The Effect of Interpolated Questions in Video Modeling Examples on Learning Outcomes

Lisette Bosveld (5385628)

Faculty of Social Sciences, Universiteit Utrecht

Master's thesis

V. Hoogerheide

Second assessor: C. Hulshof

30 June 2023

Word count: 7342

Abstract

Instructional videos that explain a procedure are a popular way of delivering instruction. Literature shows that a sequence of examples and practice problems is equally effective as only studying examples. To smooth the transition from example study to problem-solving, questions can be interpolated in a video modeling example (VME). This study researches the questions: Does enriching VMEs with interpolated questions improve students' learning outcomes? Does asking questions during or after a VME improve learning outcomes more? 237 secondary school students were randomly allocated to one of three experimental conditions: interpolated questions, questions after the VME, and a control condition. A posttest assessed the students' comprehension and application. Interpolating questions significantly increased overall performance ($F(2,233) = 3.133$, $p = .045$, $\eta^2 = .026$), comprehension ($F(2,233) = 3.133$, $p = .045$, $\eta^2 = .026$) and application ($F(2,233) = 4.377$, $p = .014$, $\eta^2 = .036$). Asking questions after the VME increased application significantly. Fading, increased engagement, decreased mental effort and the testing effect can explain this conclusion. School and group also influenced learning outcomes. Limitations include ecological validity and lack of a delayed posttest. With these findings, teachers can increase the effectiveness of their instruction.

The Effect of Interpolated Questions in Video Modeling Examples on Learning Outcomes

In 2020, the Dutch government decided to spend 15 million EUR on the digitalization of secondary schools (VO Raad, 2020). It has become much more common to use digital educational material. Digital educational materials are easily accessible and flexible. One part of digital educational material consists of instructions. According to De Koning et al. (2018), videos are a popular way of delivering online instructions. Video instructions commonly aim to convey procedural content (Bétrancourt & Benetos, 2018; Fiorella & Mayer, 2018). Usually, the instructor provides an example in the video.

Example-Based Learning

Examples are central to example-based learning, specifically examples that show students how to solve certain problems (van Gog et al., 2011). In example-based learning two types of examples exist: worked examples (written step-by-step solutions; Baars et al., 2014; van Gog & Rummel, 2010) and modeling examples (van Gog et al., 2014). This study focusses on video modeling examples (VMEs), a subcategory of modeling examples.

In a VME, an instructor explains or demonstrates a way of performing a task (Fiorella et al., 2017; van Wermeskerken et al., 2018). VMEs frequently involve verbal explanations, but these are not essential (van Wermeskerken et al., 2018). The instructor is not necessarily visible (van Wermeskerken et al., 2018). VMEs can be used in highly structured domains, such as science, technology, engineering and mathematics, as well as in in less structured situations, like social and meta-cognitive skills (van Gog & Rummel, 2010).

Effectiveness of Example-Based Learning

Observing how others perform a task is a powerful way of learning, because the learner has to actively interpret the behavior and integrate it with their prior knowledge (Fiorella et al., 2017). VMEs have proven to be effective in acquiring new skills (Hoogerheide et al., 2016). Modeling examples seem to be particularly effective for novices (e.g. Kirschner et al., 2006; van Gog & Rummel, 2010).

An explanation for the effectiveness of examples for novices can be found in cognitive load theory, which suggests that studying a worked example requires less mental effort than problem-solving (Kant et al., 2017). Mental effort is the cognitive capacity that is used to deal with the task's demands (Paas et al., 2003). Van Harsel et al. (2020) expound on this explanation by suggesting that modeling examples require less mental effort from novices than solving practice problems, because novices have not yet developed strong problem-solving strategies and acquired the necessary cognitive schemata. When novices need to invest less mental effort in searching for the right strategy, more mental capacity remains to understand the process and acquire problem-solving strategies.

Downsides of Example-Based Learning

Yet, one might argue that only providing examples is not the most beneficial to learners' results. The arguments listed below can be used to support said statement. First, watching VMEs is a more passive form of learning than solving practice problems (van Harsel et al., 2020). Active learning strategies lead to better academic achievement than passive learning strategies, because active learning strategies require, rather than stimulate, the learner to build cognitive schemata (Brame, 2016; de Jong, 2019). According to Chi (2009), an activity can be considered active when the student can interact with the new information or environment. Passive learning strategies do not allow interaction.

Second, studying examples might lead to learners overestimating their understanding of the subject matter (Baars et al., 2017; Lowe, 2004). Lowe (2004) suggests that learners who studied a VME tend to overestimate their understanding, because watching a video requires little investment of mental effort. Learners use the invested amount of mental effort to assess their understanding of the subject matter (van Gog et al., 2020). Thus, when learners watch a VME, which requires little mental effort, they assume they understand the subject matter. Another explanation is that the learner recognizes the example and therefore pays less attention to it (Anderson et al., 2019).

Third, heavily guided instruction, such as example study, can hamper learning (K. Huang et al., 2015; Y.-H. Huang et al., 2015; Rey & Buchwald, 2011). When learning, students build problem-solving schemata (van Harsel et al., 2020). Learners who already acquired such schemata do not need the guidance of an example (Rey & Buchwald, 2011; Sweller, 2011). Novices do not have these schemata yet (van Harsel et al., 2020), therefore needing to invest more mental effort to understand the subject matter (Kirschner et al., 2006). This effect is known as the expertise reversal effect (Atkinson & Renkl, 2007; Kirschner et al., 2006). Literature is inconclusive on the cause of the expertise reversal effect (Rey & Buchwald, 2011).

The Added Value of Practice Problems

Example-based learning solved said issues by adding practice problems to a learning sequence. A typical learning sequence thus involves both examples and practice problems. Practice problems are usually similar to the shown example(s) and do not include instructions on how to solve the problem (Baars et al., 2017; Carpenter, 2012; van Harsel et al., 2019). A learning sequence overcomes the downsides of example study in the following ways.

The passivity of example study is complemented by the activity of solving a practice problem. Practice problems stimulate deep cognitive processing, as it integrates new knowledge in existing structures, and strengthens the connection between fragments of knowledge by stimulating organization (Fiorella & Mayer, 2016). Practice problems allow the learner to interact with the subject matter, therefore transforming the passive learning activity into an active one and increasing learning gains (Chi et al., 2018).

Second, practice problems stimulate students to accurately assess their learning. Adesope et al. (2017) suggest that practice problems increase results, because a practice problem uses similar mental processes as a final examination. Practice problems provide the learner with the opportunity to reflect on their understanding, which would result in more accurate judgements of learning (Baars et al., 2017).

The third issue with examples is the occurrence of the expertise-reversal effect. Practice problems are less guided than examples and therefore accommodate the needs of more experienced learners. Also, solving practice problems stimulates organization of knowledge (Fiorella & Mayer, 2016). Organizing the new and existing knowledge into a comprehensive whole is essential to understanding the subject matter (Mayer, 2014).

Sequencing Examples and Practice Problems

Van Gog et al. (2011) studied what learning sequence would yield the best learning gains. They concluded that studying worked examples only and example-problem pairs were significantly more efficient than practice problems only and problem-example pairs, as did Kant et al., (2017), Van Gog et al. (2015) and Van Harsel et al. (2020). This conclusion contrasts with the literature described above, which suggests that example study and practice problems support each other, leading to increased learning gains compared to example study only. This study will look into the effects of combining example study and practice problems in a different way, namely by interpolating the two.

Interpolating Questions

Interpolating practice problems and example study can be done in various ways. Huang, Lin et al. (2015) presented the problem and the example simultaneously. Mirriahi et al. (2021) asked questions during the video about how the information applied to other contexts. In this study, interpolated questions are pop-up questions during the VME, in line with Mirriahi et al. (2021). These interpolated questions concern the procedure explained in the instructional video. Arguments for interpolating questions in examples include engagement, mental effort and the fading principle.

Increased Student Engagement

Interpolating questions is said to increase engagement (Haagsman et al., 2020; Rice et al., 2019). Students tend to lose focus while watching educational videos (Szpunar et al., 2013). When students do not pay attention, they are not engaged, whilst attention is critical for academic functioning (Anderson et al., 2019). Some say the term 'engagement' is so widely used for different constructs, that the term in itself has become meaningless (Whitton & Moseley, 2014). A common factor in definitions of engagement is that it involves the student performing an action. By interpolating questions, the students have to actively click the correct answer, therefore increasing engagement compared to watching.

Engagement influences learning gains (Nair & Mathew, 2022). It appears that learning gains increase when students' engagement increases (Chi et al., 2018; de Jong, 2019; Imlawi, 2021). Cai and Liem (2017) suggest engagement is essential for academic achievement. Little is known, however, about how engagement increases learning gains (Cohen et al., 2018). Based on the current understanding, engaging students by interpolating questions in a VME is likely to yield positive results.

Decreased Mental Effort

A second reason to interpolate questions in VMEs is that failing to do so increases the amount of mental effort the learner needs to invest. Learning will increase if problem-solving would require less mental effort (Agarwal, 2019). Interpolating pauses or questions prevent an activity from requiring too much mental effort (Mayer & Pilegard, 2014). The human brain can process small amounts of information at once (Ibrahim et al., 2012). Interpolating questions provides the learner with time to process the subject matter (Fiorella & Mayer, 2018; Ibrahim et al., 2012) and structure (Merkt et al., 2018). This segmenting effect seems to yield positive results, especially for procedural learning (Biard et al., 2018; Ibrahim et al., 2012; Mayer, 2021). Thus, by interpolating questions in the VME, the example is segmented and structured, which would lead to greater learning gains.

Fading Principle

Most students use ineffective strategies when studying examples (Atkinson et al., 2003; Renkl et al., 2004). Fading, a procedure during which instructional support decreases (Eiriksdottir & Catrambone, 2011), integrates example study and problem-solving, allowing for opportunities to interact with the subject matter (Renkl et al., 2002). Fading presents information in small chunks that increase in size (Lange & Costley, 2019), which requires learners to gradually invest more mental effort (Eiriksdottir & Catrambone, 2011). Fading offers structure (de Jong, 2019; Hmelo-Silver et al., 2007) and forces the student to think about what they are doing, therefore stimulating effective processing (Hmelo-Silver et al., 2007). Fading reduces errors during learning and increases transfer (Eiriksdottir & Catrambone, 2011; Renkl et al., 2002). It does not increase time on task significantly, but does increase learning gains (Atkinson et al., 2003) while accounting for students' prior knowledge (van Gog & Rummel, 2010). Taking the students' prior knowledge into account prevents the expertise-reversal effect, while making difficult tasks accessible to novices (Hmelo-Silver et al., 2007). In summary, interpolating questions is a way of fading the transition from example study to problem-solving, therefore stimulating effective processing of VMEs because the superfluous mental effort decreases and the student is forced to thoroughly process the example.

Present Study

The present study aims to answer the question: Does enriching VMEs with interpolated questions improve students' learning outcomes? The second research question elaborates hereon: Does asking questions during or after a VME improve learning outcomes more? To answer these questions, secondary school students studied a VME and answered questions during or after the video, or none at all. A posttest measured the effects on learning outcomes, specifically the student's ability to solve a similar problem on their own (comprehension) and their ability to use the same procedure in a situation that differs from the examples (application).

I hypothesize that asking questions during VMEs positively impacts learning outcomes when compared to not asking questions during the VME. My second hypothesis is that interpolating questions increases learning outcomes more than asking questions after the instruction, with either timing yielding greater gains than not asking any questions at all.

