# Making the abstract concrete: Visualizing mathematical solution procedures

Katharina Scheiter, Peter Gerjets, Richard Catrambone

## HAL Id: hal-00197364
## https://telearn.hal.science/hal-00197364

Making the Abstract Concrete: Visualizing Mathematical Solution Procedures

Katharina Scheiter

University of Tuebingen

Peter Gerjets

Knowledge Media Research Center

Richard Catrambone

Georgia Institute of Technology

Correspondence concerning this article should be addressed to Katharina Scheiter,

Applied Cognitive Psychology and Media Psychology, Konrad-Adenauer-Strasse 40,

72072 Tuebingen, Germany. Electronic mail may be sent to k.scheiter@iwm-kmrc.de.

Abstract

This article reports on an experiment investigating the effects of different kinds of computer-based visualizations on the acquisition of problem-solving skills in the domain of probability theory. Learners received either purely text-based worked examples, text plus an instruction to mentally imagine the examples' contents, or the possibility of retrieving either static pictures or concrete animations that depicted the problem statement and the problem states achieved by applying a specific solution-step. It could be shown that frequent use of both static pictures and imagining the examples' contents improved performance on isomorphic problems. Frequent use of animations, in contrast, led to substantial increases in learning time and a decrease in performance. Thus, the use of concrete animations to visualize solution procedures was more harmful than helpful for conveying problem-solving skills. Reasons for these outcomes are discussed in the light of cognitive load theory.


Keywords: problem solving, skill acquisition, worked-out examples, visualizations, pictures, animation, imagery, cognitive load

Making the Abstract Concrete: Visualizing Mathematical Solution Procedures

Understanding mathematical solution procedures is a necessary prerequisite for solving both familiar problems, as well as novel problems requiring adaptation of acquired solution procedures. However, it has been noted that students often face severe difficulties in understanding solution procedures even when they have received elaborate instructional explanations of the individual solution-steps. This may result from the fact that the solution-steps are often conveyed in a rather abstract way so that learners experience difficulties in imagining which changes in the problem state are achieved by applying a specific solution-step to a problem.

The use of multimedia learning environments may offer ways to overcome these difficulties (Mayer, 2001). In multimedia learning environments, information presentation can be accomplished by using different representational formats (textual and pictorial) which may be processed in different sensory channels (auditory and visual). Additionally, information presentation is not restricted to static displays (e.g., diagrams, pictures, written text), but the representations used can involve changes over time (e.g., dynamic visualizations, spoken text).

This research addresses the effects of augmenting a purely text-based hypertext environment called HYPERCOMB by different kinds of visualizations. HYPERCOMB teaches students how to calculate the probability of complex events. Calculating complex-event probabilities are related to situations where the probability of randomly selecting a particular configuration of elements out of a set of elements has to be determined (cf. Figure 1).

The main instructional principle underlying HYPERCOMB is the use of worked-out examples for conveying knowledge on different problem categories. Worked-out examples consist of a problem statement and a step-by-step solution procedure. Research over the last

especially in well-structured domains like mathematics, physics, or computer programming (Atkinson, Derry, Renkl, & Wortham, 2000).

In a series of prior experiments, different example formats in HYPERCOMB (Gerjets, Scheiter, & Catrambone, 2004) were compared. Conventionally designed or molar examples focus on explaining how to categorize problems according to multiple structural task features and how to apply category-specific solution formulas. Molar examples demonstrate a fast and convenient approach to problem solving, but may also impose high cognitive demands on learners who have to simultaneously consider multiple problem features in order to solve a problem. Thus, an alternative modular example format was developed which completely avoids references to molar concepts like problem categories or formulas and where complex solutions are broken down into smaller, meaningful solution elements that can be conveyed separately. Though it was demonstrated that a modular example format reduces cognitive load and boosts performance compared to conventionally designed molar examples (Gerjets et al.), there is still space left for improvement. We assumed that this space might be claimed by the benefits achieved through the use of visualizations.

According to Mayer's multimedia principle (2001) embellishing textual learning materials by static pictures or dynamic visualizations (i.e., animations) helps promote the learners' understanding of instructions. With regard to the acquisition of problem-solving knowledge, visualizations of worked-out examples may first help learners understand the situation described in the problem statement (i.e., the initial problem state) and thus to correctly represent its meaning in a situation model (Nathan, Kintsch, & Young, 1992). Second, visualizations of the solution-steps may promote an understanding of changes with regard to the initial problem state that are achieved by applying a solution-step to a problem. Visualizing worked-out examples can be done by presenting either static pictures, animations, or an instruction to mentally imagine the examples' contents.

Static pictorial representations are known to foster the immediate and delayed retention of facts contained in the accompanying text (cf. for a review Carney & Levin, 2002; Levin, Anglin, & Carney, 1987). Moreover, with regard to the acquisition of problem-solving knowledge in domains like mathematics and physics, the added value of abstract diagrammatic and graphical representations has been acknowledged (e.g., Shah & Hoeffner, 2002; Vekiri, 2002). These types of visualizations are said to be computationally effective in that they facilitate specific inferential processes needed for some learning tasks (Larkin & Simon, 1987). However, the researchers have found no research on whether concrete pictures depicting a problem situation and its modification caused by the application of a solution procedure promote acquisition of problem-solving knowledge in abstract domains. There is research on the use of video demonstrating the application of observable procedural skills (e.g., tying nautical knots, Schwan & Riempp, 2004); however, this does not allow for predicting the effectiveness of concrete visualizations for teaching abstract mathematical skills that do not have a natural representation in the real world.

