Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations?

Peter Gerjets, Katharina Scheiter, Richard Catrambone

Running head: ENHANCING LEARNING FROM MOLAR AND MODULAR EXAMPLES

Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction*, *16*, 104-121. This the author's version of the journal article, published in Learning and Instruction © Elsevier Ltd. 2006. doi:10.1016/j.learninstruc.2006.02.007

Can Learning from Molar and Modular Worked Examples be Enhanced by Providing Instructional Explanations and Prompting Self-Explanations?

Peter Gerjets

Knowledge Media Research Center, Tuebingen, Germany

Katharina Scheiter

University of Tuebingen, Tuebingen, Germany

Richard Catrambone

Georgia Institute of Technology, Atlanta, USA

Author Note:

The work reported in this paper was supported by the Alexander von Humboldt-Foundation (TransCoop-Program). Correspondence concerning this article should be addressed to Peter Gerjets, Knowledge Media Research Center, Konrad-Adenauer-Str. 40, 72072 Tuebingen, Germany. Electronic mail may be sent to p.gerjets@iwm-kmrc.de.

Abstract

In two experiments we explored how learning from traditional molar worked-out examples - focusing on problem categories and their associated overall solution procedures - as well as from more efficient modular worked-out examples - where intrinsic cognitive load is reduced by breaking down complex solutions into smaller meaningful solution elements - can be further enhanced. Instructional explanations or self-explanation prompts were administered to increase germane cognitive load. However, both interventions were not effective for learning and prompting for self-explanations even impaired learning with modular examples. In the latter case, prompting might have forced learners to process redundant information, which they had already sufficiently understood.

Can Learning from Molar and Modular Worked Examples be Enhanced by Providing Instructional Explanations and Prompting Self-Explanations?

It has often been argued that the most important prerequisite for successful problem solving probably consists of the availability of abstract *problem-type schemas* (Gick & Holyoak, 1983; Reed, 1993), that is, representations of problem categories together with category-specific solution procedures. Once a problem has been identified by means of its *structural problem features* as belonging to a known *problem category*, the relevant schema is retrieved from memory. This schema is then instantiated with the information that is specific to the to-be-solved problem. Finally the *category-specific solution procedure* attached to the schema is executed in order to produce a solution to the problem (cf. Derry, 1989; VanLehn, 1989). Schema-based problem solving is considered to be very efficient and therefore a substantial amount of research on skill acquisition has focused on the question of how such schemas can be conveyed.

As has been demonstrated repeatedly, *worked-out examples* (i.e., example problems illustrating problem categories together with step-by-step solutions) are an important instructional device for supporting the construction of problem-type schemas – particularly in the initial phases of skill acquisition (for an overview see Atkinson, Derry, Renkl, & Wortham, 2000). This ‘worked-example effect’ is usually explained by referring to Cognitive Load Theory and its distinction between intrinsic, germane, and extraneous components of the overall cognitive workload that arises during schema acquisition (Sweller, Van Merriënboer, & Paas, 1998). The number of elements that are to be integrated into a to-be-learned schema and therefore have to be processed in working memory simultaneously is referred to as *intrinsic cognitive load*. Intrinsic cognitive load depends on the relational complexity of the to-be-learned content (number of elements that are to be integrated into an schema) and the learner’s degree of domain-specific prior knowledge (i.e., schema availability). It is usually assumed in Cognitive Load Theory that intrinsic cognitive load cannot be altered by instructional design. It has, however, been argued that there are instructional-design manipulations that reduce intrinsic cognitive load (Gerjets, Scheiter, & Catrambone, 2004; Van Merriënboer, Kirschner, & Kester, 2003).

Beyond intrinsic cognitive load there might be additional cognitive load due to the nature of the instructional materials and the activities learners engage in. This load can be influenced by instructional design and can be categorized according to whether it is beneficial for schema construction (i.e., *germane cognitive load*) or not (i.e., *extraneous cognitive load*). When intrinsic task demands leave sufficient cognitive resources available, germane cognitive load might be induced by stimulating higher-level cognitive processes required for a deeper understanding of the materials; that is, processes that stimulate integrating the elements into a schema. Extraneous cognitive load is the result of implementing “instructional techniques that require students to engage in activities that are not directed at schema acquisition” (Sweller, 1994, p. 299). Extraneous cognitive load thus impedes learning. According to this line of reasoning, worked examples are an effective means for skill acquisition because they help to avoid extraneous cognitive load. As a result, working memory resources are freed-up that can be used to engage in elaborative (germane) processes of schema construction.