This study differs from the existing literature by interpolating questions in a VME and by providing an example-based learning perspective on interpolated questions. Often, studies that incorporate examples and problems let students solve entire problems (e.g. Y.-H. Huang et al., 2015; van Harsel et al., 2020), whereas the current study integrates the two. With the conclusions of this study, teachers can improve their online practice, resulting in higher student performance.

Method

Design and Participants

The experiment consisted of three phases: a prequestionnaire, an experiment during which students watched a mathematical instruction and a posttest. The participants, in their first or second year of secondary school, were allocated to an experimental condition: a group that watched the video with interpolated questions (during; group D), a group that answered questions after the video (after; group A), or a group that only watched the video (control; group C). The students were quasi-randomly divided, controlling for gender and year of secondary school. Despite the decreased gender gap in mathematical performance in the Netherlands, boys generally outperform girls in mathematics (OECD, 2019). To indicate the difference between a common practice (Alkema et al., 2015) and the experimental groups, group C was included, albeit differing in study time.

To determine the necessary amount of participants, a G*Power test was conducted (Faul et al., 2007). The estimated effect size was set to .25, the α error probability to .05 and the expected power to .8. This resulted in an estimate of 160 participants. To obtain 160 participants, eleven secondary schools located in province Utrecht, the Netherlands, were approached. These schools were selected on the educational levels they offer. In total, 236 students from three schools participated. Demographic information can be found in table 1.

Table 1

Demographic Information

Condition	N	N_{male}	N_{female}	$N_{unknown}$	N_{class1}	N_{class2}	M_{age}	SD_{age}
After VME	75	25	48	2	53	22	12.96	0.796
During VME	82	27	49	6	59	23	12.82	0.877
Control	79	31	45	3	54	25	12.87	0.868

Note. $N_{unknown}$ is the number of students that did not wish to disclose their gender. N_{class1} is the number of students that are currently in their first year of the HAVO. N_{class2} is the number of students that are currently in their second year of the VMBO.

Instruments

Prequestionnaire

A prequestionnaire provided insight in the participants' prior knowledge. The alternative, objective testing prior to instruction, positively influences learning outcomes (Brink, 2013; Gyllen et al., 2021), therefore leading to invalid results on a posttest. A questionnaire instead of a pretest also decreases the chance of students losing motivation or self-efficacy (van Harsel et al., 2020). However, a self-report questionnaire tends to have lower reliability due to students' incapability to judge themselves accurately (Caspersen et al., 2017).

The prequestionnaire consisted of six items and was administered using Qualtrics (www.qualtrics.com). It followed the structure of Fiorella and Mayer's (2013) pretest. First, the students had to indicate on a five point Likert-scale (1 very low – 5 very high) how much prior knowledge they had of performing calculations with radicals. Next, they had to indicate if they agreed with each of the following statements: "I know what simplifying means", "I know what radicals with the same radicand are", "I know how to simplify radicals", "I know what a radical is" and "I performed calculations with radicals before."

Instructional Video

The 5:15 minute video was about addition and subtraction of radical expressions (Math with Menno, 2020) and was used with permission from the creator. A (translated) transcript can be found in Appendix B. Mathematics offer opportunities for conceptual questions and procedural questions, hence a mathematical video. The topic was chosen from HAVO year 2, to make sure that no HAVO year 1 students previously studied the topic, but had sufficient prior knowledge to understand the matter.

The video started with the learning objective, after which the process of addition of radicals was explained, followed by a VME. The experimental groups all had the same experience up till this point. Next, five more examples were provided. Group D students answered questions for all five examples, which popped up automatically before each example. These questions concerned the next step in simplifying the expressions. Group C and A continued watching the video uninterrupted. When the video was finished, group A answered the questions group D answered during the video. Group C did not answer any questions about the video. The questions can be found in Appendix C.

All students watched the instructional video with the program Playposit (https://www.playposit.com/; see figure 1), which allows teachers to interpolate questions in videos. It keeps track of and marks the students' answers and allows for automatic predetermined feedback. The researcher developed the questions, a mathematics teacher checked if these questions matched the students' level.

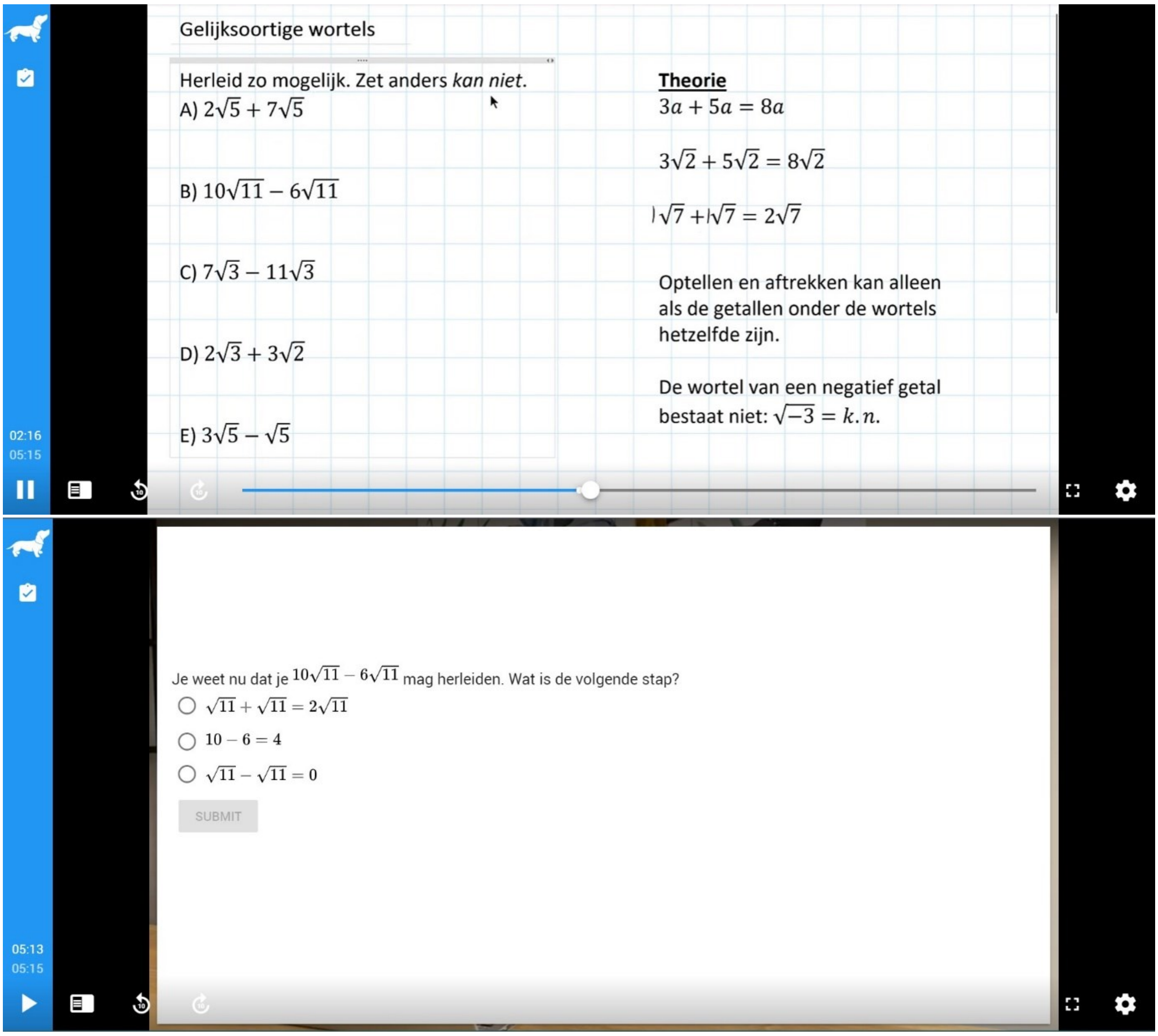


Figure 1

Screenshots of Playposit

Posttest

The posttest (Appendix D) included two assignments, one that was similar to the examples from the video, and one that had a different context than the examples in the video. According to Pellegrino et al. (2016), assessment should be based on what the students need to learn to determine tasks that elicit the responses that inform the assessor of the students' understanding. Recognizing when a certain strategy should be used and using it is one of the main aims of the Dutch mathematical curriculum (Besluit kerndoelen onderbouw VO, 2006). In the MATH taxonomy, a taxonomy for mathematical questions derived from Bloom et al.'s (1956) taxonomy, these skills can be classified as comprehension and application respectively. Smith et al. (1996) suggest these can be tested by asking the students to perform a routine task (comprehension) and by asking the students to solve a similar problem in a new context (application), also known as isomorphic problems (van Gog et al., 2015).

The posttest was presented on paper to maximize similarity to a regular summative test. Another reason for a test on paper was that Qualtrics did not offer the possibility to integrate mathematical symbols in answers.

The questions for the post-test and the marking sheet (Appendix E) were developed by the researcher. An independent teacher checked the posttest and checked the marking sheet used to mark the posttest. Additionally, 30 posttests were also marked by an independent teacher to calculate the inter-rater reliability.

Procedure

The ethical board of the UU approved the setup of this study. Two weeks before the experiment commenced, the caretakers received an e-mail (Appendix A) requesting permission for their child to participate. The students were asked for their informed consent at the start of the experiment. The students participated anonymously and their teacher had no access to the students' answers.

The experiment was conducted at the students' schools and took approximately 50 minutes. The researcher explained what the students could expect, whereafter the students received a link to the prequestionnaire. Next, a questionnaire was administered, after which the students were randomly divided over the three conditions, controlling for gender and class. The students could ask questions about the procedure, but not about the content. They were allowed to take notes, which simulates a real educational context. Students were told not to pause the video, and could not rewind the video. Pausing and rewinding was not allowed, because that would greatly influence students' study time and therefore their results. Playposit also tracked the time on task. Lastly, the students completed the post-test on paper.

Analysis

All analyses were performed using IBM SPSS Statistics (version 28.0.1.1) and with a level of significance of .05.

Prequestionnaire

The maximum score was 10, consisting of a maximum of 5 points for the general indication and one point for each statement the student agreed with. The points per student were used to carry out an analysis of variance-test. An analysis of variance can be used to determine whether the means of different experimental groups differ significantly (Field, 2018). As the prequestionnaire is a self-report questionnaire, which generally have poor validity (Boeije et al., 2009), it will be left out of further analysis.

Posttest

The posttest was marked, leading to a total score per student. These were used to carry out a one-way analysis of variance for three conditions to compare the posttest means of the three groups. An analysis of variance can be used when the independent categorical variable (experimental group, with the categories A, D and C) leads to a dependent interval or ratio variable (Troncoso Skidmore & Thompson, 2013). If the means differ significantly ($p < .05$) a post hoc analysis will be performed to determine what groups differ.

The posttest had a total Cronbach's alpha of .732 for nine items, which is acceptable (Taber, 2018). For comprehension only, the Cronbach's alpha was .705 for seven items, and for application .772 for two items, which are acceptable too. The interrater reliability was calculated with 31 posttests rated by two independent teachers and the researcher. As Cohen's kappa can only be used to calculate the inter rater reliability between two raters (Mandrekar, 2011; Nichols, 2021), an alternative had to be found. To calculate the intraclass reliability, the following settings were used: two-way random single measures, absolute agreement. This resulted in a intraclass reliability of .97, with $F(30,60) = 98.985$, $p < .001$, 95 % CI [.947 - .984]. This is indicates excellent reliability (Koo & Li, 2016). Interestingly, the intraclass reliability differed for the questions on comprehension (ICC = .989) and application (ICC = .903). In most of the cases where scores differed, it concerned how the student wrote down their calculations. The comprehension questions did not involve entire calculations, but only required a final answer. When students have to write down their calculations, there is more room for notation ambiguity. Mathematical assessment includes assessment on the students' ability to convey mathematical ideas (Görgüt & Dede, 2022). What one researcher might have found an acceptable way of communicating mathematical ideas, might be unclear for another. However, interrater reliability remained very high and therefore did not pose a problem for further analyses.