It can however be assumed that representing the problem statement in a picture will help to understand which features and interrelations of objects are relevant to the solution of the problem and thus to build a situation model of it (Nathan et al., 1992). For instance, in the sprinter example in Figure 1 seven sprinters on the racetrack are depicted out of whom three can win a medal (i.e., gold, silver, and bronze), represented by the pedestal. Additionally visualizing the solution-steps may support learners in inferring which change is achieved by applying a solution procedure when they compare the new problem state to the previous one. For instance, comparing the picture illustrating the second solution-step with the one depicting the first step helps clarify the fact that only six sprinters are eligible for the silver medal in the second solution-step, because one sprinter has already been assigned the gold medal in the first solution-step and thus already stands on the pedestal.

An animation is "any application which generates a series of frames, so that each frame appears as an alteration of the previous one, and where the sequence of frames is determined either by the designer or the user" (Betrancourt & Tversky, 2000, p. 313). Thus, animations do not only depict the current state of objects; they also deliver information concerning changes of objects and their position over time (motion) as well as information concerning the direction of these changes (trajectory, Rieber, 1990). Several findings suggest that animations can be used successfully for delivering abstract content such as mathematical rules, Newton's laws, or computer algorithms (Baek & Layne, 1988; Byrne, Catrambone, & Stasko, 1999; Catrambone & Seay, 2002; Rieber). With respect to conveying problem-solving knowledge, the visual-spatial properties of the visualization may be used to deliver information on the current problem state and its relevant structural features. Moreover, the changes over time that can be depicted in an animation may be used to reflect the changes in problem states that result from applying a solution-step to a specific problem state of the example – without the need to compare multiple representations as is necessary when using static pictures.

However, learning from animations is known to impose certain requirements on the learners that they may have difficulty meeting (Betrancourt & Tversky, 2000). Pane, Corbett, and John (1996) have demonstrated that learners often fail to use animations to a sufficient extent and thus miss important information. Lowe (1999) suggests that learners have trouble focusing on the most relevant parts of an animation and are often distracted by salient, but irrelevant details. Furthermore, due to the dynamic changes of the display, the information that has to be remembered can only be viewed for a limited amount of time and may therefore have vanished before learners have identified it. Finally, dynamic visualizations may lead to overly passive information processing and prevent learners from performing effortful cognitive processes required for a deeper understanding (Palmiter & Elkerton, 1993; Schnotz,

learners may refrain from deeply processing the contents of an animation and passively watch it like an ongoing movie.

*Imagery*

To circumvent the problem of passive information processing, learners can be instructed to imagine the contents of a text-based worked-out example by mentally envisioning objects and their relations included in the examples. In this case, learners do not receive any pictorial representations, but are told to construct their own visualizations. Hodes (1992) compared the effectiveness of imagery instructions and instructional visuals for fact recall and understanding. Both instructional methods were helpful in inducing an imagery strategy and in improving posttest performance; however, for some performance measures, achievements due to presenting external visuals were larger than for the imagery instructions. Ginns, Chandler, and Sweller (2003) showed that imagining procedures (vs. studying text-based materials) was only helpful for learners possessing sufficiently high prior knowledge in the domain. This replicates findings of Cooper, Tindall-Ford, Chandler, and Sweller (2001) who found interactions between domain-specific abilities and the type of instruction given. Ginns et al. explained these findings by assuming that "to successfully imagine a procedure or a concept, all of the relevant elements must be processed simultaneously in working memory. That may be possible only after a schema had been constructed" (p. 231). Contrary to the experiment outlined in the remainder of this paper, learners in these studies (Cooper et al.; Ginns et al.) were not instructed to imagine a worked-out example's concrete objects. Rather, they were told to envision the operations they needed to perform in order to solve a problem (e.g., adding an addition sign to an equation) and the examples did not contain any concrete elements, but only the mathematical information.

Comparing the three different visualization methods (static pictures, animations, and imagery) to each other yields insights on specific advantages as well as drawbacks associated

learning outcomes (Table 1). Learning from *static pictures* may engage learners in a more active way, because they have to compare multiple visualizations in order to understand the to-be-learned solution procedures. As this information is permanently visible, these comparisons may be conducted without overloading the cognitive system. However, learners may make erroneous inferences and may miss important information so that their internal representation of the solution procedures may be incorrect as well as incomplete. Moreover, because the external representations cannot be modified by a learner, they are not adapted to his or her preferences or prior knowledge level. *Dynamic pictures* may reduce learner activity and induce a rather passive style of processing. The cognitive demands imposed by the need to extract the relevant information from a changing display and to memorize this information may be rather high. Unless animations are highly interactive, they do not allow for any adaptation to a learner's preferences or prior knowledge. However, an advantage of dynamic visualizations of solution-steps is that all the information that is needed to understand problem states and their changes is in principle contained in the representation and thus it is correct as well as complete. Finally, while *imagery instructions* should foster learner activity in an optimal way, the need to envision all information may require rather demanding processes that may be overwhelming for learners with low cognitive prerequisite knowledge in the domain (Cooper et al., 2001; Ginns et al., 2003). There is also the danger that learners may miss important information or that they make incorrect inferences. A possible advantage of mental imagery is that, because it is based on self-generated images, these images are adapted to a learner's preferences and prior knowledge level.

These trade-offs between possible promises and drawbacks were investigated in the experiment outlined in the remainder of the article.