Unfortunately, it has often been observed that learners who study conventionally designed worked-examples usually do not spontaneously engage in these profitable elaboration processes (e.g., Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Gerjets & Scheiter, 2003; Gerjets, Scheiter, & Schuh, 2005; Gerjets, Scheiter, & Tack, 2000; Renkl, 1997). Accordingly, without additional instruction support learners often tend to experience serious difficulties in example-based learning resulting in the acquisition of rather shallow representations of problem categories and solution procedures.

One reason for this problem might be that conventionally designed worked-examples tend to teach problem categories and category-specific solution procedures in a *molar* way by treating problem categories and solution procedures as the basic units of analysis that cannot be broken down any further. These examples can easily overwhelm learners because they have to consider numerous structural problem features and solutions steps at once resulting in a high level of intrinsic cognitive load. To counteract this problem, we developed a *modular* example format that focuses on the role of individual structural problem features and individual solution steps below the category level (Catrambone, 1994; Gerjets, Scheiter, & Catrambone, 2004). By designing worked examples in a modular way we intended to reduce task-related, *intrinsic* load, thereby increasing the probability that learners would have free cognitive resources at their disposal for germane processes.

Molar versus Modular Examples: Reducing Intrinsic Cognitive Load

Molar examples focus on the information that is related to the main components of problem-type schemas, namely on information related to problem-category membership, structural task features, and category-specific solution procedures. They are often found in textbooks on knowledge-rich and well-structured domains like physics, mathematics, or programming. For instance, Atkinson, Catrambone, and Merrill (2003) have noted that mathematical problem solving is often characterized by ‘computationally-friendly’ molar solution approaches in which multiple solution steps are collapsed into a single formula that represents the solution procedure. These ‘recipe-like’ formulas allow one to easily calculate solutions by simply inserting the correct variable values. In order to enable learners to apply these ‘recipes,’ the examples demonstrate how to categorize problems by considering multiple (and often abstract) structural task features.

To demonstrate how conventional molar examples are usually designed, we will refer to the domain of probability word problems that was used in the two studies reported in this paper. Probability word problems deal with calculating the probability of individual and complex events. The probability of some individual event in a random experiment can be calculated by dividing the number of acceptable outcomes by the number of possible outcomes. A series of random experiments – where each random experiment consists of a selection process that yields one individual event – results in a complex event. How the probability of complex events is calculated, can be explained by either using a molar example format or a modular example format (for an illustration see Table 1).

According to the molar solution approach the probability of complex events is calculated by dividing the number of acceptable complex events by the number of possible complex events. Category-specific solution formulas are needed for calculating the number of possible complex events. In the materials used for experimentation we distinguish between four different problem categories (permutations and combinations, each with and without replacement) that differ with regard to two structural features. The first is whether the order in which elements are selected is important (yes: permutations; no: combinations); the second is whether selected elements are replaced after selection. The solution procedure illustrated by molar examples comprises four steps (see left column of Table 1), namely (1) identify task features, (2) apply formula, (3) insert values, and (4) calculate probability.

This solution approach is a convenient and fast way of calculating complex-event probabilities, in particular if the problems to be solved contain large numbers. However, it might have some serious limitations. For instance, formulas are usually restricted to solving a narrow range of problems that fall into predefined problem categories corresponding to the solution formula. Additionally, the strong focus on problem categories might cause learners to “memorize stereotypic solutions to problems based on their categorization” (Reed, 1999, p. 95) rather than thoughtfully considering the procedure. Finally, molar solution approaches are often characterized by a high level of relational complexity of the to-be-learned content, because they require the learner to attend to multiple structural task features and multiple solution steps at the same time. Thus, molar examples tend to impose a high level of intrinsic cognitive load, which might prevent learners from engaging in profitable example elaborations.

Therefore, we tried to find a way to reduce *intrinsic* cognitive load in example-based learning by abandoning the molar structure of conventionally designed worked-out examples and by presenting a more modular solution approach instead. This more modular solution procedure requires learners to keep only a limited number of elements active simultaneously in working memory. Modular examples avoid references to cognitively demanding molar concepts like problem categories, clusters of structural task features, and category-specific solution procedures. Solution procedures are broken down into smaller meaningful groups of solution steps that can be understood in isolation and that can be separately transferred when solving novel problems. Accordingly, the calculation of a complex event probability by means of a formula can be decomposed into a sequence of simpler calculations, that is, calculations of individual event probabilities. When calculating a particular individual event probability one has to take into account how the number of possible and acceptable choices change from the preceding to the current trial. These changes depend on whether previously selected objects are replaced or not after having been selected (i.e., *problem with or without replacement*), and on whether there is more than one acceptable choice in a given trial (i.e., *order of selection important or not*). Thus, the calculation of individual event probabilities allows one to directly relate individual structural task features and individual characteristics of solution steps without the need for prior problem categorization. This direct correspondence between structural features and solution steps makes it easier to adapt the modular approach to novel problems. The solution procedure we developed based on the modular approach is illustrated in the worked example in the right column of Table 1. In this example the probability of a complex event is calculated by determining the probabilities of all individual events that make up the complex event (steps 1 to 3) and then multiplying these individual event probabilities to calculate the overall probability (step 4).

