Observational Learning from Animated Models: Effects of Modality and Reflection on Transfer

Pieter Wouters*¹, Fred Paas²,³, and Jeroen J. G. van Merriënboer²

¹Institute of Information and Computing Sciences
Utrecht University, the Netherlands

²Educational Technology Expertise Centre,
Open University of the Netherlands, Heerlen, the Netherlands

³Institute of Psychology,
Erasmus University Rotterdam, the Netherlands

Key words: cognitive load theory, modality, cognitive modeling, reflection

* Correspondence concerning this article should be addressed to Pieter Wouters, Utrecht University, Institute of Information and Computing Sciences, P.O. Box 80.089, 3508 TB Utrecht, the Netherlands. E-mail: pj.m.wouters@cs.uu.nl, Tel. +31 30 253 9219

Author Note

This research project is funded by the Netherlands Organization for Scientific Research (NWO, The Hague, project number: 411-02-163)
Abstract

Animated models use animations and explanations to teach how a problem is solved and why particular problem-solving methods are chosen. Often spoken explanations are proposed to accompany animations in order to prevent overloading the visual channel (i.e., the modality effect). In this study we adopt the hypothesis that the inferior performance of written text compared to spoken text is due to the fact that written text receives less attention and, consequently, less effortful processing. In a 2 x 2 factorial experiment ($N = 96$) with the factors modality (written, spoken) and reflection prompts (yes, no) the hypothesis is tested that prompted reflection requires learners to explicitly attend to written explanations and carefully process them, thus yielding higher transfer performance, whereas for spoken explanations prompted reflection would have no effect on transfer performance. The results indeed showed the hypothesized interaction between modality and reflection prompts. They suggest that the modality effect can be compensated for when learners explicitly attend to the information and effortfully process it. This has implications for learning situations in which spoken explanations are no option, such as education for the hearing-impaired.
We refer to the combined use of animations with explanatory text and pedagogical agents in modeling as ‘animated models’. These animated models illustrate the solving of, for instance, scientific problems (e.g., solving a problem about gravity), mathematical problems (e.g., probability calculation problems), and search problems (e.g., finding information on the Internet). The pedagogical agent functions as a social model and guides the learner through the animation, by moving around the screen and guiding the learner’s attention to specific parts of the animation, by addressing the learner in a personalized style and/or by showing which errors typically occur and how they may be avoided by the learner. For example, in solving a problem in the domain of probability calculation, it is important to know whether it is a ‘drawing with or without replacement’. For novices this concept may be rather abstract and difficult to understand. An animation can visualize the concept by showing what is happening, for instance, in a situation with mobile phones. Imagine a mobile factory where in an assembly line six mobiles -each with a distinct color- are packed in a box. A controller blindly selects two mobiles to check them for deficiencies. The learner has to calculate the probability that the controller draws a yellow mobile and a blue mobile from the box. The animated model may show a box with six mobiles. The first mobile drawn from the box can be put away from the box. As shown in Figure 1, the pedagogical agent may move to the mobile drawn and explain that a mobile that is drawn should not be put back because you do not want to draw an already checked mobile again. Then the group of remaining mobiles in the box becomes encircled. The pedagogical agent moves to the box with mobiles and explains that the second mobile will be selected from the remaining mobiles.

Figure 1
Animated models can be effective instructional methods for a number of reasons. To start with, they are in line with the current focus on lifelong learning and flexibility in task performance that increasingly emphasize the modeling of cognitive skills, such as problem solving and reasoning in a variety of domains (Jonassen, 1999). This type of modeling, also referred to as cognitive modeling, concerns covert cognitive processes that have to be explicated in order to become observable for learners. This enables learners not only to observe how a problem is solved, but also why a particular method is chosen. This type of information has proven to be beneficial for problem solving skills (Collins, 1991; van Gog, Paas, & van Merriënboer, 2004). Secondly, computer-based animations with verbal explanations are increasingly used to explicate the covert processes in cognitive modeling and seem to be in particular successful in learning abstract concepts and processes (Casey, 1996; Chee, 1995; Collins, 1991). Finally, developments in computer technology have facilitated the authoring and application of pedagogical agents, that is, computer-based characters that support learners with verbal feedback and guidance in order to engage them in more active learning (Clarebout, Elen, Johnson, & Shaw, 2002).