## Experiment

*Method*

*Participants*. Participants were 124 students (88 female, 36 male) at the University of Tuebingen, Germany, who participated for either course-credit or payment. One participant in the static-pictures condition had to be excluded from the analysis, because learning time was more than four standard deviations above the mean learning time in this condition. Twenty-nine participants were in the text-only condition, 30 in the static-pictures condition, 32 in the animation condition, and 32 in the imagery condition. The average age was 23.60 years ($SD = 4.28$). Most of the participants were familiar with the domain of probability theory and were experienced computer users.

*Materials and procedure.* HYPERCOMB consisted of a technical instruction, a short introduction to the domain, an example-based learning phase, and a subsequent test phase. Before starting with the experiment, participants filled in a multiple-choice questionnaire with eleven questions on important concepts and definitions from the field of probability theory. This questionnaire was used to measure prior knowledge. In the first part of HYPERCOMB, learners were given a short technical introduction to the system and to the experiment. After that, the basic notion of random experiments and the general rationale behind calculating the probability of outcomes were explained in a short introduction to the domain. In the subsequent example-based learning phase, learners had to acquire knowledge on four different problem categories, where each category was explained by means of two worked-out examples. Depending on the experimental condition, participants also received an imagery instruction or could retrieve static pictures or dynamic visualizations that augmented the worked-out examples. Participants were not forced to process these visualizations; rather, they could choose to select them by clicking on a button in order to view them. The visualizations depicted the contents of the worked-out examples in a concrete way (cf. Figure 1). For every worked-out example there was always one visualization of the problem statement and one of every worked-out solution-step. In accordance with the coherence

cluttered with irrelevant details. Participants could decide by themselves when to start

working on the test problems. The instructional materials were no longer available during

problem solving.

Before solving the test problems, learners had to give an estimate of the cognitive load

they had been experiencing during learning (see 'dependent measures' for details). For the 11

test problems transfer distance with respect to the worked-out examples was varied by

presenting five isomorphic and six novel problems. Isomorphic test problems differed from

the instructional examples only with regard to their surface features. Novel test problems were

constructed in a way that two complex-event probabilities had to be considered whose

outcomes had to be multiplied in order to calculate the required probability. An example of a

novel test problem would be:

> At a soccer stadium, there are two dressing rooms for the two opposing teams.
>
> The 11 players from Oxford wear T-shirts with odd numbers from 1 to 21 and
>
> the 11 players from Manchester have even numbers from 2 to 22. Because the
>
> aisle from the dressing rooms is very narrow only one player at a time can enter
>
> the field. The players of the two teams leave their rooms alternately with a
>
> player from Oxford going first. What is the probability of the first five players
>
> entering the field having the numbers 5, 2, 13, 8, and 1 (i.e., the first has the
>
> number 5, the second has the number 2, and so on)?

*Design and dependent measures.* Participants studied in one of four instructional

conditions. In the *text-only condition* only the written worked-out examples were available. In

the *imagery condition* learners were additionally told to mentally imagine the contents as

vividly and with as many details as possible. Mental imagery was trained at the beginning of

the experiment by instructing learners to imagine the contents of a short text passage

describing a traffic situation involving multiple cars approaching an intersection from

the example-based learning phase by a sticker "imagine the situation" attached to the frame of

the computer screen. In the *pictures condition,* static visualizations could be retrieved for each

component of the worked-out examples, whereas in the *animation condition* clicking the play-

button for any of the example components resulted in the presentation of a dynamic

visualization. In the sprinter animation, for example, the sprinters were first shown entering

the racetrack (problem statement) and were then running around the track. The visualization

of the solution-steps always depicted one sprinter passing the finishing line and ascending the

pedestal. The last frame of the animation was always used as the static picture (cf. Figure 1)

depicting the problem state resulting from the application of the respective solution-step.

The dependent measures were problem-solving performance for isomorphic and novel

test problems, time spent learning (learning time), and cognitive load. For each of the 11 test

problems, 1 point was assigned for a correct answer; no partial credits were given. The sums

across the five isomorphic problems and across the six transfer problems were each

transformed into a percentage for ease of interpretation. Cognitive load was assessed by

means of a modified version of the NASA-TLX (Hart & Staveland, 1988), which had been

successfully used in order to distinguish between different aspects of cognitive load in prior

studies (Gerjets et al., 2004). This cognitive load measure consisted of three items, each rated

on a scale from 0 (very low cognitive load) to 100 (very high cognitive load). The following

subscales were used: '*task demands*' (How much mental and physical activity was required,

e.g., thinking, deciding, calculating, remembering, looking, searching etc.), '*effort*' (How hard

the participant had to work to understand the contents of the learning environment), and

'*navigational demands*' **(**How much effort the participant had to invest to navigate the

learning environment). According to Cognitive Load Theory (Sweller, van Merriënboer, &

Paas, 1998) task demands are caused by inherent properties of the learning task (intrinsic

cognitive load), higher-level processes for achieving a deeper understanding - as reflected in

directed to learning such as decision processes required for navigation and information selection (extraneous load). Thus, a mapping is assumed between the theoretical assumptions of Cognitive Load Theory and the items of the modified version of the NASA-TLX.

The control measures were learners' intrinsic motivation prior to learning determined by a shortened version of the Intrinsic Motivation Inventory (Ryan, 1982), domain-specific prior knowledge determined by a questionnaire included in the environment, and spatial abilities determined by the mental rotation test (Vandenberg & Kuse, 1978).