The reasoning exemplified in the modular example format should help learners to understand relations below the category level that hold irrespective of category membership. What is even more important – in contrast to the category-based approach - is that this reasoning can be understood by holding only a rather limited amount of information in working memory simultaneously. Thus, this format should impose less intrinsic cognitive load than a molar example format and accordingly free cognitive resources that can then be used by learners to engage in example elaborations.

In line with these hypothesized advantages, experimental evidence was found for the superiority of modular as compared to molar examples with regard to problem-solving performance, learning time, and cognitive load (Catrambone, 1994; Gerjets, Scheiter, & Catrambone, 2004; Gerjets, Scheiter, & Kleinbeck, 2004). The positive effects of modular examples were found regardless of the number of problem categories taught, the type of learning task, and the learners' level of domain-specific prior knowledge. The results with regard to the lower levels of cognitive load imposed by modular examples seem to corroborate the idea that at least one of the advantages of modular examples is that solution procedures are broken down into smaller meaningful solution elements that can be understood in isolation without holding large amounts of information active in working memory.

Despite these very encouraging results, the data for transfer performance on novel test problems indicated that there still is room for improvement. Accordingly, we aimed at further enhancing learning from molar as well as from modular worked-out examples by providing additional instructional explanations and by prompting self-explanations.

Providing Instructional Explanations and Prompting Self-Explanations: Increasing Germane Cognitive Load

According to the instructional rationale of Cognitive Load Theory, reducing extraneous cognitive load (due to providing worked-out examples instead of practice problems) and reducing intrinsic cognitive load (due to providing modular instead of molar examples) can be seen as optimal preconditions for a subsequent stimulation of profitable (i.e., germane) example-elaboration strategies.

Example elaborations may in particular be helpful for fostering learners' skills in solving novel problems that do not fall into known problem categories and that require adapting learned procedures. Research on the role of example elaborations in learning has demonstrated that drawing inferences concerning the structure of example solutions, concerning the rationale behind solution procedures, and concerning the goals that are accomplished by individual solution steps (e.g., by relating example-specific information to more abstract information) are necessary ingredients of meaningful learning processes (Chi et al., 1989; Pirolli & Recker, 1994; Renkl, 1997).

To investigate whether stimulating processes of example elaboration might further improve transfer performance of learners provided with modular or molar examples, two experimental studies were conducted. In these studies it was tested whether providing instructional explanations (Experiment 1) or prompting self-explanations (Experiment 2) might support learners in their attempts to elaborate on instructional examples, thereby improving their transfer performance. In both approaches the focus lies on a type of example elaborations that Renkl (1997) describes as *principle-based self-explanations*. These explanations are characterized by the assignment of "meaning" to the individual steps in the solution procedure. Principle-based explanations try to answer the question of why a particular solution step is appropriate to solve the current problem, both in terms of the underlying principles that justify a step and in terms of the subgoal that the step aims to achieve in the solution procedure. This is similar to the process proposed in Catrambone's (1998) subgoal learning model, in which learners are hypothesized to be induced to self-explain the purpose of a set of steps when these steps are grouped together.

According to Renkl (1999, 2002) instructional explanations and self-explanations might be suitable in complementary situations as they can be characterized by a differential pattern of advantages and disadvantages during learning. Renkl postulates that – if appropriate – self-explanations should be more effective than instructional explanations, and that learners should therefore rely on self-explanations as much as possible. To justify this claim, Renkl argues that self-explanations are better adapted to the individual *prior knowledge* of learners, that the *timing* of self-explanations is superior because they occur when they can be integrated into ongoing cognitive activities, and that self-explanations profit from a *generation effect* according to which self-generated information is better remembered than presented information (cf. Lovett, 1992). At least three problems are, however, associated with self-explanations (Conati & VanLehn, 2000; Renkl, 1999, 2002): First, students often overestimate their understanding of the examples, that is, they may suffer from so called *illusions of understanding*, and thus refrain from self-explanation activities. Second, even if they have noticed gaps in their knowledge, they may not be able to generate explanations that are helpful to overcome those gaps. Third, learners may not necessarily generate explanations that are helpful for learning. That is, there may be limitations regarding the quality of self-explanations.