However, observing a poor designed animated model may easily overload the cognitive system. In this respect the split-attention effect is often cited in multimedia research. Split attention occurs when information from two (or more) sources must be processed simultaneously in order to derive meaning from the subject matter. Take for instance an animated model in which the written explanations are physically separated from the pictorial information. The learner has to mentally search, match, and integrate both sources of information (i.e., the animation and the text). It is impossible for the learner to attend both to the animation and the explanatory text when they are physically separated. This will cause much visual search, which is likely to pose a high load on the cognitive system without contributing to learning.

Cognitive load theory (CLT) tries to align the structure of information and the way it
is presented with human cognitive architecture (Paas, Renkl, & Sweller, 2003, 2004; Sweller, 1988, 1999, 2004; Sweller, van Merriënboer, & Paas, 1998; van Merriënboer & Sweller, 2005). CLT distinguishes between different categories of cognitive load. Intrinsic load is related to the complexity of the domain, whereas extrinsic load is determined by the manner in which the information is presented to learners. The load imposed by information and activities that hinder the learning process is called ‘extraneous’, whereas the load related to information and activities that foster the learning processes is called ‘germane’. Intrinsic, extraneous, and germane load are considered additive in that, taken together, the total load cannot exceed the memory resources available, if learning is to be maximized (see Paas, Tuovinen, Tabbers, & Van Gerven, 2003). An important objective of CLT is to decrease extraneous cognitive load and to enable learners to engage in learning activities required to perform effectively in tasks that impose germane cognitive load.

The split-attention effect can be regarded as an important source of extraneous cognitive load, after all the learner has to mentally search, match, and integrate the textual and pictorial information, which takes up much cognitive capacity without learning to commence. An instructional design guideline to overcome this split-attention effect that follows from cognitive load theory is the modality principle: Providing explanations in spoken format rather than in written format. Understanding new information involves the construction of separate mental representations for the verbal and the pictorial information and referential connections between these representations (Mayer, 2001; Mayer & Moreno, 2003). When verbal material is presented in spoken rather than in written format, cognitive demands on the visual channel are reduced which enables the learner to process the visual material and construct an adequate pictorial representation. Consequently, the verbal channel has sufficient cognitive resources to construct a verbal mental representation. Hence, the combined use of the visual channel for pictorial learning material and the verbal channel for the explanation of this material increases effectively available working memory capacity and facilitates learning.
Observational Learning from Animated (Ginns, 2005; Mousavi, Low, & Sweller, 1995).

Although the modality effect has proven its effectiveness (for a review, see Ginns, 2005), the use of the verbal channel is not always feasible. To start with, animated models with spoken explanations are technically more difficult to produce and demand additional hardware (speakers, headphones) that makes them more expensive. Secondly, for some groups of students, such as deaf or hearing-impaired students spoken explanations are no option, whereas the combination of pictorial information and textual explanations may nevertheless enhance their performance. Thirdly, there are complex tasks that put such a high demand on the auditory system that training with spoken explanation would interfere with the task. This may be the case in present-day air traffic where pilots not only have to deal with numerous visual tasks (e.g., monitoring the flight instruments), but also have to engage in voice communication with the air traffic controller. In other words, there are a number of reasons to raise the question whether an instructional method can be developed that makes the use of written explanations in animated models as effective as spoken explanations.

One of the main reasons of the modality effect to occur is the limited time for learners to process the written explanations and the visual information. In this respect a method enabling learners to process the written explanation as effective as possible can be helpful. Such a method may follow from research on attentional processing of text that has revealed that spoken text automatically receives more conscious attention than written text (Hasher & Zacks, 1979; Melara & O’ Brien, 1987; Patching & Quinlan, 2002; Posner, Nissen, & Klein, 1976). Recently, Foos and Goolkasian (2005) found that participants who were prompted to effortful and attentional processing of written words performed equally on recall to participants receiving spoken words. In their experiments they enforced effortful and attentional processing by either presenting so-called degraded words (e.g., grey bars through the word) or by having participants mentally rehearse the words. These studies indicate that learners are inclined to more effortful processing when they are prompted to pay attention to
written explanations. In multimedia the limited cognitive capacity has to be effectively distributed between the written text and the pictorial information. A method supporting an effective use of this limited cognitive capacity by prompting learners to pay extra attention to and effortful processing of the presented information in the available time may be promising. In this study we use prompts to reflect on the information for this purpose. When learners know they have to reflect on the information in the animated model they will generate thoughts and this requires them to rehearse, at least part of, the information.

