.

Van Gog, T., & Kester, L. (2012). A test of the testing effect: Acquiring problem solving skills from worked example study. *Cognitive Science, 36*, 1532–1541. http://dx.doi.org/10.1111/cogs.12002.

Van Gog, T., Kester, L., Dirkx, K., Hoogerheide, V., Boerboom, J., & Verkoeijen, P. P. J. L. (2015). Testing after worked example study does not enhance delayed problem-solving performance compared to restudy. *Educational Psychology Review, 27*, 265–289. http://dx.doi.org/10.1007/s10648-015-9297-3.

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*, 212–218. http://dx.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*, 16–26. http://dx.doi.org/10.1080/00461520701756248.

Van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22*, 155–174. http://dx.doi.org/10.1007/s10648-010-9134-7.

Van Gog, T., Verveer, I., & Verveer, L. (2014). Learning from video modeling examples: effects of seeing the human model's face. *Computers & Education, 72*, 323–327. http://dx.doi.org/10.1016/j.compedu.2013.12.004.

Van der Meij, H., & Van der Meij, J. (2013). Eight guidelines for the design of instructional videos for software training. *Technical Communication, 60*, 205–228.

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychological values by medical students: a test of self-determination theory. *Journal of Personality and Social Psychology, 70*, 767–779. http://dx.doi.org/10.1037/0022-3514.70.4.76.

Yi, M. Y., & Hwang, Y. (2003). Predicting the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *International Journal of Human-Computer Studies, 59*, 431–449. http://dx.doi.org/10.1016/S1071-5819(03)00114-9.

Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology, 94*, 660–668. http://dx.doi.org/10.1037/0022-0663.94.4.660.

Zmyj, N., Aschersleben, G., Prinz, W., & Daum, M. (2012). The peer model advantage in infants' imitation of familiar gestures performed by differently aged models. *Frontiers in Psychology, 3*, 252. http://dx.doi.org/10.3389/fpsyg.2012.00252.

Zmyj, N., & Seehagen, S. (2013). The role of a model's age for young children's imitation: a research review. *Infant and Child Development, 22*, 622–641. http://dx.doi.org/10.1002/icd.1811.