Different conclusions can be drawn from these findings: First, rather than expecting self-explanations to occur spontaneously, learners should be prompted to engage in self-explanatory activities. Second, in addition to prompts, scaffolds can be provided to ensure the availability of self-explanations that are suited to overcome knowledge gaps and that are of sufficiently high quality. These scaffolds may include (adaptive) feedback in case of incorrect self-explanations (Aleven & Koedinger, 2002; Conati & VanLehn, 2000) and the provision of instructional explanations to exploit their specific advantages (Renkl, 1999, 2002), particularly for learners with *very low knowledge prerequisites*. Instructional explanations are usually *correct* and may help learners to *overcome comprehension difficulties* when they cannot master these difficulties on their own. Finally, Renkl assumes that instructional explanations can show learners that they do not yet have a sufficient understanding of the content to be learned, thereby *avoiding illusions of understanding*.

However, scaffolding by means of feedback and instructional explanations is controversial (e.g., Conati & VanLehn, 2000; Schworm & Renkl, in press). For instance, Chi (2000) has argued that no feedback should be given in case of incorrect self-explanations, because it may obstruct self-explanatory activities triggered by discovering flaws in one's knowledge and aimed at fixing those flaws. Similarly, instructional explanations can even be harmful because they hinder learners in generating explanatory justifications of solution steps by themselves (Kulhavy, 1977). For instance, Schworm and Renkl (in press) demonstrated that instructional explanations reduced the students' self-explanation activities and thus learning outcomes. Similarly, Aleven and Koedinger (2000) have noted that students using their computer-based training tool asked directly for instructional help without attempting to generate an explanation by themselves. This has led several authors to conclude to allow for as much self-explanatory activities as possible and to provide only as many scaffolds as absolutely necessary (Conati & VanLehn, 2000; Renkl, 1999, 2002). One open question with regard to this issue is how elaborated and detailed instructional explanations should be in order to be more helpful than hindering.

Beyond prior knowledge, the availability of cognitive resources that can be devoted to generate self-explanations may affect whether self-explanations prompts or different types of instructional explanations are the more effective instructional method in a given situation. Because there are distinct differences regarding the level of intrinsic cognitive load imposed by either molar or modular worked examples, we expect differential effects with regard to whether learning from either example format can be further enhanced by providing instructional explanations or by prompting self-explanations.

It is important to note that we are interested in finding ways of improving learning from both example formats, because both solution approaches illustrated in these examples may contain distinct advantages. On the one hand, fostering understanding of molar examples may be important, because being able to apply category-specific solution procedures like formulas can – once the general rationale of probability problems is sufficiently understood – be very helpful for specific tasks. For instance, formulas provide a fast and convenient way to solve problems that contain large numbers. The applicability of the modular approach is rather limited in this case due to the large amount of error-prone computation necessary. On the other hand, the modular approach may help to foster the understanding of the general rationale of probability problems, which is a prerequisite for applying formulas. Thus, we believe that true problem-solving expertise can be conveyed only by a combination of modular examples that teach understanding and molar examples that allow for routine problem solving – together with the appropriate prompts and scaffolds for each example format.

With regard to the differential effectiveness of these prompts and scaffolds we expect that providing instructional explanations should be especially helpful for learners presented with molar examples. This should be the case because instructional explanations should support learners in coping with comprehension difficulties due to the high relational complexity of the molar solution approach. The modular solution approach, however, might be sufficiently easy to understand already, thus instructional explanations probably will not enhance learning any further.

On the other hand, prompting self-explanations should mainly be helpful if there are sufficient cognitive resources available for engaging in explanatory activities that are associated with germane cognitive load. Thus, self-explanation prompts should be more effective for instructional conditions where intrinsic cognitive load is low as it is the case for modular worked-out examples, but not for molar ones. Learners with molar examples might suffer from cognitive overload when trying to understand the complex solution procedures by themselves.

These assumptions can be derived from the basic instructional-design rationale of Cognitive Load Theory, according to which obstacles in instruction should first be removed (i.e., by reducing extraneous and intrinsic cognitive load) before higher-level cognitive processes can be productively stimulated, thereby fostering germane cognitive load. The role of instructional explanations and self-explanation prompts in learning from molar and modular examples was investigated in two experimental studies.