For learning from animated models in the domain of probability calculation, we hypothesize that the provision of reflection prompts stimulates learners to engage in effortful processing of written explanations and enables them to construct a coherent mental representation. In contrast, reflection prompts will have no effect for spoken explanations and thus not facilitate the construction of a coherent mental representation. To test this hypothesis, a factorial design was used with the factors modality (written, spoken) and reflection (reflection prompts, no reflection prompts). If the hypothesis is true, we expect learners in the conditions with written explanations to benefit from reflection prompts yielding higher transfer performance, whereas learners in the conditions with spoken explanations are not expected to profit from reflection prompts and thus will not enhance their transfer performance.

Method

Participants

Participants were 98 pupils of pre-university education in the Netherlands (50 females and 48 males). Their mean age was 15.8 years ($SD = .73$). The participants were paid 10 euro for their collaboration. A factorial design was used with the factors modality (written, spoken) and reflection prompts (yes, no). The participants were randomly assigned to one of the four conditions. Due to technical problems, the data of two pupils in the condition with written text and no reflection prompts was excluded. This resulted in 25 participants in the condition with
written text and reflection prompts; 22 participants in the condition with written text and no reflection prompts; 25 participants in the condition with spoken explanations and reflection prompts, and 24 participants in the condition with spoken text and no reflection prompts.

*Computer-based learning environment*

The computer-based learning environment was developed with Flash MX. The computer-based learning environment consisted of the following parts: A demographic questionnaire, a prior-knowledge test, an instructional component, and an assessment component. The participants could decide themselves how much time they spent on each part. The experiment started with a demographic questionnaire with questions about gender, age, the profile of their study, and the mathematics subjects they engaged in as well as the difficulty level of these mathematics subjects. The prior-knowledge test following the demographic questionnaire consisted of 8 open questions and 4 multiple-choice questions of varying difficulty. An example of an open question is:

‘You are playing a game with some friends and it is your turn to throw a die. If you throw sixes you win. What is the probability that you throw sixes?’

An example of a multiple choice question is:

‘You have a deck of cards from which you select 4 cards. You want to get an ace, king, queen and jack in this specific order. Does it matter whether you put back the selected cards before each new selection or not?

a. Yes, your chances increase when you put back the selected cards

b. Yes, your chances decrease when you put back the selected cards

c. No, your chances remain the same whether you put back the selected cards or not

d. This depends on the number of jokers in the deck of cards’

The instructional component consisted of an introduction to probability calculation and the experimental treatment. The introduction comprised a brief explanation of concepts in probability calculation, such as randomization, individual events, complex events, and how
Observational Learning from Animated counting can be used in calculating the probability. After this introduction, which was identical for all four groups, participants received condition-specific information about the learning environment. In this part the participants in the conditions with reflection prompts were notified beforehand that they had to reflect on the animated models. With a continue button the participants could start the experimental treatment which consisted of eight animated models demonstrating and explaining how to solve a particular probability calculation problem. An example of such a problem is

‘In a factory mobile phones are produced. On a production line the mobiles receive a cover in one of six colors before they are packed in a box. Each box contains six mobiles in the colors red, black, blue, yellow, green, and pink. Before a box leaves the factory two mobiles are selected randomly and checked on deficiencies. What is the probability that you select the yellow and the blue mobile from one box?’

The animated models were grouped in four problem categories which resulted from two important characteristics in probability calculation: The order of drawing (relevant vs. irrelevant) and replacement of drawing (without replacement vs. with replacement). For each problem category two animated models were presented to enable learners to recognize structural similarities and dissimilarities between problems and thus learn not only how to solve problems but also when to apply which procedure. Table 1 shows the order in which the animated models were presented.

Insert Table 1 here

In all conditions participants were presented the same eight animated models. All animated models were continuous and learner paced, that is, participants could use a pause and play button and they could restart the animated model from the beginning. Each animated model depicted the problem-solving process and was completed with supportive spoken explanations (conditions with spoken modality) or written explanations (conditions with
written modality) by a pedagogical agent that was implemented as a dolphin. Table 2 provides an example of the explanatory text, which was used for the mobile animated model.