*Results*

The statistical analysis is divided into three parts. First, all four instructional conditions were compared in an overall ANOVA in order to detect possible differences between them with regard to problem-solving performance, learning time, and cognitive load. In a second step, a closer look was taken at the utilization of external visualizations (i.e., static pictures and animations) and the way this affects the aforementioned variables. Finally, the external-visualization conditions were compared to the imagery condition to answer the question of whether learners can generate helpful visualizations on their own by envisioning the contents of the worked-out examples. Within the last two steps of analysis the respective five conditions (i.e., frequent/sparse use of static pictures/animations, imagery) were also contrasted with the text-only condition to test whether using a specific type of external visualization either frequently or sparsely and whether instructing learners to imagine the worked-out examples was superior to presenting only text.

*Overall analyses.* Prior knowledge, intrinsic motivation, and spatial abilities were comparable across all four instructional conditions (prior knowledge: $F < 1$, intrinsic motivation: $F < 1$; spatial abilities: $F(3, 119) = 1.65$; $MSE = 51.22$; $p > .10$).

In a first step, problem-solving performance and learning time across all four conditions were analyzed by means of one-factor ANOVAs (Table 2). Performance on

condition ($F$(3, 119) = 2.32; *MSE* = 769.06; $p$ < .10), while performance on novel problems was left unaffected by the experimental manipulation ($F$ < 1). Learning time increased rather naturally with the more instructional materials being available for processing ($F$(3, 119) = 6.21; *MSE* = 48052.96; $p$ = .001). Furthermore, there were no differences between the four instructional conditions with regard to task demands, effort, and navigational demands experienced by learners (all *F*s < 1).

Thus, at first sight it seems that although presenting external visualizations increased the time learners devoted to learning, these increases were not accompanied by gains in performance. Contrarily, performance for isomorphic problems was worst in the animation condition. However, it is not clear whether the finding that problem-solving performance was only slightly affected by variations of the instructional materials is due to the ineffectiveness of these variations or whether it is due to the fact that learners did not sufficiently use the external visualizations. To address this question, the frequency by which learners retrieved these visualizations was analyzed in a second step.

*Effects of utilizing external visualizations.* For every participant it was possible to determine how often (s)he had retrieved visualizations. Though there were a total of 31 visualizations available in each of the conditions, these visualizations were seldom retrieved (see Figure 2). The figure displays the number of participants using a specific number of visualizations in each of the two conditions. For instance, nine participants in the animation condition and 11 participants in the static pictures condition used only one or two visualizations. On average, static pictures were viewed 7.0 times, animations 10.3 times. Thus, less than one quarter of the available static pictures and only one third of the animations were used for learning.

For the further course of the statistical analyses a median split within each of the two conditions was conducted to distinguish between learners who sparsely used external

condition the median for visualization retrieval was $Mdn = 3$, with 18 participants being classified as sparse users and 12 as frequent users. In the animation condition the median was $Mdn = 9$ with 17 participants classified as sparse users and 15 as frequent users of visualizations. The resulting variable was used as a second factor in an ANOVA (instructional condition x visualization utilization).

Analyzing performance for isomorphic problems in the two conditions with external visualizations (Table 3) by means of this ANOVA yielded a significant main effect of instructional condition in favor of learning from static pictures ($F(1, 58) = 5.56$; $MSE = 743.58$; $p < .05$) as well as a marginally significant main effect for the utilization of visualizations ($F(1, 58) = 3.08$; $MSE = 743.58$; $p < .10$). Moreover, there was a significant interaction ($F(1, 58) = 7.20$; $MSE = 743.58$; $p < .01$) indicating that – when frequently used – static pictures were superior to animations ($t(25) = 3.44$; $p < .01$), while there were no differences when the visualizations were sparsely used ($t(33) = -0.24$; $p > .80$).

There were no significant main effects for novel problems (instructional condition: $F < 1$; information utilization: $F(1, 58) = 1.30$; $MSE = 474.28$; $p > .20$) nor interactions ($F(1, 58) = 1.12$; $MSE = 474.28$; $p > .20$). Learning times were longer for students in the animation condition than in the static pictures condition ($F(1, 58) = 7.97$; $MSE = 44333.61$; $p < .01$) and increased dramatically with frequent use of visualizations ($F(1, 58) = 18.02$; $MSE = 44333.61$; $p < .001$). Moreover, a significant interaction ($F(1, 58) = 8.78$; $MSE = 44333.61$; $p < .01$) revealed that frequently using static pictures compared to using them sparsely had no impact on learning time ($t(33) = 0.11$; $p > .90$), while frequent use of animations almost doubled the time spent with the instructional materials ($t(25) = -3.95$; $p = .001$).

Interestingly, learners did not experience any differences on cognitive load during learning. That is, the task demands, amount of effort, and navigational demands were judged

behavior (all $F$s < 1 with one exception: instructional condition x visualization utilization for

task demands: $F(1, 58) = 1.05$; $MSE = 545.14$; $p > .30$).

In a final step the problem-solving performance, learning times, and cognitive load of

learners in the four visualization groups (Table 3) were compared to those of learners who

had only text at their disposal (i.e., the baseline, Table 2) by means of multiple $t$-tests.