Results

An overview of posttest scores can be found in table 2. For the analyses of variance, η^2 is used to indicate effect size, as is common for analyses of variance (Norouzian & Plonsky, 2018). Effect sizes were deemed small if over .01, medium if over .09 and large if over .25. These numbers are based on the explained variance norms for Pearson's r, as can be found in Field (2018). The effect sizes for multiple regression are reported as Pearson's r. Effect sizes were deemed small if over .10, medium if over .30 and large if over .50 (Field, 2018).

For post-hoc procedures, Gabriels and Games-Howell were chosen, because sample sizes varied slightly and Games-Howell procedure is robust against unequal population variances (Field, 2018). To determine the effect sizes, Cohen's d was calculated with unpooled standard deviations. The effect size was deemed small if over 0.2, medium if over 0.5 and large if over .08.

Table 2

Overview of Posttest Scores

	M	SD	$M_{Comprehension}$	$SD_{Comprehension}$	$M_{Application}$	$SD_{Application}$
After	8.61	3.304	5.08	1.819	3.13	2.056
During	8.87	3.377	5.83	1.871	3.04	2.111
Control	7.33	3.445	5.48	2.030	2.25	1.958
Total	8.27	3.443	5.47	1.927	2.81	2.072

Prior Knowledge

A one-way analysis of variance was carried out after all assumptions were checked. All assumptions were met. Prior knowledge, as tested in the prequestionnaire, did not differ significantly per condition ($F(2,233) = 0.799$, $p = .451$, $\eta^2 = .007$). This indicates that a difference in self-reported prior knowledge is unlikely to have influenced the outcome of other analyses.

Distribution of Sex Over Conditions

Because boys generally outperform girls concerning mathematics (OECD, 2019), a Chi-Square test was conducted to determine if the genders were equally distributed over the conditions. All assumptions were met, excluding one. The number of students that did not wish to disclose their gender per condition was too small, therefore violating the requirement of sample size (McHugh, 2013). However, since over 80 % of the cells of the test could be filled, the test could be carried out (McHugh, 2013). The analysis was non-significant (X^2 (4, 236) = 2.889, $p = .577$), indicating that the genders were equally distributed and therefore is unlikely to have influenced the outcome of other analyses.

Time Spent on Learning per Condition

To determine if the time spent on the learning phase of the experiment, studying the VME and answering questions, was equal across conditions, a one-way analysis of variance was carried out after all assumptions were checked. All assumptions were met. Only cases that had a realistic time on task, between 5 and 50 minutes, were included. The time students spent on watching the video and answering the (interpolated) questions differed significantly per condition ($F(2,243) = 13.816$, $p < .001$, $\eta^2 = .102$). Post hoc tests revealed that Group C differed significantly from group D, with a mean difference (MD) of -268.644, $p < .001$, $d = 0.68$. Group C also differed significantly from group A ($MD = -246.564$, $p < .001$, $d = .70$). Group D and A did not differ significantly ($MD = 22.080$, $p = .970$, $d = .07$).

Effect of Condition on Overall Performance

A one-way analysis of variance was carried out to determine the overall effect of condition on posttest performance. All assumptions were met. A significant effect of condition on posttest performance was found ($F(2,233) = 4.731$, $p = .010$) with a small effect ($\eta^2 = .039$). Hence, post hoc tests were conducted. Group D performed significantly better than group C ($MD = 1.537$, $p = .013$, $d = 0.45$). Group A did not perform significantly better than group C ($MD = 1.284$, $p = .056$, $d = 0.37$) and not significantly worse than Group D ($MD = 0.253$, $p = .884$, $d = 0.08$).

Effect of Condition on Comprehension

After checking the assumptions, a one-way analysis of variance was conducted to determine whether condition influenced students' comprehension. All assumptions were met. Condition significantly influenced students' comprehension ($F(2,233) = 3.133$, $p = .045$), with a small effect size ($\eta^2 = .026$). Hence, a post-hoc test was conducted. Group D scored significantly higher on comprehension than group C ($MD = 0.753$, $p = 0.39$, $d = 0.17$). Group

A did not score significantly worse than group D (*MD* = -0.349, $p = .464$, $d = -0.37$) and not significantly better than group C (*MD* = 0.404, $p = .310$, $d = -0.20$).

Effect of Condition on Application

A one-way analysis of variance was carried out to determine whether condition influenced students' application. All assumptions were met. Condition significantly influenced application ($F(2,233) = 4.377$, $p = .014$, $\eta^2 = .036$) with a small effect size. Gabriels post hoc test was used to determine the cause of the significant result. Group D significantly outperformed group C (*MD* = 0.783, $p = .046$, $d = 0.40$), as did group A (*MD* = 0.880, $p = .024$, $d = 0.45$). Group D and A did not perform significantly different (*MD* = -0.097, $p = .954$, $d = 0.05$).

Other Analyses

To eliminate as many possible causes for a significant result aside from condition, several other analyses were carried out. While conducting the experiment, it was observed that the learning climate in each of the schools and groups differed. Despite the conditions being allocated equally in each of the groups and schools, school and group still might have influenced the posttest performance. Hence, the effect of school and group were analyzed as well.

Effect of School and Group on Posttest Performance

To determine the effect of school on posttest performance a linear regression analysis was performed. All assumptions were met. An overview of posttest scores per school can be found in table F1. School explained 18.9 % of the variance ($R^2 = .189$, $F(2,234) = 27.352$, $p < .001$) and had a medium effect ($r = .435$). For Comprehension, school explained 14.2 % of the variance ($R^2 = .142$, $F(2,234) = 19.316$, $p < .001$, $r = .38$). For Application, school explained 13.8 % of the variance ($R^2 = .138$, $F(2,234) = 18.687$, $p < .001$, $r = .37$).

To determine if group made a difference, a linear regression analysis was conducted. All assumptions were met. Table F2 provides an overview of posttest scores per group. Group explained 26.3 % of the variance in overall posttest scores ($R^2 = .263$, $F(11,225) = 7.286$, $p < .001$) and has a large effect ($r = .513$). Group explained 20.9 % of the variance of Comprehension scores ($R^2 = .209$, $F(11,225) = 5.398$, $p < .001$, $r = .46$) and 19.4 % of the variance of Application ($R^2 = .194$, $F(11,225) = 4.926$, $p < .001$, $r = .44$).

Condition explained, together with group and school, 30.9 % of the variance ($R^2 = .309$, $F(13,223) = 7.670$, $p < .001$) and had a large effect on overall posttest performance ($r = .556$). Condition, school and group together explained 24.2 % of the variance of Comprehension ($R^2 = .242$, $F(13,223) = 5.469$, $p < .001$, $r = .49$) and 23.5 % of the variance of Application ($R^2 = .235$, $F(13,223) = 5.259$, $p < .001$, $r = .48$). In these last three models, condition was added first, thereafter school and lastly group.

Other Observations

Aside from the data collection, the researcher observed some peculiarities. These are listed in this paragraph. Some students said they found this topic easier than their regular subject matter. One student even asked if the addition and subtraction of radicals could be the subject of their next test because she found it so easy. Some classes found it very hard to work quietly or individually. Multiple students mentioned they found studying the video and answering questions was more fun than their regular lessons. The teachers that cooperated expected that interactive videos are more effective than regular videos. Not all students took notes. This simulates a real life environment, but could have influenced results. Some students continued studying the video without sound even after repeatedly asking and telling them to listen to the video as well. These cases were rare.

Discussion

In order to offer a different perspective on integrating examples and practice problems and improve video instructions, this study aimed to answer the question: Does enriching VMEs with interpolated questions improve students' learning outcomes? The results indicate that asking questions during video modeling examples (VMEs) positively impacts learning outcomes when compared to not asking questions during the VME. This is in line with the first hypothesis of this study. The second research question was: Does asking questions during or after a VME improve learning outcomes more? Interpolating questions increases learning outcomes more than not asking questions, but asking questions after the VME does not differ significantly from interpolated questions and not asking questions. Interpolating questions increases both comprehension and application, whereas asking questions after the VME increases application only. The second hypothesis is therefore partially confirmed.

It appears that interpolated questions help students understand the subject matter. The question remains how interpolated questions increase learning gains. From an example-based learning perspective, arguments for including and interpolating questions included fading, decreased mental effort and increased engagement. Fading is the gradual reduction of support during the learning process (Eiriksdottir & Catrambone, 2011). This study used forward fading, in which the student is guided through the problem chronologically (Renkl et al., 2004). Fading provides structure in knowledge acquisition (de Jong, 2019; Hmelo-Silver et al., 2007) and requires the learner to solve a problem increasingly on their own (Hmelo-Silver et al., 2007; van Gog & Rummel, 2010). Thus, interpolated questions make practice problems accessible to novices that have not yet developed strong problem-solving strategies, which is a risk of practice problems (van Harsel et al., 2020). Interpolated questions provide the learner with time to process the subject matter and therefore decrease the amount of mental effort needed to solve a problem (Fiorella & Mayer, 2018; Ibrahim et al., 2012). Interpolating might also increase learning outcomes because it increases engagement (Haagsman et al., 2020), because the student has to actively respond to the VME. Interpolating the questions transforms the passive activity of studying a VME into an active one, because students have to physically interact with the subject matter (Chi, 2009).

Another possible explanation for the effectiveness of interpolated questions can be found in assessment literature. By interpolating questions, the student can assess their understanding during the instruction, thereby counteracting the downside of example-based learning that students easily overestimate themselves when studying an example (Baars et al., 2017). The questions cause the student to reflect on their understanding and form a judgement of learning (Baars et al., 2017). Pulukuri and Abrams (2021) found that interpolated questions increase university students' monitoring accuracy, a student's ability to assess their understanding of the subject matter. Students can use this judgement to adjust their learning process (Weurlander et al., 2012), for example by paying extra attention to the next part of the instruction. This can be viewed as a form of formative assessment (Weurlander et al., 2012), in which assessment leads to adjustment of learning (Black & Wiliam, 2010; Clark, 2012). Formative assessment has proven to have a positive influence on learning, and is often characterized as an interaction between student and teacher (Heritage & Heritage, 2013). It can help the student reflect on their learning and fosters self-regulated learning (Clark, 2012; Pulukuri & Abrams, 2021). Van Alten et al. (2020) found that interpolating videos with prompts that stimulate self-regulated learning increases learning outcomes in secondary school students. Self-regulated learning is seen as a key competency for lifelong learning (Dignath & Veenman, 2021) and is associated with greater achievement (Jansen et al., 2019; van Alten et al., 2020).