Insert Table 2 here

The animated pedagogical agent moved across the screen to focus the learners’ attention while one of two possible problem-solving methods was explained and demonstrated. The method of individual events was applied to four animated expert models. This method implies that first the probability of individual events is calculated separately after which the complex event is calculated by multiplying the individual events. In the mobile phone problem the probability of selecting the yellow and the blue mobile was calculated first (respectively 2/6 and 1/5) and these two probabilities were subsequently multiplied for calculating the probability of the complex event. The method of counting was applied in the other four animated expert models. This method implies that all possible combinations are balanced by the correct number of combinations. For example, suppose someone calculates the probability to guess a PIN code consisting of 4 figures. For each figure 10 different numbers (0 up to and including 9) can be chosen, whereas for 4 figures 10*10*10*10, that is 10,000, possible combinations can be chosen of which only one combination is correct. In the animated expert models the pedagogical agent explicated which considerations underlie the choice of one of the two methods.

In the spoken explanations-reflection prompts condition the participants could listen to a narrated animated model which was spoken by a male voice without accent. Immediately after the last animated model in each problem category, that is, after the ‘running contest’, ‘PIN code’, ‘Checking mobiles in a factory’, and ‘Finding figures in a cereal box’ animated models, participants received a screen with the question ‘Please, write down how the problem in the last animation was solved’. Their reflection had to be written in a textbox on the screen and was logged by the computer. With a continue button they could then proceed to the next animated model. The spoken explanations-no reflection prompts condition was identical to
the spoken explanations-reflection prompts condition except that no reflection prompts and thus no textbox appeared after the last animated model in each problem category. The written explanations-reflection prompts condition was identical to the spoken explanations-reflection prompts condition. The exception was that the explanations were written and appeared in a text balloon -originating from the animated pedagogical agent- very close to the place in the animated model it was referring to (see Figure 2 for a screen shot of the written explanations condition). Finally, the written explanations-no reflection prompts condition was identical to the written explanations-reflection prompts condition except that no reflection prompts and thus no textbox appeared after the last animated model in each problem category.

Insert Figure 2 here

In all conditions, after each animated model, the participants were asked to score the mental effort they perceived when they studied the animated model on a one-item 9-point rating scale based on Paas (1992; see also Paas et al., 2003). This scale ranged from ‘very, very little effort’ to ‘very, very much effort’. After the instructional component an assessment component followed consisting of twelve transfer tasks. Of the twelve tasks, eight tasks were near transfer tasks. The near transfer items were analogous to the problems solved in the animated models. The following is an example of a near transfer task:

‘In a pop music magazine you see an ad under the heading FOR SALE in which a ticket for a spectacular concert by your favorite pop group is offered. Unfortunately the last 2 digits of the telephone number, where you can obtain information about the ticket, are not readable anymore. You really like to have the ticket and decide to choose the 2 digits randomly. What is the probability that you dial the correct digits on your first trial?’

The remaining four items were far transfer items. These far transfer items were different from the problems solved in the animated models. Take for instance the following example of a far transfer task:
‘In order to determine the final mark for a subject your teacher uses two complementary methods. First, you have to perform a practice task, followed by a test consisting of 8 multiple-choice questions. One out of five possible practice tasks (named A, B, C, D and E) is randomly assigned to you. You know you have done practice on task E a month ago. For the multiple-choice questions your teacher uses a large pool of 100 different multiple-choice questions from which he randomly selects 8 questions for you. You have made a test before with 8 questions from this pool. What is the probability that you are assigned practice task E as well as the 8 questions you have had before?’

This far transfer task comprises a problem from a specific problem category, that is, order is not important and without replacement (the drawing of the multiple choice items), which has to be combined with one individual event (the practice task). After each near and far transfer task the participants were asked to score the mental effort they perceived when they solved the transfer task on the one-item 9-point rating scale described above.

**Procedure**

The experiment was conducted in one session and was run in the computer rooms of the participating schools. Each computer had a headset to listen to the verbal explanations. After welcoming the participants, the experimenter gave them a code to log in to the experimental environment. When the participants entered the environment, the purpose of the experiment was explained on the computer screen and an outline was given of the different parts of the experiment. First, participants had to fill out the demographic questionnaire on the computer. Then, the prior-knowledge test was conducted. The instruction phase started after the prior-knowledge test with the brief introduction to probability calculation. After reading the introduction participants could press a continue button to study the animated models. After each animated model, they were asked to score their perceived mental effort. By pressing a button they could proceed to the next animated model. After two animated models,
participants in the conditions with reflection prompts received a screen in which they had to reflect on the last presented animated model. Following the instruction phase, the transfer test was administered. Participants could use a calculator as well as scrap paper during the transfer test. All input to the calculator was logged and the scrap paper was collected after the experiment. After each transfer item they were asked to score their invested mental effort. Finally, the participants were debriefed and thanked for their participation.