Frequently using static pictures was clearly associated with a better performance on

isomorphic problems compared to learning from text only ($t(39) = 2.27$; $p < .05$), while there

were no differences for novel problems ($t(39) = -0.60$; $p > .50$). At the same time, this

frequent use did not require more learning time ($t(39) = 1.23$; $p > .20$) nor did it coincide with

increases in cognitive load experienced by learners (task demands: $t(39) = -0.58$; $p > .50$;

effort: $t(39) = -0.22$; $p > .80$; and navigational demands: $t(39) = 1.23$; $p > .20$).

Whenever static pictures were only sparsely used there were no effects on problem-

solving performance (isomorphic problems: $t(45) = -1.42$; $p > .15$; novel problems:

$t(45) = -0.57$; $p > .50$), learning time $t(45) = 0.15$; $p > .80$, or cognitive load (task demands:

$t(45) = 0.18$; $p > .80$; effort: $t(45) = 0.60$; $p > .50$; and navigational demands: $t(44) = -0.12$;

$p > .90$).

Frequently using animations was associated with a significant decrease in performance

for isomorphic problems compared to learning from text only ($t(42) = -2.11$; $p < .05$). There

were no differences for novel problems ($t(42) = -1.09$; $p > .20$). Not astonishingly, a frequent

use of animations increased the time for learning ($t(42) = 5.61$; $p < .001$). No significant

differences could be detected for the three cognitive load measures (task demands:

$t(42) = 0.47$; $p > .60$; effort: $t(42) = 0.90$; $p > .30$; and navigational demands: $t(42) = 1.06$;

$p > .20$).

Sparsely using animations was not related to any changes in performance (isomorphic

problems: $t(44) = -1.13$; $p > .20$; novel problems: ($t(44) = 0.62$; $p > .50$) or in learning time

differences with regard to cognitive load (task demands: $t(44) = -0.58$; $p > .50$; effort:

$t(44) = -0.22$; $p > .80$; and navigational demands: $t(44) = 1.23$; $p > .20$).

To sum up, using static pictures frequently was superior to learning from animations in

that it was associated with a better performance for isomorphic problems while at the same

time requiring less learning time. Based on the comparisons to the text baseline, static pictures

proved to be a helpful augmentation of verbal materials. On the contrary, frequently using

animations led to substantial increases in learning time while at the same time an inferior

performance as compared to learning from text. Sparse users of external visualizations

produced the same outcomes as learners who did not have visualizations at their disposal.

Therefore, whereas concrete visualizations of the contents of the worked-out examples

supported learning, the dynamics contained in the animations were unnecessary or even

harmful. If complex dynamic visualizations cannot be recommended for learning, the

question arises whether there is a need for external visualizations at all. That is, if simple

visualizations like static pictures of the content help to achieve an understanding of the

principles, maybe learners are able to generate these images by themselves. The question

whether external visualizations are more helpful than the instruction to imagine the contents

of the worked-out examples is addressed in the remainder of this article.

*The benefits of imagery compared to external visualizations.* In this final analysis, the

imagery condition was compared to the two external-visualization conditions by means of

multiple *t*-tests whereby in the latter two sparse and frequent users of visualizations were

distinguished.

The imagery condition had no impact on any of the performance measures when

compared to the group of learners who had frequently used static pictures (isomorphic

problems: $t(42) = 1.42$; $p > .10$; novel problems: $t(42) = -0.78$; $p > .40$). However, learning

from static pictures took more time than imagining the contents of the examples ($t(42) = 2.30$;

demands: $t(42) = -1.30; p > .20$; effort: $t(42) = 0.41; p > .60$; and navigational demands: $t(42) = 0.34; p > .70$).

Instructing learners to imagine the contents of the worked-out examples improved performance on isomorphic problems slightly, though not significantly, compared to a sparse use of static pictures ($t(48) = -1.96; p < .10$), while there were no differences for novel test problems ($t(48) = -0.76; p > .40$). There were no differences with regard to learning time between the two groups ($t(48) = 0.88; p > .30$) nor with regard to any of the cognitive load measures (task demands: $t(48) = -0.87; p > .30$; effort: $t(48) = 0.18; p > .80$; and navigational demands: $t(48) = -0.48; p > .60$).

Mental animation was thus more effective and efficient compared to a sparse use of static pictures in that it slightly improved performance without requiring more time. Moreover, it was more efficient than frequent use of static pictures as it led to the same problem-solving performance with lower time demands. It can be expected that this pattern of results will be even more pronounced when comparing the mental animation condition to the groups of learners using external animations.

Learners in the mental animation condition clearly outperformed learners who had frequently used external animations with regard to isomorphic problems ($t(45) = -2.51$; $p < .05$), while the differences for novel problems were not pronounced ($t(45) = -1.34$; $p > .10$). Moreover, frequently using external animations led to prolonged learning times ($t(45) = 7.32; p < .001$). Again, none of these differences was reflected in learners' judgments of cognitive load (task demands: $t(45) = -0.62; p > .50$; effort: $t(45) = 0.59; p > .50$; and navigational demands: $t(45) = 0.61; p > .50$).

Finally, imagery was slightly superior to sparse use of dynamic visualizations (isomorphic problems: $t(47) = -1.70; p < .10$; novel problems: $t(47) = 0.57; p > .50$), while requiring the same amount of learning time ($t(47) = 0.71; p > .40$). There were no differences

for the cognitive load measures (task demands: $t(47) = -1.64$; $p > .10$; effort: $t(47) = -0.76$; $p > .40$; and navigational demands: $t(43) = 0.77$; $p > .40$).

Thus, imagining the contents of the examples was more effective than learning from dynamic visualizations and more efficient in that it improved performance while requiring equal or even less amounts of learning time.