Experiment 1

Method

Participants

Ninety-six students (63 female, 33 male; mean age 24.64 years, *SD* = 6.28) from different majors at the University of Tuebingen in Germany participated in this experiment for either course credit or payment. While most students had been taught elementary concepts of probability theory in high school, only very few had dealt with calculating complex event probabilities in their university studies.

Materials and Procedure

A computer-based learning and problem-solving environment (HYPERCOMB) was used that teaches students how to calculate the probability of complex events. The environment consisted of a technical instruction on how to use the system, a short introduction to the domain, an example-based learning phase, and a subsequent test phase. Participants first completed a multiple-choice questionnaire with 11 questions on concepts and definitions that are important to know as a prerequisite for understanding how the calculation of the probability of complex events works. Next, participants were given a technical introduction to the HYPERCOMB system and to the experiment. After that, the general rationale behind calculating the probability of complex events was explained in a short introduction to the domain.

In the subsequent example-based learning phase, learners were told that they had to acquire knowledge on four different problem categories, where each category was explained by means of two worked-out examples. The worked-out examples, which had been constructed by the authors of the paper, consisted of a problem statement and a step-by-step solution procedure. The eight worked-out examples were presented in a linear order, where the sequence of problem categories was as follows: permutation without replacement - permutation with replacement - combination without replacement - combination with replacement. For instructional reasons, two different types of cover stories were used for the problem statements: The first worked-out example of each problem category always was a so-called urn example that dealt with selecting marbles from an urn. The second worked-out example of each problem category was a so-called daily-life example, which was related to situations that might occur in real life and which, therefore, was richer with regard to the cover story it was embedded in. Whereas the surface features were kept constant across problem categories for the urn examples, they varied across categories for the daily life examples, where each example was related to a different situation. These variations in cover stories were supposed to enable profitable processes of example comparison within and across problem categories (for an extensive discussion of the role of surface features for schema induction, see Quilici & Mayer, 1996; Scheiter & Gerjets, 2005; Scheiter, Gerjets, & Schuh, 2004). The numbers mentioned in the examples were of comparable size. Moreover, we aimed at keeping the complexity of the wording comparable across the different examples. Thus, no systematic differences in example difficulty can be expected. Participants could go back and forth between the examples by means of navigation buttons contained in the environment.

Depending on the experimental condition the solution procedure illustrated by these examples was either based on the molar or on the modular solution approach. Additionally, the amount of instructional explanations was varied across conditions. When a participant felt he or she had studied the examples for a sufficiently long time, he or she could start working on the test problems. The instructional materials were not available during problem solving.

Before solving the test problems, learners had to give an estimate of the cognitive load and related variables (i.e., feeling of success, stress) they had been experiencing during learning (see ‘dependent measures’ for details). Following this questionnaire, participants were instructed to solve nine probability problems for which the transfer distance was varied. The test problems were presented in a linear order from which participants could not deviate. Participants had a calculator available in order to ensure that their answers would not be inflated by calculation mistakes. The length of the experimental sessions depended on the time the participants took for studying the examples and solving the test problems.

Design and Dependent Measures.

As a first independent variable the solution approach illustrated by the worked-out examples was varied between subjects. The examples were either presented in the modular or the molar example format. As a second independent variable the amount of instructional explanations was varied between subjects resulting in three levels of elaboration (cf. Table 1). The instructional examples with highly-elaborated instructional explanations provided detailed justifications for why a solution step had been chosen. In an intermediate version facts concerning the solution steps were mentioned (e.g., individual-event probabilities, variable values), but no further justifications were given (medium-elaborated), while in a condensed version only the mathematical information was given without providing any verbal explanations at all (low-elaborated). The design of the explanations was based on a careful task analysis conducted by the authors to ensure that they provided information relevant to understanding each of the solution approaches. The highly-elaborated explanations illustrated the reasoning of a proficient problem solver when solving a problem.

For the resulting 2 x 3 design, 18 students were assigned to the modular approach/low-elaborated condition, 15 to the modular approach/medium-elaborated condition, 17 to the modular approach/highly-elaborated condition, 15 to the molar approach/low-elaborated condition, 17 to the molar approach/medium-elaborated condition, and 14 to the molar approach/highly-elaborated condition.