**Scoring**

For each open question of the prior-knowledge test a list of correct answers was formulated. For each correct answer 1 point was assigned, otherwise 0 points. Computational errors were ignored and no partial credits were awarded. For each correct multiple-choice question participants received 1 point, otherwise they received 0 points. In total the maximum score on the prior-knowledge test could be 12 points. The mental effort scores after studying the animated models were summed across all eight animated models and divided by 8, resulting in an average score on mental effort ranging from 1 to 9. For each near and far transfer task a list of correct answers was formulated. Computational errors were ignored and no partial credits were awarded. Each near and far transfer item was assigned 1 point when it was correct and 0 points when it was incorrect. The maximum score for the near transfer test was therefore 8 points, and for the far transfer test 4 points. The mental effort scores after solving the near and far transfer tasks were summed across the eight near and the four far transfer tasks and divided by respectively eight and four, resulting in an average score on mental effort on near and far transfer ranging from 1 to 9. Instruction time (in s) was defined as the time participants needed for the introduction (the basic theory of probability calculation) and the instruction component (the time spent on observing the animated models). For the conditions who received reflection prompts, the time spent on reflection was included in instruction time. The time (in s) needed to accomplish the transfer tasks was
Observational Learning from Animated
logged by the computer. The computer logged both the start time and the end time of the instruction.

Results

The dependent variables under investigation were instruction time (s), mental effort during instruction (score 1-9), performance on transfer (score 0-12), mental effort on transfer (score 1-9), and time on transfer tasks (s). For all statistical tests a significance level of .05 was applied. Effect sizes are expressed in terms of omega-squared ($\omega^2$). We began our analysis with testing the dependent measures that could be used as covariates for further analyses. Table 3 shows the mean scores and standard deviations of performance on the prior-knowledge test and the dependent variables for all conditions.

The mean score on the prior knowledge test was 6.00 ($SD = 2.05$), indicating that the participants had some knowledge (the maximum score was 12) regarding the subject matter. Research has shown that the level of prior knowledge interacts with the effectivity of instructional material, that is, design guidelines that are beneficial for novice learners can be ineffective or even detrimental when applied to experts (Kalyuga, 2005; Kalyuga, Ayres, Chandler, & Sweller, 2003). Moreover, it is found that in the domain of probability calculation the level of prior knowledge has an effect on the quality of self-explanations (Atkinson, Renkl, & Merrill, 2003). Therefore the prior-knowledge score was included as a covariate (see also Atkinson, 2002). For instruction time, an ANOVA with the between-subjects factors modality and reflection showed a significant effect of reflection ($F(1, 92) = 51.01, MSE = 100,769.26, p = 0, \omega^2 = 34\%$), indicating that learners in the reflection conditions needed more instruction time than learners in the no reflection conditions ($M = 2,124$ and $SD = 382$ vs. $M = 1,659$ and $SD = 219$). No main effect of modality ($F(1, 92) < 1, ns, \omega^2 = 0\%$) nor an interaction between modality and reflection was found ($F(1, 92) < 1, ns, \omega^2 = 0\%$). Next, we tested for time on transfer tasks to determine whether it should be used as a covariate in further analyses. No
differences were found on time on transfer tasks for modality, \( F(1, 92) < 1, ns, \omega^2 = 0 \), and reflection, \( F(1, 92) = 2.76, MSE = 213,218.43, ns, \omega^2 = 1\% \). No interaction for modality and reflection was found on time on transfer tasks \( (F(1, 92) = 3.04, ns, \omega^2 = 2\%) \). For each measure the homogeneity of regression was tested and all results were found to be nonsignificant, \( F < 2.3 \). Therefore, scores were analyzed with ANCOVAs with the between-subjects factors modality (spoken explanations vs. written explanations) and reflection (reflection prompts vs. no reflection prompts), and the covariates prior knowledge and instruction time.