Finally comparing mental animation to the text-only condition yielded no significant differences with regard to performance (isomorphic problems: $t(59) = -0.88$; $p > .30$; novel problems: $t(59) = -0.14$; $p > .80$), learning time ($t(59) = 0.79$; $p > .40$), or cognitive load (task demands: $t(59) = -1.21$; $p > .20$; effort: $t(59) = -0.57$; $p > .50$; and navigational demands: $t(59) = -0.44$; $p > .60$).

## Discussion

Evidence for the differential effectiveness of external and internal visualizations for cognitive skill acquisition has been presented here. The results support the assumption that learners might benefit from a concrete visualization of problem states tied to worked-out examples. Interestingly, generating these visualizations by oneself in the imagery condition produced learning outcomes similar to those in the "best" external visualization condition (i.e., frequent use of static pictures) and was more efficient than any of the external visualization conditions by requiring less learning time. Unfortunately, the effects of imagery were not strong enough to yield significant differences with the text-only condition. Before addressing possible reasons for the latter result, the differential effectiveness of external visualizations observed in this experiment are first discussed.

While it was demonstrated that a frequent use of static pictures was associated with performance improvements on isomorphic problems, frequently retrieving animations coincided with the opposite effect in that it was related to a significant decrease in problem-solving performance. The sparse use of external visualizations was no different to not having

strategic variables (i.e., frequency of using different types of external visualizations) is important, as these variables may moderate the relationship between design variables and learning outcomes (Gerjets & Scheiter, 2003).

The initial idea behind using animations had been that the dynamics of an animation might be helpful for illustrating information on changes in problem states that occur due to applying a specific solution-step. However, it seems that explicitly representing these changes was more harmful than helpful in that it may have distracted learners. These results are in line with prior findings (Lowe, 1999; Pane et al., 1996).

Interestingly, in contrast to results of prior experiments (Gerjets et al., 2004), no differences in cognitive load as measured with the modified version of the NASA-TLX could be detected. Rather than doubting its sensitivity in the current context, two factors which have effects in opposite directions may have contributed to this result. On the one hand, based on the findings by Lowe (1999) or Pane et al. (1996) one might assume that dynamic visualizations should be associated with higher navigational demands and effort ratings. Salomon's findings (1983, 1984), on the other hand, suggest that the amount of invested mental effort is influenced by the learners' perception of the medium and the demands it imposes. Viewing an animation – similar to watching TV as in the Salomon study (1984) - is often perceived as being "easy" and thus learners may decide to refrain from a deeper processing of the presented information. Accordingly, from a prescriptive point of view learners would need to invest more resources into learning from animations; however, from a descriptive point of view they might not experience this need and may refrain from performing effortful cognitive processes required for a deeper understanding (Schnotz et al., 1999). Accordingly, it does not necessarily have to be the case that learners report higher cognitive load values for the animation condition.

This research, thus, pursues the idea of further reducing the cognitive demands

visualizations. A reduction of cognitive load is aimed at by using abstract rather than concrete dynamic visualizations of the worked-out examples. These abstract animations are characterized by the fact that the visualizations of all examples share a common representation of objects and of the relevant relations among them. That is, irrespective of an example's cover story, objects such as the sprinters are represented by marbles selected from an urn. This common representation across examples should help learners to focus on the structural similarities and differences between the examples while being able to ignore their surface features. Thus, there is not only less information that needs to be processed in total, the ratio between mathematically relevant and irrelevant information is improved compared to the concrete animations investigated in the current paper. Thus, this simplified representation should be less demanding and should leave cognitive resources free in order to cope with the dynamics of the visualization.

At the same time, instructional methods which reduce the danger of a superficial processing of the presented animations and which force learners to engage in a deeper processing of the instructional materials and invest more mental effort in the learning task will be investigated. Until now, the animations have been non-interactive; that is, learners had no opportunity to stop or replay the visualizations, which may enhance the impression of an ongoing movie whose presentation is predetermined. Accordingly, it is planned to implement interactive animations in a subsequent experiment. Interactivity seems to be a value in its own right and produces positive results rather consistently (Betrancourt & Tversky, 2000; Ferguson & Hegarty, 1995; Gonzales, 1996; Mayer & Chandler, 2001; Tversky, Bauer Morrison, & Betrancourt, 2002). In particular, it affords learners in thinking about whether they have grasped the information presented or whether they should replay the animation.

Additionally, the work with regard to cognitive skill acquisition from static pictures and imagery will be continued. First, research will be carried out on ways of improving the

guide learners to use static pictures more often will be tested. Prior studies demonstrated that prompting learners to retrieve profitable information units is an effective means to foster problem-solving performance, especially for students with low prior knowledge (Gerjets, Scheiter, & Schuh, in press). Second, once static pictures have been retrieved, their processing should be supported by additional instructional guidance. In particular, learners may receive *instructions to compare multiple static pictures* to enable them to more easily discover changes in problem states that have occurred due to applying a solution-step.

Finally, the use of imagery might be further improved. In the current experiment, learners were instructed to use imagery only at the beginning of the experiment and the reminder they received during the learning phase could be easily overlooked. This may also explain why giving an imagery instruction neither extended learning times nor increased effort ratings compared to the text-only condition, which one might expect if learners are asked to envision an example's content in addition to merely reading the text. The argument is similar to the one made with regard to the observed lack of differences for cognitive load ratings between the animation condition and the other instructional conditions. That is, if learners in the imagery condition had invested the amount of mental effort required to benefit from imagining worked-out examples, then this should have expressed itself in longer learning times and higher cognitive load ratings. The fact that these differences with the text-only condition were not observed indicates that learners did not engage in the imagination task to a sufficient degree. In order to circumvent this problem, computer-based prompts that frequently remind learners to envision the examples' contents are to be investigated in future experiments.