Interestingly, interpolating questions increases comprehension and application, whereas asking questions after a VME increases application only. Interpolated questions might have lead to better comprehension than asking questions after a VME, because students who answered interpolated questions did not have to remember the problem solving procedure for as long as the group that answered questions after the VME. Remembering large pieces of information requires greater investment of mental effort than smaller pieces (Ibrahim et al., 2012). Additionally, the students could not interact with the subject matter while they still remembered details of the problem solving procedure. Interacting with subject matter helps students understand it (Chi et al., 2018). The students who answered interpolated questions on the other hand were guided through the problem solving process. The interpolated questions required the student to perform more of the problem solving procedure step by step (Lange & Costley, 2019), which requires learners to gradually invest more mental effort (Eiriksdottir & Catrambone, 2011) and allows students to interact with the subject matter. Recalling the entire problem solving process might have required more mental effort than the student could invest, therefore hindering the student's comprehension (Kant et al., 2017).

An alternative explanation for the conclusion that interpolating questions leads to better comprehension than asking questions after a VME is that it reduced the split attention effect. The split attention effect occurs when the student needs the information from two or more spatially or temporally separate sources in order to solve a problem (Sweller et al., 2011; Tindall-Ford et al., 2015). Mutlu-Bayraktar et al. (2022) observed more brain activity in students when the split-attention effect occurred. When the sources are integrated, as is the case with interpolated questions, fewer mental effort is needed and learning increases (Pouw et al., 2019; Schroeder & Cenkci, 2018). However, Schroeder and Cenkci (2020) found no evidence in their meta-analysis for differences in mental effort when the split-attention effect occurred. Instead, they suggest that the theoretical basis for the split-attention effect has to do with the student actively engaging with and organizing the information, such as Mayer (2014) proposed. It might be the case that interpolating questions in VMEs helps students organize the problem solving procedure, therefore leading to better comprehension.

The current study argued that interpolating questions in VMEs increases learning outcomes in secondary school students and used an example-based learning perspective to support the hypotheses. The results confirmed that interpolated questions in VMES increase learning outcomes, in line with previous research by Pulukuri and Abrams (2021). They compared the effect of textbook learning with interactive videos on learning outcomes, monitoring accuracy and self-regulated learning in 81 university students. Their results indicate that interpolated questions increase learning outcomes, lead to more accurate judgements of learning and to better self-regulated learning. These results support the explanation of the positive effect of interpolated questions on learning outcomes because of better judgement of learning and increased self-regulated learning. In contrast, Mirriahi et al. (2021) compared the effect of in-video annotations to interpolated questions on self-efficacy and learning of 93 university students in a laboratory setting. They found that annotations and interpolated questions did not increase learning outcomes, but did influence self-efficacy. The current study differs from previous work by using a larger sample from a different population, namely secondary school students instead of university students.

Limitations

A limitation of this study concerns the duration. The posttest was conducted immediately after the students watched the instruction. However, in the classroom it seldom happens that students are tested immediately after instruction. Research has shown that conducting a posttest immediately after the instruction can lead to different results than a delayed posttest (Carpenter, 2012; Latimier et al., 2021). This holds true especially for retrieval practice, in which students need to recall information exactly as it was presented (Agarwal, 2019; Latimier et al., 2021). For retrieval practice, it has been concluded that a delayed posttest showed better retention of information than an immediate posttest (Agarwal, 2019; Cepeda et al., 2009; Karpicke, 2012; Roediger & Karpicke, 2006). However, the posttest in the current study only concerned comprehension and application questions, which are higher order skills than retrieval of information (Agarwal, 2019; Kinnear et al., 2020). Future research needs to determine if a delayed posttest leads to similar conclusions about the increase in comprehension and application after interpolating questions in a VME.

In a similar vein, one can question the design of the control condition. The control condition did not answer any questions other than the posttest and thus spent less time learning than the other conditions. Increased learning time can increase learning outcomes (Scherer et al., 2015), while spending more time-on-task can also indicate poor learning outcomes (Vörös et al., 2021). Goldhammer et al. (2014) suggest that the effect of time on task differs per individual, based on the difficulty of the task and the individual's skill level. The analysis in the current study lead to the conclusion that the control condition spent significantly less time-on-task, but the conditions that answered questions did not differ in time-on-task significantly. Because of the insignificant difference between the two experimental conditions, time-on-task is controlled for. The control condition was included as a starting point, because it is common practice in education (Alkema et al., 2015). Simply asking students to watch a video is the easiest way to implement VMEs in education, since it does not require much effort from the teacher. The comparison between watching a VME and interpolating questions shows how the simplest way of using VMEs can be enhanced, while controlling for time-on-task. However, future research can investigate other ways of controlling for time-on-task.

An aspect of this study that can be seen as both a limitation and a strength is that the experiment was conducted in real-life classrooms. Using a real-life setting, in this case classrooms, increases the ecological validity (Holleman et al., 2020). Ecological validity concerns the relation between the experimental setting and the real world (Schmuckler, 2001) and is often defined as the extent to which the characteristics of the experimental setting are similar to the setting in which the studied behavior naturally occurs (Bronfenbrenner, 1977; Kihlstrom, 2021). According to Holleman et al. (2020), it is important to describe and compare the exact cues and context that are used in the experiment to the setting in which the studied behavior naturally occurs, in order to determine the ecological validity of the experiment. The classroom is one of the environments in which VMEs are used, thus involves similar cues and context as the natural setting, suggesting high ecological validity. At the same time, using a real-life situation does not guarantee high ecological validity (Kihlstrom, 2021). A classroom is not as controlled as a laboratory setting (Falk & Heckman, 2009). For example, a few students continued studying the video without sound even after repeatedly asking them to listen to the video as well. Without sound, the students missed the explanation, which might have influenced some of the results. Tight control of the experimental setting reduces noise in the data (Falk & Heckman, 2009) and thus can increase content validity (Gerritsen-van Leeuwenkamp et al., 2017).

Related to the ecological validity of the study is its generalizability. Results showed that school and group influenced learning outcomes. However, only three schools agreed to participate. The finding that school influences learning outcomes is in line with research on school culture (Thapa et al., 2013; Wang & Degol, 2016). School culture generally concerns four domains, namely academic climate, community, safety and institutional environment (Thapa et al., 2013; Wang & Degol, 2016). School culture affects learning through those domains. The exact effect of school on learning outcomes, combined with the effect of interpolated questions, can be further researched.

Similarly, group also influenced learning outcomes. Theory on how exactly group influences learning outcomes is limited, but three suggestions are the learning climate, timing and group composition. A first explanation could be the classroom learning climate. For example, background noise, such as talking classmates, might prevent students from performing to their best ability (Lamotte et al., 2021; Massonnié et al., 2022) and annoys students (Lamotte et al., 2021). Another explanation that springs to mind is the timing of the experiment. Klein (2004) found that secondary school students achieve less in the early afternoon. He attributes this finding to the students' biological rhythm, which is in line with Hershner's (2020) finding. Lastly, group composition might have played a role as well. Burke and Sass (2013) found that peers have a small significant effect on learning outcomes. It was beyond the scope of the current study to determine how school and group influenced the relationship between timing of questions and learning outcomes.

Directions for Future Research

This study aimed to determine if interpolating questions impacts learning outcomes. Now that the results indicate that interpolating questions does indeed affect learning outcomes positively, it can be researched what mechanisms drive the effectiveness. For example, one could investigate how the questions could best be interpolated. The current study used forward fading, in which the student is guided through the problem chronologically (Renkl et al., 2004), but backward fading might lead to a different result. Other factors that need more investigation for their possible influence are prior knowledge and motivation. In this study, prior knowledge was only measured by self-report survey. A more objective way of measuring prior knowledge will provide a more accurate estimate of the influence of prior knowledge. A factor that was left out of this study, is motivation. Motivation in many forms is known to affect student performance (Howard et al., 2021; Richardson & Abraham, 2009). Interactive learning materials have shown to influence motivation to achieve (Li et al., 2018). Therefore it should be researched if motivation influences the correlation between interpolated questions and learning outcomes. Another interesting analysis that would shed light on the effect of interpolated questions on learning outcomes, is the comparison with regular classroom instruction.

Recommendations

The current study indicates that interpolating questions in instructional videos increases learning outcomes. Instructional videos remain a popular way of delivering instructions to students (de Koning et al., 2018), enhancing those will result in greater efficiency for both teacher and student. The teacher might have to spend more time creating the instruction, but it will help the student comprehend and apply the knowledge. That means that students will require less individual instruction during class time.

Another benefit of interpolating questions is that teachers can use the students' answers to identify struggling students and personalize instruction (Zou & Xie, 2019). Personalized instruction has proven to increase learning outcomes (Connor et al., 2018). This can help teachers focus their attention on those students and possibly decrease the performance gap between low-performing and high-performing students. As suggested before, interpolating questions might also help students regulate their own learning process. This suggests that even when a student cannot attend a lesson, the student can still receive an indication of their understanding of the subject matter, and the teacher can monitor the student's progress through the interpolated questions. Interpolating questions thus benefits both teachers and students.

References

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, *87*(3), 659–701. https://doi.org/10.3102/0034654316689306

Agarwal, P. K. (2019). Retrieval practice & Bloom's taxonomy: Do students need fact knowledge before higher order learning? *Journal of Educational Psychology*, *111*(2), 189–209. https://doi.org/10.1037/edu0000282

Alkema, E., Kuipers, J., Lindhout, C., & Tjerkstra, W. (2015). *Meer dan onderwijs: Theorie en praktijk van het onderwijs in de basisschool* (8th ed.). Van Gorcum.

Anderson, V., Northam, E., & Wrennall, J. (2019). *Developmental neuropsychology: A clinical approach* (Second edition). Routledge.

Atkinson, R. K., & Renkl, A. (2007). Interactive example-based learning environments: Using interactive elements to encourage effective processing of worked examples. *Educational Psychology Review*, *19*(3), 375–386. https://doi.org/10.1007/s10648-007-9055-2

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, *95*(4), 774–783. https://doi.org/10.1037/0022-0663.95.4.774

Baars, M., van Gog, T., de Bruin, A., & Paas, F. (2014). Effects of problem solving after worked example study on primary school children's monitoring accuracy. *Applied Cognitive Psychology*, *28*(3), 382–391. https://doi.org/10.1002/acp.3008

Baars, M., van Gog, T., de Bruin, A., & Paas, F. (2017). Effects of problem solving after worked example study on secondary school children's monitoring accuracy.

Educational Psychology, *37*(7), 810–834. https://doi.org/10.1080/01443410.2016.1150419

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*, 471–475. https://doi.org/10.1016/j.chb.2018.08.035

Biard, N., Cojean, S., & Jamet, E. (2018). Effects of segmentation and pacing on procedural learning by video. *Computers in Human Behavior*, *89*, 411–417. https://doi.org/10.1016/j.chb.2017.12.002

Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *92*(1), 81–90. https://doi.org/10.1177/003172171009200119

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (Eds.). (1956). *Cognitive domain*. Longman Green.

Boeije, H., Hart, H. ’t, & Hox, J. (Eds.). (2009). *Onderzoeksmethoden* (8th ed.). Boom Lemma.

Brame, C. J. (2016). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE—Life Sciences Education*, *15*(4), es6. https://doi.org/10.1187/cbe.16-03-0125

Brink, A. G. (2013). The impact of pre- and post-lecture quizzes on performance in intermediate accounting II. *Issues in Accounting Education*, *28*(3), 461–485. https://doi.org/10.2308/iace-50445

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*(7), 513–531.

Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics*, *31*(1), 51–82. https://doi.org/10.1086/666653

Cai, E. Y. L., & Liem, G. A. D. (2017). 'Why do I study and what do I want to achieve by studying?' Understanding the reasons and the aims of student engagement. *School Psychology International*, *38*(2), 131–148. https://doi.org/10.1177/0143034316686399

Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science*, *21*(5), 279–283. https://doi.org/10.1177/0963721412452728

Caspersen, J., Smeby, J.-C., & Olaf Aamodt, P. (2017). Measuring learning outcomes. *European Journal of Education*, *52*(1), 20–30. https://doi.org/10.1111/ejed.12205

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental Psychology*, *56*(4), 236–246. https://doi.org/10.1027/1618-3169.56.4.236

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*, 73–105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Chi, M. T. H., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., Levy, R., Li, N., McEldoon, K. L., Stump, G. S., Wylie, R., Xu, D., & Yaghmourian, D. L. (2018). Translating the ICAP theory of cognitive engagement into practice. *Cognitive Science*, *42*(6), 1777–1832. https://doi.org/10.1111/cogs.12626

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, *24*(2), 205–249. https://doi.org/10.1007/s10648-011-9191-6

Cohen, S. S., Madsen, J., Touchan, G., Robles, D., Lima, S. F. A., Henin, S., & Parra, L. C. (2018). Neural engagement with online educational videos predicts learning performance for individual students. *Neurobiology of Learning and Memory*, *155*, 60–64. https://doi.org/10.1016/j.nlm.2018.06.011

Connor, C. M., Mazzocco, M. M. M., Kurz, T., Crowe, E. C., Tighe, E. L., Wood, T. S., & Morrison, F. J. (2018). Using assessment to individualize early mathematics instruction. *Journal of School Psychology*, *66*, 97–113. https://doi.org/10.1016/j.jsp.2017.04.005

de Jong, T. (2019). Moving towards engaged learning in STEM domains: There is no simple answer, but clearly a road ahead. *Journal of Computer Assisted Learning*, *35*(2), 153–167. https://doi.org/10.1111/jcal.12337

de Koning, B. B., Hoogerheide, V., & Boucheix, J.-M. (2018). Developments and trends in learning with instructional video. *Computers in Human Behavior*, *89*, 395–398. https://doi.org/10.1016/j.chb.2018.08.055

Dignath, C., & Veenman, M. V. J. (2021). The role of direct strategy instruction and indirect activation of self-regulated learning—Evidence from classroom observation studies. *Educational Psychology Review*, *33*(2), 489–533. https://doi.org/10.1007/s10648-020-09534-0

Eiriksdottir, E., & Catrambone, R. (2011). Procedural instructions, principles, and examples: How to structure instructions for procedural tasks to enhance performance, learning, and transfer. *Human Factors*, *53*(6), 749–770. https://doi.org/10.1177/0018720811419154

Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, *326*(5952), 535–538. https://doi.org/10.1126/science.1168244

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*, 175–191.

Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). Sage.

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, *38*(4), 281–288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

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

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*, *109*(5), 653–665. https://doi.org/10.1037/edu0000161

Gerritsen-van Leeuwenkamp, K. J., Joosten-ten Brinke, D., & Kester, L. (2017). Assessment quality in tertiary education: An integrative literature review. *Studies in Educational Evaluation*, *55*, 94–116. https://doi.org/10.1016/j.stueduc.2017.08.001

Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, *106*(3), 608–626. https://doi.org/10.1037/a0034716

Görgüt, R. Ç., & Dede, Y. (2022). Mathematics teachers’ assessment of mathematical understanding. *Studies in Educational Evaluation*, *75*, 101190. https://doi.org/10.1016/j.stueduc.2022.101190

Gyllen, J. G., Stahovich, T. F., Mayer, R. E., Entezari, N., & Darvishzadeh, A. (2021). Priming productive study strategies with preparatory quizzes in an engineering course. *Applied Cognitive Psychology*, *35*(1), 169–180. https://doi.org/10.1002/acp.3750

Haagsman, M. E., Scager, K., Boonstra, J., & Koster, M. C. (2020). Pop-up questions within educational videos: Effects on students' learning. *Journal of Science Education and Technology*, *29*(6), 713–724. https://doi.org/10.1007/s10956-020-09847-3

Heritage, M., & Heritage, J. (2013). Teacher questioning: The epicenter of instruction and assessment. *Applied Measurement in Education*, *26*(3), 176–190. https://doi.org/10.1080/08957347.2013.793190

Hershner, S. (2020). Sleep and academic performance: Measuring the impact of sleep. *Current Opinion in Behavioral Sciences*, *33*, 51–56. https://doi.org/10.1016/j.cobeha.2019.11.009

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107. https://doi.org/10.1080/00461520701263368

Holleman, G. A., Hooge, I. T. C., Kemner, C., & Hessels, R. S. (2020). The 'real-world approach' and its problems: A critique of the term ecological validity. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00721

Hoogerheide, V., Staal, J., Schaap, L., & van Gog, T. (2019). Effects of study intention and generating multiple choice questions on expository text retention. *Learning and Instruction*, *60*, 191–198. https://doi.org/10.1016/j.learninstruc.2017.12.006

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., Visee, J., Lachner, A., & van Gog, T. (2019). Generating an instructional video as homework activity is both effective and enjoyable. *Learning and Instruction*, *64*, 101226. https://doi.org/10.1016/j.learninstruc.2019.101226

Howard, J. L., Bureau, J., Guay, F., Chong, J. X. Y., & Ryan, R. M. (2021). Student motivation and associated outcomes: A meta-analysis from self-determination theory. *Perspectives on Psychological Science*, *16*(6), 1300–1323. https://doi.org/10.1177/1745691620966789

Huang, K., Chen, C.-H., Wu, W.-S., & Chen, W.-Y. (2015). Interactivity of question prompts and feedback on secondary students' science knowledge acquisition and cognitive load. *Journal of Educational Technology & Society*, *18*(4), 159–171.

Huang, Y.-H., Lin, K.-C., Yu, X., & Hung, J. C. (2015). Comparison of different approaches on example-based learning for novice and proficient learners. *Human-Centric Computing and Information Sciences*, *5*(1), 29. https://doi.org/10.1186/s13673-015-0048-8

Ibrahim, M., Antonenko, P. D., Greenwood, C. M., & Wheeler, D. (2012). Effects of segmenting, signalling, and weeding on learning from educational video. *Learning, Media and Technology*, *37*(3), 220–235. https://doi.org/10.1080/17439884.2011.585993

Imlawi, J. (2021). Students' engagement in e-learning applications: The impact of sound's elements. *Education and Information Technologies*, *26*(5), 6227–6239. https://doi.org/10.1007/s10639-021-10605-0

Jansen, R. S., van Leeuwen, A., Janssen, J., Jak, S., & Kester, L. (2019). Self-regulated learning partially mediates the effect of self-regulated learning interventions on achievement in higher education: A meta-analysis. *Educational Research Review*, *28*, 100292. https://doi.org/10.1016/j.edurev.2019.100292

Kant, J. M., Scheiter, K., & Oschatz, K. (2017). How to sequence video modeling examples and inquiry tasks to foster scientific reasoning. *Learning and Instruction*, *52*, 46–58. https://doi.org/10.1016/j.learninstruc.2017.04.005

Karpicke, J. D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science*, *21*(3), 157–163. https://doi.org/10.1177/0963721412443552

Kihlstrom, J. F. (2021). Ecological validity and "ecological validity." *Perspectives on Psychological Science*, *16*(2), 466–471. https://doi.org/10.1177/1745691620966791

Kinnear, G., Bennett, M., Binnie, R., Bolt, R., & Zheng, Y. (2020). Reliable application of the MATH taxonomy sheds light on assessment practices. *Teaching Mathematics and Its Applications: An International Journal of the IMA*, *39*(4), 281–295. https://doi.org/10.1093/teamat/hrz017

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Klein, J. (2004). Planning middle school schedules for improved attention and achievement. *Scandinavian Journal of Educational Research*, *48*(4), 441–450. https://doi.org/10.1080/0031383042000245825

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lamotte, A.-S., Essadek, A., Shadili, G., Perez, J.-M., & Raft, J. (2021). The impact of classroom chatter noise on comprehension: A systematic review. *Perceptual and Motor Skills*, *128*(3), 1275–1291. https://doi.org/10.1177/00315125211005935

Lange, C., & Costley, J. (2019). How sequencing and fading affects the relationship between intrinsic and germane cognitive loads. *Distance Education*, *40*(2), 243–261. https://doi.org/10.1080/01587919.2019.1600366

Latimier, A., Peyre, H., & Ramus, F. (2021). A meta-analytic review of the benefit of spacing out retrieval practice episodes on retention. *Educational Psychology Review*, *33*(3), 959–987. https://doi.org/10.1007/s10648-020-09572-8

Li, S., Yamaguchi, S., & Takada, J. (2018). The influence of interactive learning materials on self-regulated learning and learning satisfaction of primary school teachers in Mongolia. *Sustainability*, *10*(4), 1093. https://doi.org/10.3390/su10041093

Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, *14*(3), 257–274. https://doi.org/10.1016/j.learninstruc.2004.06.003

Mandrekar, J. N. (2011). Measures of interrater agreement. *Journal of Thoracic Oncology*, *6*(1), 6–7. https://doi.org/10.1097/JTO.0b013e318200f983

Massonnié, J., Mareschal, D., & Kirkham, N. Z. (2022). Individual differences in dealing with classroom noise disturbances. *Mind, Brain, and Education*, *16*(3), 252–262. https://doi.org/10.1111/mbe.12322

Math with Menno (Director). (2020, May 13). *Wortels—Gelijksoortige wortels (2 HAVO/VWO & 2 VWO)*. https://www.youtube.com/watch?v=prRr9jyZUME

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 43–71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

Mayer, R. E. (2021). Evidence-based principles for how to design effective instructional videos. *Journal of Applied Research in Memory and Cognition*, *10*(2), 229–240. https://doi.org/10.1016/j.jarmac.2021.03.007

Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 316–344). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.016

McHugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica*, *23*(2), 143–149. https://doi.org/10.11613/BM.2013.018

Merkt, M., Ballmann, A., Felfeli, J., & Schwan, S. (2018). Pauses in educational videos: Testing the transience explanation against the structuring explanation. *Computers in Human Behavior*, *89*, 399–410. https://doi.org/10.1016/j.chb.2018.01.013

Besluit van 7 juni 2006, houdende vaststelling van de kerndoelen voor de onderbouw van het voortgezet onderwijs alsmede aanpassing van het Inrichtingsbesluit W.V.O. (Besluit kerndoelen onderbouw VO), (2006). https://open.overheid.nl/repository/ronl-6efa18ce27ee2b5f9738380fd1244dc267d6bafe/1/pdf/besluit-kerndoelen-onderbouw-vo.pdf

Mirriahi, N., Jovanović, J., Lim, L.-A., & Lodge, J. M. (2021). Two sides of the same coin: Video annotations and in-video questions for active learning. *Educational Technology Research and Development*, *69*(5), 2571–2588. https://doi.org/10.1007/s11423-021-10041-4

Mutlu-Bayraktar, D., Ozel, P., Altindis, F., & Yilmaz, B. (2022). Split-attention effects in multimedia learning environments: Eye-tracking and EEG analysis. *Multimedia Tools and Applications*, *81*(6), 8259–8282. https://doi.org/10.1007/s11042-022-12296-2

Nair, S., & Mathew, J. (2022). Learning through play: Gamification of learning, a systematic review of studies on gamified learning. *Journal of Information Technology Management*, *14*(1), 113–126. https://doi.org/10.22059/jitm.2021.322193.2779

Nichols, M. (2021). Development of an approved learning analytics ethics position. *Open Learning: The Journal of Open, Distance and e-Learning*, *0*(0), 1–14. https://doi.org/10.1080/02680513.2021.1986376