For performance on transfer no difference could be observed for either modality \( (F(1, 92) < 1, ns, \omega^2 = 0) \) or reflection \( (F(1, 92) < 1, ns, \omega^2 = 0\%) \). However, the interaction between modality and reflection on performance on transfer, which is depicted in Figure 3, was significant, \( F(1, 92) = 6.81, MSE = 3,2307, p = .011, \omega^2 = 5\% \). The interaction indicates that written explanations yielded better transfer performance with reflection prompts than without reflection prompts \( (M = 5.12, SD = 2.06 with reflection vs. M = 4.54, SD = 1.50 with no reflection) \), whereas for spoken explanations no difference was found between reflection prompts and no reflection prompts \( (M = 4.79, SD = 2.26 with no reflection vs. M = 4.16, SD = 2.39 with reflection) \).

Post-hoc multiple comparisons based on the adjusted means were conducted using Bonferroni’s procedure. This analysis showed that learners in the condition with written explanations and reflection prompts performed better on transfer than the learners with written explanations without reflection prompts. The other comparisons revealed no differences.

For mental effort during instruction no difference was found for either modality \( (F(1, 92) = 2.96, MSE = 1,189, ns, \omega^2 = 1\%) \) or reflection \( (F(1, 92) = 1.95, ns, \omega^2 = 0\%) \). Neither an interaction between modality and reflection was observed \( (F(1, 92) < 1, ns, \omega^2 = 0\%) \). Finally,
no differences were observed on mental effort on transfer for either modality ($F(1, 92) = 1.20$, $MSE = 2,706$, $ns$, $\omega^2 = 0\%$) or reflection ($F(1, 92) < 1$, $ns$, $\omega^2 = 0\%$). Neither an interaction between modality and reflection was found ($F(1, 92) < 1$, $ns$, $\omega^2 = 0\%$).

Discussion

The aim of this study was to investigate whether having learners attend and effortfully process written explanations could compensate for the modality effect. In line with our hypothesis, learners who received written explanations with reflection prompts yielded higher performance on transfer tasks than those who received written explanations without reflection prompts. For spoken explanations, no difference on transfer performance between reflection prompts and no reflection prompts was observed. In line with the modality effect a difference could have been expected between the conditions with spoken and written explanations when no reflection prompts are involved. Although the two conditions differ on transfer performance, the difference is not statistically significant ($p = .07$). The learners in this study were not real novices, they possessed some prior knowledge. It is possible that their prior knowledge enabled them to deal with the written text in conjunction with the pictorial information. No effects were found on mental effort during instruction and mental effort on transfer. The mental effort measure used did not differentiate between mental effort due to perceived difficulty of the subject matter, presentation of the instructional material, or being engaged in relevant learning activities. It is possible that the effect on the perceived mental effort of the varying design guidelines, that is, modality and reflection prompts, have neutralized each other. For example, during the instruction the reflection prompts may have yielded an increase in germane load for the written explanation condition, but at the same time have decreased the intrinsic cognitive load due to a better understanding of the subject matter. The absence of reflection prompts in the written explanation with no reflection prompts condition, on the other hand, may neither have yielded an increase in germane cognitive load nor a decrease of intrinsic cognitive load.
From a theoretical point of view these results contribute to a better understanding of the modality effect. The modality effect in multimedia learning assumes that working memory is used more effectively when the verbal channel is used for spoken explanations and the visual channel for the pictorial or visual information. According to this view, accompanying complex visual materials with written explanations may result in too much load on the visual channel and thus to inferior processing of the instructional materials (Ginns, 2005; Mousavi, Low, & Sweller, 1995). Research focusing on attentional processing of text that has revealed that written text automatically receives less conscious attention than spoken text (Foos & Goolkasian, 2005). When learners are stimulated to attend to this information by mentally rehearsing or repeating the textual information, spoken text is no longer superior to written text. The results of the present study suggest that the role of attention in modality also pertains to multimedia learning. These results corroborate the findings of other studies indicating that the modality effect can be compensated for. For example, Tabbers, Martens, and van Merriënboer (2002, 2004) found that written explanations were more effective than spoken explanations when learners had control over the pace of presentation. This was ascribed to the lack of time pressure in learned-controlled conditions, which enabled learners to take full advantage of the characteristics of written media and read the written explanations in such a way that they could select relevant parts of the text and skip irrelevant parts. Apparently, there are conditions (self-pacing, prompting attention) under which the modality effect does not hold true anymore.