<div align="center">References</div>

Atkinson, R. K., Derry, S. A., Renkl, A., & Wortham, D. (2000). Learning from examples:

Instructional principles from the worked examples research. *Review of Educational*

Baek, Y. K., & Layne, B. H. (1988). Color, graphics, and animation in a computer-assisted learning tutorial lesson. *Journal of Computer-Based Instruction, 15*, 131-135.

Betrancourt, M., & Tversky, B. (2000). Effect of computer animation on users' performance: A review. *Le Travail Humain, 63*, 311-329.

Byrne, M. D., Catrambone, R., & Stasko, J. T. (1999). Evaluating animations as student aids in learning computer algorithms. *Computers & Education, 33*, 253-278.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review, 14*, 5-26.

Catrambone, R., & Seay, F. A. (2002). Using animations to help students learn computer algorithms. *Human Factors, 44*, 495-511.

Cooper, G., Tindall-Ford, S., Chandler, P., & Sweller, J. (2001). Learning by imagining. *Journal of Experimental Psychology: Applied, 7*, 68-82.

Ferguson, E. L., & Hegarty, M. (1995). Learning with real machines or diagrams: Application of knowledge to real-world problems. *Cognition and Instruction, 13*, 129-160.

Gerjets, P. & Scheiter, K. (2003). Goal configurations and processing strategies as moderators between instructional design and cognitive load: Evidence from hypertext-based instruction. *Educational Psychologist, 38*, 33-41.

Gerjets, P., Scheiter, K., & Catrambone, R. (2004). Designing instructional examples to reduce intrinsic cognitive load: Molar versus modular presentation of solution procedures. *Instructional Science, 32*, 33-58.

Gerjets, P., Scheiter, K., & Schuh, J. (2005). Instruktionale Unterstützung beim Fertigkeitserwerb aus Beispielen in hypertextbasierten Lernumgebungen [Instructional support for skill acquisition from examples in hypermedia-based learning environments]. *Zeitschrift für Pädagogische Psychologie, 19*, 25-38.

Ginns, P., Chandler, P., & Sweller, J. (2003). When imagining information is effective.

Gonzales, C. (1996). Does animation in user interfaces improve decision making? In R. Bilger, S. Guest & M. J. Tauber (Eds.), *Proceedings of the CHI 96 Conference on Human Factors in Computing Systems* (pp. 27-34). New York: ACM Press.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of experimental and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-183). Amsterdam: North Holland.

Hodes, C. L. (1992). The effectiveness of mental imagery and visual illustrations: A comparison of two instructional variables. *Journal of Research and Development in Education, 26*, 46-56.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65-99.

Levin, J. R., Anglin, G. L., & Carney, R. N (1987). Validating functions of pictures in prose. In D. M. Willows & H. A. Houghton (1987). *The psychology of illustration* (Vol. 1: Basic research, pp. 51-86). New York: Springer.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education, 14*, 225-244.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge, MA: Cambridge University Press.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*, 390-397.

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognition and Instruction, 9*, 329-389.

Palmiter, S., & Elkerton, J. (1993). Animated demonstrations for learning procedural computer-based tasks. *Human-Computer Interaction, 8*, 193-216.

Pane, J. F., Corbett, A. T., & John, B. E. (1996). Assessing dynamics in computer-based instruction. In M. J. Tauber (Ed.), *Proceedings of the ACM Conference on Human Factors in Computing Systems* (pp. 797-804). Vancouver, Canada: ACM.

Rieber, L. P. (1990). Animation in computer-based instruction. *Educational Technology Research & Development, 38*, 77-86.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*, 450-461.

Salomon, G. (1983). The differential investment of mental effort in learning from different sources. *Educational Psychologist, 18*, 42-50.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology, 76*, 647-658.

Schnotz, W., Boeckheler, J., & Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. *European Journal of Psychology of Education, 14*, 245-265.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction, 14*, 293 – 305.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review, 14*, 47-69.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Tversky, B., Bauer Morrison, J., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies, 57*, 247-262.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional

Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology*

   *Review, 14*, 261-312.

Author Note

Table 1

*Advantages (+) and Drawbacks (-) of Visualization Methods*

|  | Static pictures | Dynamic pictures | Imagery |
|---|---|---|---|
| Learner activity | + | - | + |
| Cognitive load | + | - | - |
| Correctness / completeness | - | + | - |
| Adaptation to learner preferences / knowledge | - | - | + |

Table 2

*Means and Standard Deviations for Learning Time (in Sec), Problem-Solving Performance (in % correct) and Cognitive Load as a Function of Instructional Condition*

|  | Instructional condition | | | |
|---|---|---|---|---|
|  | Text | Static pictures | Animation | Imagery |
| Performance: | 49.7 | 51.3 | 38.1 | 55.6 |
| Isomorphic problems | (23.1) | (33.5) | (24.0) | (29.1) |
| Performance: | 19.5 | 15.6 | 18.8 | 20.3 |
| Novel problems | (23.2) | (17.5) | (25.3) | (20.6) |
| Learning time | 493.6 | 530.8 | 678.3 | 457.4 |
|  | (204.9) | (184.5) | (302.3) | (152.2) |
| Task demands | 34.8 | 34.2 | 34.1 | 42.0 |
|  | (23.6) | (25.5) | (20.8) | (22.8) |
| Effort | 52.1 | 57.0 | 54.4 | 55.3 |
|  | (23.6) | (19.7) | (24.6) | (19.0) |
| Navigational demands | 16.6 | 18.3 | 24.2 | 19.1 |
|  | (19.6) | (24.0) | (24.0) | (24.2) |