Norouzian, R., & Plonsky, L. (2018). Eta- and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. *Second Language Research*, *34*(2), 257–271. https://doi.org/10.1177/0267658316684904

OECD. (2019). *PISA 2018 results (Volume II): Where all students can succeed* (Vol. 2). OECD. https://doi.org/10.1787/b5fd1b8f-en

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8

Pellegrino, J. W., DiBello, L. V., & Goldman, S. R. (2016). A framework for conceptualizing and evaluating the validity of instructionally relevant assessments. *Educational Psychologist*, *51*(1), 59–81. https://doi.org/10.1080/00461520.2016.1145550

Pouw, W., Rop, G., de Koning, B. B., & Paas, F. (2019). The cognitive basis for the split-attention effect. *Journal of Experimental Psychology: General*, *148*(11), 2058–2075. https://doi.org/10.1037/XGE0000578

Pulukuri, S., & Abrams, B. (2021). Improving learning outcomes and metacognitive monitoring: Replacing traditional textbook readings with question-embedded videos. *Journal of Chemical Education*, *98*(7), 2156–2166. https://doi.org/10.1021/acs.jchemed.1c00237

Renkl, A., Atkinson, R. K., & Große, C. S. (2004). How fading worked solution steps works – A cognitive load perspective. *Instructional Science*, *32*(1), 59–82. https://doi.org/10.1023/B:TRUC.0000021815.74806.f6

Renkl, A., Atkinson, R. K., Maier, U. H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *The Journal of Experimental Education*, *70*(4), 293–315. https://doi.org/10.1080/00220970209599510

Rey, G. D., & Buchwald, F. (2011). The expertise reversal effect: Cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied*, *17*(1), 33–48. https://doi.org/10.1037/a0022243

Rice, P., Beeson, P., & Blackmore-Wright, J. (2019). Evaluating the impact of a quiz question within an educational video. *TechTrends*, *63*(5), 522–532. https://doi.org/10.1007/s11528-019-00374-6

Richardson, M., & Abraham, C. (2009). Conscientiousness and achievement motivation predict performance. *European Journal of Personality*, *23*(7), 589–605. https://doi.org/10.1002/per.732

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Scherer, R., Greiff, S., & Hautamäki, J. (2015). Exploring the relation between time on task and ability in complex problem solving. *Intelligence*, *48*, 37–50. https://doi.org/10.1016/j.intell.2014.10.003

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy*, *2*(4), 419–436. https://doi.org/10.1207/S15327078IN0204_02

Schroeder, N. L., & Cenkci, A. T. (2018). Spatial contiguity and spatial split-attention effects in multimedia learning environments: A meta-analysis. *Educational Psychology Review*, *30*(3), 679–701. https://doi.org/10.1007/s10648-018-9435-9

Schroeder, N. L., & Cenkci, A. T. (2020). Do measures of cognitive load explain the spatial split-attention principle in multimedia learning environments? A systematic review. *Journal of Educational Psychology*, *112*(2), 254–270. https://doi.org/10.1037/edu0000372

Smith, G., Wood, L., Coupland, M., Stephenson, B., Crawford, K., & Ball, G. (1996). Constructing mathematical examinations to assess a range of knowledge and skills. *International Journal of Mathematics Engineering Science and Technology*, *27*(1), 65–77. https://doi.org/10.1007/0-306-47231-7_9

Sweller, J. (2011). Cognitive Load Theory. In *Psychology of Learning and Motivation* (Vol. 55, pp. 37–76). Elsevier. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Sweller, J., Ayres, P., & Kalyuga, S. (2011). The Split-Attention Effect. In J. Sweller, P. Ayres, & S. Kalyuga (Eds.), *Cognitive Load Theory* (pp. 111–128). Springer. https://doi.org/10.1007/978-1-4419-8126-4_9

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences*, *110*(16), 6313–6317. https://doi.org/10.1073/pnas.1221764110

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, *83*(3), 357–385. https://doi.org/10.3102/0034654313483907

Tindall-Ford, S., Agostinho, S., Bokosmaty, S., Paas, F., & Chandler, P. (2015). Computer-based learning of geometry from integrated and split-attention worked examples: The power of self-management. *Journal of Educational Technology & Society*, *18*(4), 89–99.

Troncoso Skidmore, S., & Thompson, B. (2013). Bias and precision of some classical ANOVA effect sizes when assumptions are violated. *Behavior Research Methods*, *45*(2), 536–546. https://doi.org/10.3758/s13428-012-0257-2

van Alten, D. C. D., Phielix, C., Janssen, J., & Kester, L. (2020). Self-regulated learning support in flipped learning videos enhances learning outcomes. *Computers & Education*, *158*, 104000. https://doi.org/10.1016/j.compedu.2020.104000

van Gog, T., Hoogerheide, V., & van Harsel, M. (2020). The role of mental effort in fostering self-regulated learning with problem-solving tasks. *Educational Psychology Review*, *32*(4), 1055–1072. https://doi.org/10.1007/s10648-020-09544-y

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*(2), 265–289. https://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*(3), 212–218. https://doi.org/10.1016/j.cedpsych.2010.10.004

van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review*, *22*(2), 155–174. https://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. https://doi.org/10.1016/j.compedu.2013.12.004

van Harsel, M., Hoogerheide, V., Verkoeijen, P., & van Gog, T. (2019). Effects of different sequences of examples and problems on motivation and learning. *Contemporary Educational Psychology*, *58*, 260–275. https://doi.org/10.1016/j.cedpsych.2019.03.005

van Harsel, M., Hoogerheide, V., Verkoeijen, P., & van Gog, T. (2020). Examples, practice problems, or both? Effects on motivation and learning in shorter and longer sequences. *Applied Cognitive Psychology*, *34*(4), 793–812. https://doi.org/10.1002/acp.3649

van Wermeskerken, M., Ravensbergen, S., & van Gog, T. (2018). Effects of instructor presence in video modeling examples on attention and learning. *Computers in Human Behavior*, *89*, 430–438. https://doi.org/10.1016/j.chb.2017.11.038

VO Raad. (2020, December 18). *60.000 extra laptops en tablets voor onderwijs op afstand—VO-raad*. https://www.vo-raad.nl/nieuws/60-000-extra-laptops-en-tablets-voor-onderwijs-op-afstand

Vörös, Z., Kehl, D., & Rouet, J.-F. (2021). Task characteristics as source of difficulty and moderators of the effect of time-on-task in digital problem-solving. *Journal of Educational Computing Research*, *58*(8), 1494–1514. https://doi.org/10.1177/0735633120945930

Wang, M.-T., & Degol, J. L. (2016). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review*, *28*(2), 315–352. https://doi.org/10.1007/s10648-015-9319-1

Weurlander, M., Söderberg, M., Scheja, M., Hult, H., & Wernerson, A. (2012). Exploring formative assessment as a tool for learning: Students' experiences of different methods of formative assessment. *Assessment & Evaluation in Higher Education*, *37*(6), 747–760. https://doi.org/10.1080/02602938.2011.572153

Whitton, N., & Moseley, A. (2014). Deconstructing engagement: Rethinking involvement in learning. *Simulation & Gaming*, *45*(4–5), 433–449. https://doi.org/10.1177/1046878114554755

Zou, D., & Xie, H. (2019). Flipping an English writing class with technology-enhanced just-in-time teaching and peer instruction. *Interactive Learning Environments*, *27*(8), 1127–1142. https://doi.org/10.1080/10494820.2018.1495654

Appendix A

Letter of Informed Consent

Proefpersoneninformatie voor deelname aan (sociaal)-wetenschappelijk onderzoek

'The effect of interpolated questions in video modeling examples on learning outcomes'

Zeist, 10 maart 2023

Geachte heer, mevrouw,

Middels deze brief wil ik u toestemming vragen om uw kind mee te laten doen aan het onderzoek "The effect of interpolated questions in video modeling examples on learning outcomes" (NL: Het effect van vragen stellen in instructievideo's op leerprestaties). Dit onderzoek heeft als doel om erachter te komen of vragen stellen tijdens een instructievideo invloed heeft op leerprestaties, en wanneer die vragen gesteld moeten worden. Uiteindelijk draagt dit onderzoek daarmee bij aan het ontwikkelen van goed digitaal onderwijsmateriaal.

Wat wordt van uw kind als participant verwacht

Nadat u en uw kind toestemming voor deelname hebben verleend, beantwoordt de leerling digitaal enkele vragen over het onderwerp van de video. Deze vragen hebben als doel om vast te stellen hoeveel voorkennis de leerling heeft. Vervolgens krijgt de leerling een korte video (minder dan 10 minuten) te zien over een wiskundig onderwerp. Dit onderwerp is zo gekozen dat de gemiddelde leerling het onderwerp kan begrijpen. Mogelijk moet de leerling tijdens of na de video vragen beantwoorden over het onderwerp van de video. Daarna maakt de leerling twee opgaven het onderwerp van de video. In totaal duurt deelname aan het experiment ongeveer 30 minuten.

Deelname is eenmalig en er zijn geen consequenties verbonden aan deelname. Deelname vindt plaats onder schooltijd en vraagt geen andere voorbereiding dan het verlenen van toestemming. Nadat het experiment volledig is afgerond, zal de uitwerking van de twee opgaven met de leerling gedeeld worden.

Er zijn geen voor- en nadelen verbonden aan dit onderzoek. Er is ook geen beloning of vergoeding voor deelname.

Vertrouwelijkheid verwerking gegevens

Voor dit onderzoek is het nodig dat ik een aantal persoonsgegevens van uw kind verzamel. Deze gegevens hebben ik nodig om de onderzoeksvraag goed te kunnen beantwoorden. De persoonsgegevens worden op een andere computer opgeslagen dan de onderzoeksgegevens zelf (de zgn. ruwe data). De computer waarop de persoonsgegevens worden opgeslagen is volgens de hoogste normen beveiligd en alleen betrokken onderzoekers hebben toegang tot deze gegevens. De gegevens zelf zijn ook beveiligd d.m.v. een beveiligingscode. Bovendien worden de gegevens geanonimiseerd, door het toewijzen van een willekeurig gegenereerde code. Persoonsgegevens zullen verzameld worden met het gebruik van het online vragenlijstprogramma Qualtrics.

Voor deelname aan het onderzoek is het school-e-mailadres van uw kind nodig. Het e-mailadres zal eenmalig gebruikt worden om uw kind de link te verstrekken om deel te kunnen nemen. Deelname aan het onderzoek is anoniem, het is dus niet te herleiden welke gegevens van uw kind zijn. De contactgegevens zullen na deelname aan het onderzoek verwijderd worden.

De geanonimiseerde gegevens van uw kind zullen voor minimaal 10 jaar bewaard worden. Dit is volgens de daartoe bestemde richtlijnen van de VSNU. Meer informatie over privacy kunt u lezen op de website van de Autoriteit Persoonsgegevens:

https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/avg-europese-privacywetgeving

Vrijwilligheid deelname

Deelname aan dit onderzoek is vrijwillig. Uw kind kan op elk gewenst moment stoppen met het onderzoek, zonder opgave van reden en zonder voor u of uw kind nadelige gevolgen. De tot dan toe verzamelde gegevens worden wel gebruikt voor het onderzoek, tenzij u expliciet aangeeft dit niet te willen.