On the practical side the results of this study may have implications for instructional designers when spoken explanations are no option such as in the education of the hearing-impaired. Written explanations combined with visual material can be an effective learning arrangement provided that learners are stimulated to process these written explanations, for example by prompting them to reflect. Conversely, reflection takes more instruction time, but it does not pay back in better performance when it is combined with spoken explanations.
This is all the more important since the creation of spoken explanations for animations is more time-consuming and expensive than the creation of written explanations.

The findings and conclusions also provide directions for future research. To start with, it should be noted that learners in the condition with reflection prompts had to reflect in a written format (i.e., they typed their considerations in a text box) regardless of the modality of the explanations. From this point of view, learners in the condition with written explanations may have had an advantage because the way they could express their reflection was more in line with the modality of received explanations. For this reason a replication study is required in which not only the modality of the explanations is varied but also the format in which the learners reflect (i.e., spoken or written).

Secondly, the reflection prompts used in our study asked learners to reflect on how the problems were solved, that is, it prompted learners to focus on the method that was used to solve the problem. However, reflection can also be implemented by asking learners to explain a correct answer or solution in a multimedia learning environment (Moreno & Mayer, 2005). An interesting avenue for future research therefore might be to compare the effects of reflection and modality when learners are prompted to think about how a problem was solved, why in their opinion the solution was correct, or a combination of both.

Thirdly, modeling is about observing someone performing a complex skill with the intention to perform the problem solving skill yourself in a later stage. Reflection can be regarded as a link between observing and performing: Learners do not yet perform the problem solving skill, but they actively think about the solution without solving the problem. A next step might be to alternate between observing the problem-solving process and independently solving a novel problem. Modeling research in the domain of motor skill acquisition has shown that learners benefit from alternating between observing and practicing (Shea, Wright, Wulf, & Whitacre, 2000; Weeks & Anderson, 2000). Consequently, comparing an arrangement in which learners only observe performance in the domain of
problem solving with a situation in which observation and practicing are alternated could
ground a better understanding of the relation between observing and practicing.

Finally, it was found that learners with spoken explanations and reflection prompts
performed slightly worse on transfer tasks than those who did not receive reflection prompts
\( (M = 4.29 \text{ for reflection prompts vs. } M = 4.99 \text{ for no reflection prompts}) \). This suggests that
the reflection prompts may have interfered with the processing of the spoken explanations.
Some support for this is provided by a series of studies conducted by De Beni and colleagues
(De Beni, Moè, & Cornoldi, 1997; De Beni & Moè, 2003; Moè & De Beni, 2005). In their
studies they combined the modality of text (spoken vs. written ) with instruction strategy
(imagery vs. mental rehearsal) and found that on a free recall test written texts yielded better
performance when a mental rehearsal strategy was used, whereas spoken texts yielded better
performance with an imagery strategy. They argued that the mental rehearsal of written text
would take place in the verbal channel that would have sufficient resources available. An
imagery strategy, on the other hand, would impose load on the visual channel and thus
interfere with the processing of the written texts. With spoken texts the pattern is reversed:
Mental rehearsal may interfere with the processing of the spoken text in the verbal channel,
whereas imagery facilitates learning since it takes place in the visual channel. This suggests
that the modality effect is contingent on the kind of active processing (rehearsal vs.
imagination) that learners engage in. For this reason more research is required on the
interaction between the modality of presentation and design guidelines that foster active
processing.

Finally, the results of this study have some clear limitations. To start with, the
instructional material was only used with one particular type of learners (i.e., pupils of pre-
university education) who also had some relevant prior knowledge. Secondly, a specific
domain was used, that is, probability calculation, which is procedural in nature rather than
describing a causal chain of events (e.g., science). Thirdly, the learner reflections were limited
in both their modality (i.e., learners could only reflect in written format) and their quantity (i.e., learners were prompted to reflect only four times, after the last animated model in each problem category). In addition, the results are limited because the experiment included two different aspects regarding reflection. Beside the fact that the participants had to reflect, they were also informed beforehand that they had to reflect. More research is needed to uncover the exact relationship between these two aspects.