The dependent measures were learning time (in minutes spent in the example-based learning phase), frequency of example retrieval, cognitive load, problem-solving time (in minutes), and problem-solving performance for three isomorphic and six novel test problems. The frequency of example retrieval indicated how many examples a learner had accessed. The minimum value of this variable was eight, because there were eight examples available in each condition that had to be processed by learners. A value above eight indicated that the learner had retrieved examples multiple times.

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, Scheiter, & Catrambone, 2004). This measure consisted of three cognitive load items plus two additional items assessing feelings of success and experienced stress during learning. The cognitive load items were:

‘task demands’ (how much mental and physical activity was required to accomplish the *learning task*, 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 et al., 1998) cognitive demands are caused by inherent properties of the *learning task* (intrinsic cognitive load), higher-level processes for achieving a deeper *understanding* – as reflected in the availability of sophisticated and automated schemata (germane load), and activities not 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. In addition, learners had to rate how successful they felt in understanding the contents and how much stress they had experienced during learning. All items had to be rated on a scale from 0 (very low) to 100 (very high).

For each of the nine test problems, one point was assigned for a correct answer; no partial credit was given. Answers were also scored as correct if a student had set up the equation without mistakes, but had not determined the resulting probability as a final answer. The sums across the three isomorphic problems and across the six novel problems were each transformed into a percentage for ease of interpretation. 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, the outcomes of which 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)?

For the 11 items of the declarative pretest one point was assigned for every correct answer and the sums across all items were transformed into a percentage.

Results and Discussion

The means and standard deviations of all dependent measures are provided in Table 2. We report Cohen's *f* as a measure of effect size. According to Cohen (1988) an effect size of .10 corresponds to a small effect, .25 to a medium, and .40 to a large effect.

We first tested whether learners were comparable across the six experimental conditions with regard to prerequisite knowledge by means of an ANOVA (Solution Approach x Amount of Instructional Explanations). There were no significant main effects, nor an interaction between the two factors (all *F*s < 1).

The ANOVA (Solution Approach x Amount of Instructional Explanations) analysing the time learners spent on processing the worked-out examples revealed that learning with the molar approach took almost double the time than learning with the modular approach ($F(1, 90) = 48.24$, $MSE = 31.14$, $p < .001$, $f = .73$). Moreover, the learning time differed slightly as a function of the amount of instructional explanations provided ($F(1, 90) = 2.57$, $MSE = 31.14$, $p = .08$, $f = .24$). Posthoc Tukey tests revealed that learners took slightly less time to study low-elaborated examples than highly-elaborated examples ($p < .10$), while there were no time differences between medium- and low-elaborated examples ($p > .10$) or medium and highly-elaborated examples ($p > .90$). There was no interaction between the two factors ($F(2, 90) = 1.17$, $MSE = 31.14$, $p = .32$, $f = .16$).

The increase in learning time in the molar example conditions might go back at least partly to a more frequent retrieval of examples ($F(1, 90) = 8.81$, $MSE = 61.89$, $p = .004$, $f = .31$). In addition, the frequency of example retrieval depended on the amount of instructional explanations ($F(1, 90) = 5.20$, $MSE = 61.89$, $p = .007$, $f = .34$). Posthoc Tukey tests showed that retrieval frequency was greater in the low-elaborated conditions compared to the medium-elaborated example conditions ($p < .05$) and highly-elaborated conditions ($p < .05$), whereas there were no differences in example retrieval between medium- and highly-elaborated examples ($p > .90$). Again, no interaction could be observed ($F < 1$).

In a next step the time needed for problem solving, the problem-solving performance, and the cognitive load questionnaire data were analysed as a function of solution approach and amount of instructional explanations. There were no significant effects for problem-solving time (Solution Approach: $F(1, 90) = 1.98$, $MSE = 134.14$, $p = .16$, $f = .15$; Amount of Instructional Explanations/interaction: $F < 1$). With regard to problem-solving performance, participants solved more isomorphic problems correctly after they had studied modular examples ($F(1, 90) = 12.82$, $MSE = 1169.73$, $p = .001$, $f = .38$). There was no effect of providing different amounts of instructional explanations nor was there an interaction (both Fs < 1). The same pattern of results could be observed for novel problems: Students in the modular-example conditions solved more novel problems correctly than students in the molar-example conditions ($F(1, 90) = 20.28$, $MSE = 737.10$, $p < .001$, $f = .47$). Again, there were no effects of the amount of instructional explanations, nor an interaction effect (both Fs < 1).