Uw rechten

- u heeft het recht een kopie op te vragen van alle persoonsgegevens gebruikt zijn in het onderzoek
- u heeft het recht om te verzoeken uw persoonsgegevens te corrigeren bij onjuistheid
- u heeft het recht om bezwaar te maken tegen de verwerking van de persoonsgegevens en recht op gegevensoverdraagbaarheid
- u heeft het recht op verwijderen van persoonsgegevens (dit kan niet in alle situaties)
- u heeft het recht om toestemming voor het verwerken van de persoonsgegevens in te trekken. Analyses die tot dat moment zijn gemaakt met de betreffende persoonsgegevens zullen wel gebruikt blijven worden voor het onderzoek.

U kunt uw privacyrechten uitoefenen door contact op te nemen met de onderzoekers of via privacy@uu.nl

Onafhankelijk contactpersoon en klachtenfunctionaris

Als u vragen of opmerkingen over het onderzoek heeft, kunt u contact opnemen met edu.acma.thesis@uu.nl

Als u een officiële klacht heeft over het onderzoek, dan kunt u een mail sturen naar de klachtenfunctionaris via klachtenfunctionaris-fetcsocwet@uu.nl

Contactgegevens Functionaris Gegevensbescherming:

https://www.uu.nl/organisatie/praktische-zaken/privacy/functionaris-voor-gegevensbescherming of privacy@uu.nl

Als u na het lezen van deze informatiebrief besluit tot deelname aan het onderzoek verzoek ik u, indien mogelijk beide ouders/verzorgers, toestemming te verlenen door onderstaand hokje aan te kruisen en het school-e-mailadres van uw kind te verstrekken.

Vriendelijke groet,

Lisette Bosveld

t.e.bosveld@students.uu.nl

Toestemmingsverklaring:

Door onderstaand vakje aan te kruisen verklaar ik de informatiebrief m.b.t. onderzoek "The effect of interpolated questions in video modeling examples on learning outcomes" gelezen te hebben en akkoord te gaan met deelname aan het onderzoek van mijn kind.

Dit betekent dat ik toestemming voor mijn kind geef om deel te nemen aan het onderzoek.

Het school-e-mailadres van mijn kind is:

Appendix B

Transcripts Instructional Video (Dutch and English)

Nederlands

Hey wat leuk dat je kijkt! Welkom bij weer een nieuwe uitleg video van Math with Menno. Laten we beginnen!

Deze video gaat over gelijksoortige wortels. In deze video ga ik je laten zien hoe je wortels bij elkaar kunt optellen en van elkaar kunt aftrekken. Ik ga het uitleggen aan de hand van een aantal voorbeelden, maar voordat we naar de voorbeelden gaan kijken, kijk je eerst even naar een stukje theorie.

We gaan het dus hebben over optellen en aftrekken met wortels en om te begrijpen hoe dat werkt, kijken we eerst even naar iets wat we wel vaker hebben gezien en dat is dit sommetje: als je 3a plus 5a doet, dan is de uitkomst daarvan 8a. Je mag dit bij elkaar optellen, want ze hebben allerlei de a. Je doet 3 plus 5 dat is 8 en die a die blijft hetzelfde. Nou, dat principe van het optellen van letters gebruiken we ook bij het optellen van wortels. Bijvoorbeeld $3\sqrt{2} + 5\sqrt{2}$. Dan ga je dus eerst kijken: Zijn de wortels hetzelfde? Allebei wortel 2, dus dat zit wel goed. Dan doen we 3 + 5 dat wordt 8 en die wortel die blijft, net als die letter, gewoon hetzelfde, dus die blijft de wortel van 2. Het klinkt ook wel logisch, hè, want als je zegt ik heb 3 keer de wortel van 2, en dan komen er 5 bij. Dan heb je dus in totaal acht keer de wortel van 2, dus dat klopt helemaal. Maar belangrijk is dus wel dat die wortels precies hetzelfde moeten zijn, want als de wortels anders zijn, dus stel hier staat de wortel van 7, dan zijn de wortels niet hetzelfde en dan is het antwoord kan niet.

Soms heb je een situatie waarbij er voor een wortel niks staat, zoals hier. Dan heb je $\sqrt{7} + \sqrt{7}$. En als er niks staat, staat er eigenlijk een één. Dus er staat eigenlijk één wortel 7 plus één wortel zeven. En dan doen we net als hier 1 + 1 = 2. Die wortel 7 die blijft hetzelfde, dus het antwoord is dan 2 wortel 7.

Dus samenvattend kunnen we zeggen dat eigenlijk hetzelfde werkt als met letters. De wortels moeten dus precies hetzelfde zijn, anders mag je niet optellen en aftrekken. De getallen voor de wortels, die doe je bij elkaar. De wortel blijft hetzelfde en dan heb je het antwoord gevonden.

Nu gaan we dat even toepassen bij wat voorbeelden. De opdracht is: herleid zo mogelijk zet anders kan niet. We beginnen bij A. Er staat $2\sqrt{5} + 7\sqrt{5}$. Het is allebei wortel 5 dus we kunnen dit wel herleiden. En wat doe je dan? Nou, je doet 2 plus 7 dat is gelijk aan 9. Dus je krijgt = 9 en de wortel van 5? Die blijft gewoon staan, want die blijft de hele tijd hetzelfde, net als hier.

Dan gaan we naar B. $10\sqrt{11} - 6\sqrt{11}$ Ook nu kijken we weer: Zijn de wortels hetzelfde? Dat is het geval, want het is allebei de wortel van 11. Dan doen we 10 min 6 en 10 min 6 is 4, dus het antwoord is dan vier en die wortel van 11 zet je erachter, dus vier wortel 11 en nu hebben we vraag B al opgelost.

Dan gaan we naar C. Daar staat $7\sqrt{3} - 11\sqrt{3}$. Eerst check je weer even: Zijn de wortels hetzelfde? Het is allebei wortel drie, dus die zijn hetzelfde. Dan doen we 7 min 11. Nou, 7 min 11 dat is gelijk aan min 4. Nu ga je die wortel 3 er achter zetten, dus je krijgt min vier wortel 3. Je ziet dus dat je in zo'n wortel ook wel een negatief getal ervoor mag hebben. Min vier wortel 3 mag wel, maar zoals we eerder hebben besproken, je mag niet de wortel nemen van een negatief getal. Dus onder de wortel mag een negatief getal staan dus dit mag niet de wortel van min 3 zijn. Maar ervoor? dat kan dus wel.

Dan gaan we naar D. Daar staat $2\sqrt{3} + 3\sqrt{2}$. Eerst checken we weer even: Zijn de wortels hetzelfde? Maar je ziet dit is wortel 3 en dit is wortel 2. Ze zijn dus niet hetzelfde, dus we mogen deze twee dingen niet bij elkaar optellen. Dus het antwoord hier is ‘kan niet’ en dat mag je afkorten met k.n. Want als de wortels niet hetzelfde zijn, kun je het niet bij elkaar optellen, net als met letters. Dus nu is het antwoord kan niet.

Dan gaan we naar E. Daar staat $3\sqrt{5} - \sqrt{5}$. De wortels en hetzelfde, dus we kunnen min doen. Hier moet je bedenken er staat eigenlijk een één. Hier moet je bedenken, staat eigenlijk een één. Denk aan wat we hier hebben besproken, als er niks staat, staat eigenlijk een één, dus eigenlijk is dat 1 wortel 5. Nu doen we 3 min 1 dat is 2. De wortel van 5 schrijf je gewoon op, dus het antwoord wordt dan 2 wortel vijf. Zo kun je dus optellen en aftrekken met wortels.

Dus, wat hebben we gezien? Optellen en aftrekken met wortels mag alleen als de wortel precies hetzelfde is, anders is het antwoord ‘kan niet’, zoals bij D. Wat doe je? Het werkt hetzelfde als met letters. De getallen die ervoor staan, tel je bij elkaar op of trek je van elkaar af, en de wortel blijft de hele tijd hetzelfde. Op deze manier kun je zo'n vraag dan netjes uitwerken.

Ben je blij met mijn video’s? Abonneer dan op mijn kanaal. Wil je nog meer video's over dit hoofdstuk zien? Klik dan hiernaast. Tot de volgende keer!

English

Hey! Thanks for watching and welcome to a new video by Math with Menno. Let’s begin!

This video is about radicals, I’ll show you how you can add and subtract radicals. I’ll explain it using a few examples, but before we look at those, I’ll explain a bit of theory.

We are going to talk about simplifying radical expressions. To understand how that works, we’ll look at something we’ve seen before. That’s this equation: 3a+5a=8a. You can add this, because they have the same variable, the letter A. We add 3 to 5, which is 8, the a stays the same. We’ll use this principle of adding letters when we add radicals. For example, $3\sqrt{2} + 5\sqrt{2}$. First you ask: Are the radicals the same? Both have the radicand 2, so that’s good. We add 3 to 5, which makes 8. The radical remains the same, root 2. It makes sense, because I have three times the square root of 2, and then I add five more. That makes eight times the square root of 2. Important to note is that the radical should be the same. If the radicand differs, so if this radicand would be 7, you can't combine the radicals.

Sometimes, the radical does not have a coefficient, such as here. It says $\sqrt{7} + \sqrt{7}$. When there is no coefficient, the coefficient is one. So actually, it says one root 7, plus one root 7. Then we simply add the coefficients, just like here. The radical remains the same, so the answer is two times the square root of seven.

In short, adding radicals is very similar to adding variables. When adding or subtracting radicals, the radical should be exactly the same, otherwise you can't add or subtract. You add, or subtract, the coefficients before the radical, the radical remains the same. Then you've simplified a radical expression.

Now, let's apply this to some examples. The assignment is: simplify if possible, write 'simplified' if the expression can't be simplified.

We'll start with A. It says $2\sqrt{5} + 7\sqrt{5}$. Both have radicand 5, thus the radicals are the same, which means we can simplify this expression. So what do we do? First, you add two to seven, which equals nine. What about root 5? We write that down, because the radical does not change, just like a variable.

Let's do B. $10\sqrt{11} - 6\sqrt{11}$. First I check: Are the radicals the same? That is the case, both have radicand 11. Then we subtract 6 from 10, which equals 4. I write down '= 4' and simply write down root 11 behind it, because the radical does not change. So the simplified expression is 4 root 11.

Next up is C. It says $7\sqrt{3} - 11\sqrt{3}$. First I check: Are the radicals the same? Both have radicand 3, thus they are like radicals, so we can simplify the expression. First, I subtract 7 from 11. 11 minus 7 is negative 4. Then I write root 3 behind it, so the answer is $-4\sqrt{3}$. As you see, a radical can have a negative in front of it. Negative four root 3 is allowed, but it is impossible to have the square root of a negative. So you can’t have a negative radicand, root negative 3 is incorrect, but negative root three is possible.

We’ll look at D now. I have to simplify $2\sqrt{3}+3\sqrt{2}$. First I check if the radicals are the same. As you can see, this is root 3, and this is root 2. They do not have the same radicand, so I can’t add these radicals. The answer is ‘simplified’, because the expressions are as simplified as they can be. You can only add like radicals, radicals with the same radicand.

The last one: E. It says: $3\sqrt{5}-\sqrt{5}$. The radicals are the same, so I can subtract them. Remember, when there was no coefficient, the coefficient is one. So we have to subtract one root 5 from three root five. 3 minus 1 equals 2. I just had to write down the radical, because that did not change. The simplified version of the expression is 2 root five.

And that is how you add and subtract radicals.