Table 3

*Means and Standard Deviations for Learning Time (in Sec), Problem-Solving Performance*

*(in % correct) and Cognitive Load as a Function of Visualization Utilization in the Two*

*External Visualization Conditions*

| | Instructional condition | | | |
|---|---|---|---|---|
| | Static pictures | | Animation | |
| Visualization utilization | Sparse use | Frequent use | Sparse use | Frequent use |
| Performance: Isomorphic problems | 38.9 (28.7) | 70.0 (32.5) | 41.2 (26.9) | 34.7 (20.7) |
| Performance: Novel problems | 15.7 (20.2) | 15.3 (13.2) | 24.5 (30.7) | 12.2 (16.0) |
| Learning time | 503.1 (211.0) | 572.4 (133.3) | 495.4 (219.2) | 885.6 (244.5) |
| Task demands | 36.1 (23.4) | 31.3 (29.2) | 30.6 (24.1) | 38.0 (16.2) |
| Effort | 56.4 (21.1) | 57.9 (18.3) | 50.3 (26.8) | 59.0 (21.8) |
| Navigational demands | 15.8 (19.6) | 22.1 (30.1) | 24.7 (24.9) | 23.7 (23.8) |

Figure Captions

*Figure 1.* Screenshot of a worked-out example with external visualization

*Figure 2*: Number of participants using external visualizations and frequency of use in the

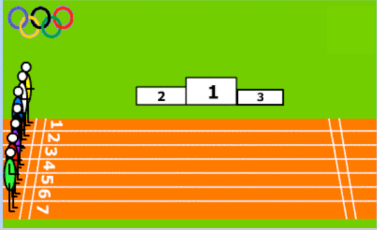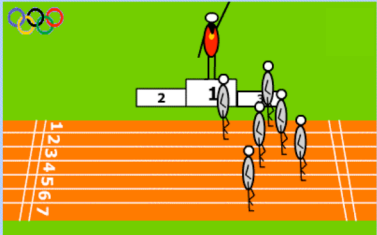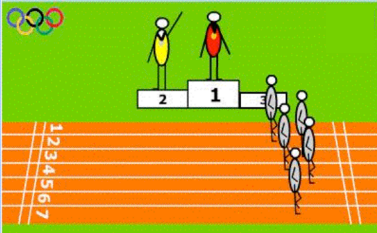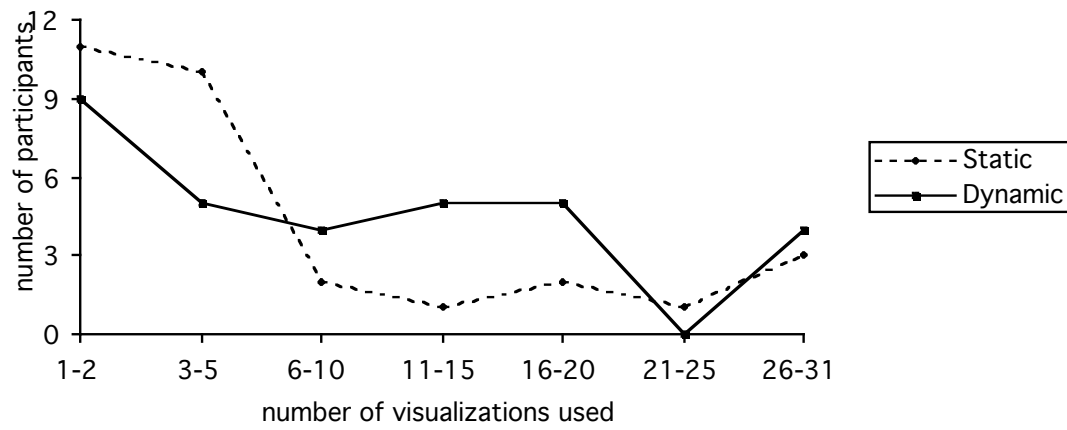static and in the dynamic visualization condition

Figure 1

**Problem: 100m-Sprint**

At the Olympics 7 sprinters participate in the 100m-sprint. What is the probability of correctly guessing the winner of the gold, the silver, and the bronze medals?

**Find 1st event probability:**

1/7

In order to find the first event probability you have to consider the number of acceptable choices and the pool of possible choices. The number of acceptable choices is 1 because only 1 sprinter can win the gold medal. The pool of possible choices is 7 because 7 sprinters participate in the 100m-sprint. Thus, the probability of correctly guessing the winner of the gold medal is 1/7.

**Find 2nd event probability:**

1/6

In order to find the second event probability you again have to consider the number of acceptable choices. The number of acceptable choices is still 1 because only 1 sprinter can win the silver medal. The pool of possible choices is reduced to 6 because only the remaining 6 sprinters participating in the sprint are eligible to receive the medal. Thus, the probability of correctly guessing the winner of the silver medal is 1/6.

Figure 2



Note: The number-of-visualizations variable was grouped into categories for facility of

inspection.