The cognitive load data were in line with this clear superiority of a modular solution approach for acquiring problem-solving skills. Learners studying modular worked-out examples reported lower task demands (F(1, 90) = 9.53, MSE = 528.92, p = .003, f = .34), less effort (F(1, 90) = 9.10, MSE = 645.12, p = .003, f = .32), and, a bit surprisingly, lower navigational demands (F(1, 90) = 6.18, MSE = 373.32, p = .02, f = .26). This latter result might be due to learners in the modular conditions having spare working memory capacity available for navigational demands due to the reduced cognitive load imposed by these materials. Moreover, learners in the modular-example conditions rated themselves as being more successful in understanding the contents (F(1, 90) = 3.86, MSE = 420.93, p = .05, f = .21). The subjective cognitive load and feeling of success was also affected by the amount of instructional explanations (task demands: (F(1, 90) = 2.49, MSE = 528.92, p = .09, f = .23; effort: (F(1, 90) = 2.90, MSE = 645.12, p = .06, f = .25), and feeling of success: F(1, 90) = 6.92, MSE = 420.93, p = .002, f = .39. Posthoc Tukey tests showed that students rated learning from low-elaborated worked-out examples as being more demanding and effortful than studying highly-elaborated examples (task demands: $p < .10$ and effort $p < .10$). For both variables, there were no differences between low- and medium-elaborated examples (task demands: $p > .40$; effort: $p > .20$) or between medium- and highly-elaborated examples (task demands: $p > .50$; effort: $p > .70$). Additionally, they felt being more successful when learning from medium-elaborated or highly-elaborated than from low-elaborated examples (both ps < .05) – however, this subjective experience was not reflected in respective performance differences. There were no reliable differences between medium- and highly-elaborated examples ($p > .70$). There were no interactions for any of the variables (all Fs < 1). Finally, students' experience of stress during learning was not affected by any of the experimental manipulations (Solution Approach: (F(1, 90) = 2.37, MSE = 790.87, p = .13, f = .16; Amount of Instructional Explanations and interaction: both Fs < 1).

To conclude, we confirmed prior findings showing a clear superiority of a modular compared to a molar solution approach (Gerjets, Scheiter, & Catrambone, 2004). Learners who studied modular examples took less time for learning, retrieved fewer examples, solved more problems correctly, and reported less cognitive load, and a higher feeling of success. With regard to the amount of instructional explanations, however, no clear benefits could be observed for providing more instructional explanations concerning the rationale behind the solution steps of both solution approaches. In fact, more elaborated examples were less efficient than low-elaborated examples in that they took more time for studying while not being associated with better problem-solving performance. With regard to cognitive load, however, more elaborated examples reduced the subjective task demands and the required effort to understand the contents, and (erroneously) increased learners' subjective feeling of success during learning.

It thus seems that while students learn better from modular worked-out examples, both groups did not benefit from the instructional support that was provided in the medium and highly-elaborated conditions. The instructional explanations might have been superfluous for both solution approaches, but for different reasons. In line with our prior assumption, learners studying modular examples seemed to have sufficient cognitive resources at their disposal to engage in effective self-explanatory activities on their own. Although they invested more time in studying elaborated examples, this effort was obviously not necessary to ensure their ability to understand the rationale behind the solution approach. Learners with molar examples, on the other hand, possibly would have benefited from carefully processing the instructional explanations provided as predicted; however their learning time data revealed that they refrained from thoroughly doing so. It can be assumed that those students suffered from a well-known shortcoming of instructional explanations, namely the impression of having been able to produce the respective justifications of solutions steps by themselves. These feelings of understanding are frequently illusionary (Renkl, 1999, 2002). Our findings thereby resemble results obtained by Aleven and Koedinger (2000) and Kulhavy (1977) according to which instructional explanations may reduce the effort learners are willing to invest in understanding the instructional materials. Instructional explanations may thus also hinder self-explanatory activities (Schworm & Renkl, in press).

In Experiment 2 we wished to explore whether instructional materials could be further improved by prompting learners to provide detailed self-explanations for the solution steps. Thus, we replaced every second example by a medium-elaborated example combined with a prompt that required learners to provide more detailed self-explanations. Our main assumption was that prompting self-explanation activity should further improve learning for participants using modular examples, who – in line with the findings from Experiment 1 – are supposed to have sufficient cognitive resources available to generate these explanations by themselves. Moreover, based on cognitive-load considerations we had initially expected that students learning from molar examples would lack the necessary resources and thus would not benefit from self-explanation prompts. Based on the results of Experiment 1, however, one might also argue that self-explanation prompts for molar examples can be a helpful means to reduce possible illusions of understanding that might have been responsible for the ineffectiveness of instructional explanations. If students learning from molar examples were forced to generate explanations on their own, they might notice their lack of understanding and as a consequence might invest more time and effort in processing the examples. Thus, it is not clear whether self-explanations prompts for molar examples will be ineffective because of a lack of cognitive resources for the respective cognitive processes or whether they will be effective because they trigger a deeper processing of the instructional materials.