Our study showed that reflection prompts might compensate for the modality effect, but more research is needed to further specify the conditions under which such compensation occurs.
References


Table 1

*Order in which the Animated Models are Presented and the Distribution across the Problem Categories*

<table>
<thead>
<tr>
<th>Order of presentation</th>
<th>Context of animated model</th>
<th>Problem category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mountainbike trip with your friend(^1)</td>
<td>Order relevant / without replacement</td>
</tr>
<tr>
<td>2</td>
<td>Running contest</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Mountainbike trip with your friend(^1)</td>
<td>Order relevant / with replacement</td>
</tr>
<tr>
<td>4</td>
<td>PIN code</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Mountainbike trip with your friend(^1)</td>
<td>Order irrelevant / without replacement</td>
</tr>
<tr>
<td>6</td>
<td>Checking mobiles in a factory</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Mountainbike trip with your friend(^1)</td>
<td>Order irrelevant / with replacement</td>
</tr>
<tr>
<td>8</td>
<td>Finding figures in a cereal box</td>
<td></td>
</tr>
</tbody>
</table>

Note: \(^1\) Animated models 1, 3, 5, and 7 share the same context. The problems that have to be solved are different.
### Table 2

**Explanatory Text for the Mobile Animated Model. At any moment, only the text in one row was visible to the participants**

<table>
<thead>
<tr>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. In a factory mobiles are produced. On a production line the mobiles receive a cover in one of six colours before they are packed. Each box contains six mobiles in the colours red, black, blue, yellow, green, and pink. Before a box leaves the factory two mobiles are selected randomly and checked on deficiencies.</td>
</tr>
<tr>
<td>2. What is the probability that you select the yellow and the blue mobile from one box?</td>
</tr>
<tr>
<td>3. The order of the drawing is not important. It only matters that both the yellow and the blue mobile are selected. Whether the yellow mobile is drawn first, and thereafter the blue mobile. Or first the blue mobile and then the yellow one, does not matter.</td>
</tr>
<tr>
<td>4. This is a drawing without replacement. When the first mobile is drawn, it can not be replaced. Otherwise you could draw a mobile that was controlled already. Of course that is not supposed to happen.</td>
</tr>
<tr>
<td>5. The order is not important in this case. Therefore you can work with individual events. The probability of the first individual event is 2/6. The first mobile can be blue or yellow. From the six possibilities, two are correct.</td>
</tr>
<tr>
<td>6. The probability of the second individual event is 1/5. Suppose that the first controlled mobile was blue. The second mobile then has to be yellow. However, only five mobiles are still in the box.</td>
</tr>
<tr>
<td>7. The two events are independent. Therefore the two probabilities have to be multiplied: $2/6 \times 1/5$. The probability of the complex event is 1/15 or 0.067</td>
</tr>
</tbody>
</table>
Table 3

*Mean Scores and Standard Deviations on Performance on the Prior Knowledge Test and Dependent Variables for all Conditions*

<table>
<thead>
<tr>
<th></th>
<th>Written explanations</th>
<th>Spoken explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reflection prompts</td>
<td>No reflection prompts</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Prior Knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior knowledge test (0 -12)</td>
<td>5.76</td>
<td>2.20</td>
</tr>
<tr>
<td>Instruction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instruction time (s)</td>
<td>2,103</td>
<td>380</td>
</tr>
<tr>
<td>Mental effort during instruction (1-9)</td>
<td>2.83</td>
<td>1.26</td>
</tr>
<tr>
<td>Transfer Test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance on transfer (0-12)</td>
<td>5.12</td>
<td>2.06</td>
</tr>
<tr>
<td>Mental effort on transfer (1-9)</td>
<td>4.15</td>
<td>1.77</td>
</tr>
<tr>
<td>Time on transfer tasks (s)</td>
<td>1,446</td>
<td>445</td>
</tr>
</tbody>
</table>

*Note:* AM means adjusted means. Scores with * differ statistically.
Figure Captions

Figure 1. Screen shot of the ‘Checking mobiles’ animated model which displays and explains why this is a ‘drawing without replacement’.

Figure 2. Screen shot of the ‘Checking mobiles’ animated model with written explanations.

Figure 3. The interaction of modality and reflection on transfer test performance (based on adjusted means).
This is a drawing without replacement. When the first mobile is drawn, it cannot be replaced. Otherwise you could draw an already controlled mobile again. Of course that is not supposed to happen.
This is a drawing without replacement. When the first mobile is drawn, it cannot be replaced. Otherwise you could draw an already controlled mobile again. Of course that is not supposed to happen.
Interaction Modality and Reflection

- **Reflection prompts**
- **No reflection prompts**

Performance on transfer (adjusted means):

- **Written**
  - Reflection prompts: 5.26
  - No reflection prompts: 4.00

- **Spoken**
  - Reflection prompts: 4.99
  - No reflection prompts: 4.29

Modality of explanation