So, what have we learned? You can only add or subtract radicals when they have the same radicand. If an expression can’t be simplified further, you write ‘simplified’, as we did with example D. So what do you do? It is similar to addition and subtraction of variables. You add or subtract the coefficients, the numbers before the radical, and the radical remains the same. And that’s how you add and subtract radical expressions.

Do you like my videos? Please subscribe. Would you like to see more videos about this unit? Click on the link.

See you next time!

Appendix C

Questions Instructional Video Group A and D (Dutch and English)

Nederlands

1. **Mag je $2\sqrt{5} + 7\sqrt{5}$ korter schrijven?**
   A. Nee, want het getal onder de wortel is niet hetzelfde.
   B. Ja, want het getal onder de wortel is hetzelfde.
   C. Nee, want het getal voor de wortel is niet hetzelfde.
   D. Ja, want het getal voor de wortel is hetzelfde.

2. **Mag je $10\sqrt{11} - 6\sqrt{11}$ herleiden?**
   A. Nee, want het getal onder de wortel is niet hetzelfde.
   B. Nee, want het getal voor de wortel is niet hetzelfde.
   C. **Ja, want het getal onder de wortel is hetzelfde.**
   D. Ja, want het getal voor de wortel is hetzelfde.
3. **Je weet nu dat je $10\sqrt{11} - 6\sqrt{11}$ mag herleiden. Wat is de volgende stap?**
   A. $\sqrt{11} + \sqrt{11} = 2\sqrt{11}$
   B. 10 – 6 = 4
   C. $\sqrt{11} - \sqrt{11} = 0$

4. **Mag je $7\sqrt{3} - 11\sqrt{3}$ herleiden?**
   A. Nee, want het getal onder de wortel is niet hetzelfde.
   B. Nee, want het getal voor de wortel is niet hetzelfde.
   C. Ja, want het getal voor de wortel is hetzelfde.
   D. **Ja, want het getal onder de wortel is hetzelfde.**
5. **Je weet nu dat je $7\sqrt{3} - 11\sqrt{3}$ mag herleiden. Wat is de volgende stap?**
   A. $\sqrt{3} + \sqrt{3} = 2\sqrt{3}$
   B. 7 – 11 = - 4
   C. $\sqrt{3} - \sqrt{3} = 0$
6. **Herleid $7\sqrt{3} - 11\sqrt{3}$**
   $7\sqrt{3} - 11\sqrt{3} = \mathbf{-4\sqrt{3}}$

7. **Mag je $2\sqrt{3} + 3\sqrt{2}$ herleiden?**
   A. Ja
   B. Nee
8. **Waarom mag je $2\sqrt{3} + 3\sqrt{2}$ niet herleiden?**
   Omdat het getal onder de wortel niet gelijk is.

9. **Herleid zo mogelijk: $3\sqrt{5} - \sqrt{5}$**
   Vink alles aan wat waar is.

A. Je kunt dit niet korter schrijven.
B. Voor de tweede $\sqrt{5}$ staat eigenlijk een 1.
C. Je mag de getallen voor de wortels van elkaar aftrekken
D. Je mag de getallen onder de wortels bij elkaar optellen.
E. De wortels zijn van dezelfde soort, namelijk $\sqrt{5}$

10. Wat is de volgende stap bij het herleiden van $3\sqrt{5} - 1\sqrt{5}$?
A. 3 – 0
B. 3 + 1
C. 3 + 5 – 1 + 5
D. 3 – 1

11. Herleid $3\sqrt{5} - 1\sqrt{5}$
$3\sqrt{5} - 1\sqrt{5} = \mathbf{2\sqrt{5}}$

English

1. Is it possible to simplify $2\sqrt{5} + 7\sqrt{5}$?
A. No, because the radicands differ.
B. Yes, because the radicands are the same.
C. No, because the coefficients are the same.
D. Yes, because the coefficients are the same.

2. Is it possible to simplify $10\sqrt{11} - 6\sqrt{11}$?
A. No, because the radicands differ.
B. No, because the coefficients differ.
C. **Yes, because the radicands are the same.**
D. Yes, because the coefficients are the same.

3. It is possible to simplify $10\sqrt{11} - 6\sqrt{11}$. What is the next step?
A. $\sqrt{11} + \sqrt{11} = 2\sqrt{11}$
B. 10 – 6 = 4
C. $\sqrt{11} - \sqrt{11} = 0$

4. Is it possible to simplify $7\sqrt{3} - 11\sqrt{3}$?
A. No, because the radicands differ.
B. No, because the coefficients differ.
C. Yes, because the coefficients are the same.
D. **Yes, because the radicands are the same.**

5. It is possible to simplify $7\sqrt{3} - 11\sqrt{3}$ What is the next step?
A. $\sqrt{3} + \sqrt{3} = 2\sqrt{3}$
B. 7 – 11 = - 4
C. $\sqrt{3} - \sqrt{3} = 0$

6. Simplify $7\sqrt{3} - 11\sqrt{3}$
$7\sqrt{3} - 11\sqrt{3} = \mathbf{-4\sqrt{3}}$

7. **Is it possible to simplify $2\sqrt{3} + 3\sqrt{2}$?**
   A. Yes
   B. No
8. **Why is it not possible to simplify $2\sqrt{3} + 3\sqrt{2}$?**
   Because the radicands differ.

9. **Simplify if possible: $3\sqrt{5} - \sqrt{5}$**

   Check the boxes that apply.

   A. It cannot be simplified.
   B. The coefficient of the second $\sqrt{5}$ is 1.
   C. You have to subtract the coefficients.
   D. You have to add the radicands.
   E. The radicals both have radicand 5.
10. **What is the next step in simplifying $3\sqrt{5} - 1\sqrt{5}$?**
   A. 3 – 0
   B. 3 + 1
   C. 3 + 5 – 1 + 5
   D. 3 – 1
11. **Simplify $3\sqrt{5} - 1\sqrt{5}$**
   $3\sqrt{5} - 1\sqrt{5} = \mathbf{2\sqrt{5}}$

Appendix D

Translated Posttest

Participant ID: … … … … …

Posttest

Write all your calculations and answers on this sheet.

Simplifying radicals

Simplify if possible, write '*simplified*' if the expression can't be simplified.

1 $\sqrt{43} + 27\sqrt{43} =$

2 $4\sqrt{7} - 4\sqrt{8} =$

3 $5\sqrt{19} - 7\sqrt{19} =$

4 $8\sqrt{10} - 3\sqrt{10} =$

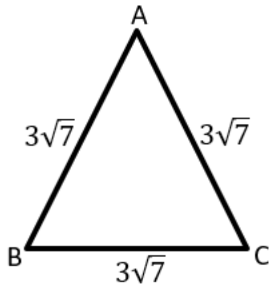
5 $7\sqrt{2} + 5\sqrt{2} =$

6 $16\sqrt{3} - 4\sqrt{3} - 2\sqrt{3} =$

7 $16\sqrt{5} - 7\sqrt{5} + 3\sqrt{13} =$

Perimeter of a triangle

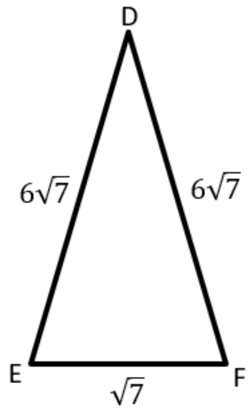
Here you see triangle ABC. Each of the sides of triangle ABC has a length of $3\sqrt{7}$.



8 Find the perimeter of triangle ABC.

9 Here you see triangle DEF. Two sides are $6\sqrt{7}$ long, and one side is $\sqrt{7}$ long.

Find the difference in perimeter between triangle DEF and triangle ABC.



Appendix E

Marking Sheet Posttest

Q	answer key	score

Simplifying radicals

1 **maximum score 1**
- ○ $\sqrt{43} + 27\sqrt{43}$
- • $= 28\sqrt{43}$ — 1

2 **maximum score 1**
- ○ $4\sqrt{7} - 4\sqrt{8}$
- • $= k.n.$ — 1

3 **maximum score 1**
- ○ $5\sqrt{19} - 7\sqrt{19}$
- • $= -2\sqrt{19}$ — 1

4 **maximum score 1**
- ○ $8\sqrt{10} - 3\sqrt{10}$
- • $= 5\sqrt{10}$ — 1

5 **maximum score 1**
- ○ $7\sqrt{2} + 5\sqrt{2}$
- • $= 12\sqrt{2}$ — 1

6 **maximum score 2**
- ○ $16\sqrt{3} - 4\sqrt{3} - 2\sqrt{3}$
- • $= 10\sqrt{3}$ — 1

7 **maximum score 2**
- ○ $16\sqrt{5} - 7\sqrt{5} + 3\sqrt{13}$
- • $= 9\sqrt{5} + 3\sqrt{13}$ — 2

One point for subtracting $16\sqrt{5} - 7\sqrt{5}$ and
one for not adding $3\sqrt{13}$ to $16\sqrt{5} - 7\sqrt{5} = 9\sqrt{5}$
No marks if 'simplified'

Perimeter of a triangle

8 **maximum score 2**
- • $3\sqrt{7} + 3\sqrt{7} + 3\sqrt{7}$ — 1
- • $= 9\sqrt{7}$ — 1

OR

- • $3\sqrt{7} \times 3$ — 1
- • $= 9\sqrt{7}$ — 1

9 **maximum score 4**
- • $6\sqrt{7} + 6\sqrt{7} + \sqrt{7}$ — 1
- • $= 13\sqrt{7}$ — 1
- • $13\sqrt{7} - 9\sqrt{7}$ — 1
- • $= 4\sqrt{7}$ — 1

Deduct one point if $9\sqrt{7} - 13\sqrt{7} = -4\sqrt{7}$

OR

- • $6\sqrt{7} + 6\sqrt{7} + \sqrt{7} - 9\sqrt{7}$ — 2
- • $= 4\sqrt{7}$ — 2

Maximum score: 14

Appendix F

Additional Tables

Table F1

Overview of Posttest Scores per School

	N	M	SD	$M_{Comprehension}$	$SD_{Comprehension}$	$M_{Application}$	$SD_{Application}$
School 1	152	9.36	2.648	6.00	1.371	3.36	1.910
School 2	64	6.55	4.004	4.59	2.473	1.95	2.027
School 3	21	5.57	2.942	4.24	1.998	1.33	1.742
Total	236	8.27	3.443	5.47	1.927	2.81	2.072

Table F2

Overview of Posttest Scores per Group

	N	M	SD	$M_{Comprehension}$	$SD_{Comprehension}$	$M_{Application}$	$SD_{Application}$
Group 1	30	9.43	2.012	6.00	0.830	3.43	1.775
Group 2	12	9.50	1.977	6.00	0.854	3.50	1.624
Group 3	13	8.54	2.876	5.77	1.481	2.77	2.204
Group 4	19	4.47	3.672	3.47	2.590	1.00	1.563
Group 5	18	6.94	4.425	4.50	2.572	2.44	2.148
Group 6	17	7.82	3.468	5.41	1.906	2.41	2.123
Group 7	24	9.50	2.167	5.92	1.283	3.58	1.840
Group 8	23	9.00	2.939	6.04	1.331	2.96	2.142
Group 9	23	10.22	3.477	6.52	1.974	3.70	2.141
Group 10	23	9.70	1.893	5.96	1.022	3.74	1.544
Group 11	14	7.00	3.883	5.14	2.476	1.86	1.916
Group 12	21	5.57	2.942	4.24	1.998	1.33	1.742
Total	236	8.27	3.443	5.47	1.927	2.81	2.072