Experiment 2

Method

Participants

Ninety-one students (46 female, 45 male; mean age 19.27 years, *SD* = 1.36) from different majors at the Georgia Institute of Technology participated in this experiment for course credit. All students were familiar with basic concepts of probability theory.

Materials and Procedure

The example-based learning phase of HYPERCOMB was modified for the current experiment while all other aspects were the same as in Experiment 1. In this modified learning phase, the solution procedure was presented either by using the molar or the modular solution approach depending on the experimental condition. The solution procedures were explained to half of the participants by means of two worked-out examples for each of the four problem categories that contained highly-elaborated instructional explanations. These experimental conditions were identical to the respective conditions of Experiment 1. The other half of the participants received highly-elaborated urn examples for each problem category; however, for each of the daily-life examples they were prompted by means of questions contained in the examples to generate these elaborations by themselves. They received feedback for every answer given (for details see “design and dependent measures”).

Design and Dependent Measures

As a first independent variable the solution approach illustrated by the worked-out examples was varied between subjects. The examples were either presented in the modular or the molar solution approach. As a second independent variable it was manipulated whether all eight worked-out examples were either highly-elaborated or whether learners were prompted to generate these elaborations by themselves for every second example (self-explanation prompts). Sample materials for the prompting conditions are provided in Figure 1 (for molar examples) and Figure 2 (for modular examples).

In the conditions with prompts, the first worked-out example of each problem category (i.e., the urn example) always contained a highly-elaborated explanation of the solution procedure. For the second example (i.e., the daily-life example), however, the solution procedure was presented in the medium-elaborated version, that is, without further instructional explanations and justifications for why the solution step had been selected. For the first solution step, learners were then prompted (by means of a question that depended on the solution approach) to provide these elaborations by themselves (cf. Figure 3 for a detailed illustration of the prompting procedure). The answer had to be typed into a field and participants then clicked a "evaluate your answer" button. On clicking this button, a feedback page appeared which contained the learner's answer and an expert's answer provided by the system. This expert answer consisted of exactly the same information that had been given in the highly-elaborated examples. Learners were asked to compare their answer to this expert answer. Once they were done comparing the two answers, another navigation button allowed them to go back to the example they had been working on. The presentation of this example had, however, changed in two important ways compared to the first time learners had seen it. First, the expert's elaborations for the first solution step had been added to the description of this step, so that this step now was displayed in the highly-elaborated format. Second, learners were now prompted to go through the same process for the second solution step. That is, they now saw a question that asked for justifications with regard to the second solution step, typed in their answer to the question, received the expert's answer as feedback, were asked to compare their answer to this expert answer, and were led back to the example, which now contained the elaborations for the second solution step and a prompt for the third solution step (in case there was one). This cycle had to be done as many times as there were solution steps and could not be terminated before the worked-out example contained highly-elaborated explanations for all solution steps. Once this was the case, the urn example for the next problem category could be accessed, which was again presented as a highly-elaborated example.

The questions that were used to prompt learners to provide elaborations depended on the solution approach taken. In the modular approach the same question was used for every solution step, namely ‘Why is the numerator [VALUE OF NUMERATOR], and the denominator [VALUE OF DENOMINATOR] in this step?’ For the molar solution approach the questions were (1) ‘Why is this a [NAME OF CATEGORY] problem?’ (for the first solution step), (2) ‘What do n and k stand for in this example?’ (for the second solution step) and (3) ‘Why does n equal [VALUE OF N] and k equal [VALUE OF K]?’ (for the third solution step). In both solution approaches no questions were asked for the final solution step, in which the overall probability of the complex event and thus the final solution was presented.

In the resulting 2 x 2 design, 21 participants were assigned to the modular/highly-elaborated condition, 22 to the molar/highly-elaborated condition, 25 to the modular/prompts condition, and 23 to the molar/prompts condition.

The dependent measures were the same as in Experiment 1.

Results

The means and standard deviations of all dependent measures are provided in Table 3. With regard to the domain-specific prerequisite knowledge as assessed by the pretest, learners were comparable across instructional conditions as confirmed by a 2 x 2 ANOVA (Solution Approach: $F < 1$; Prompts: $F(1, 87) = 1.81$, $MSE = 325.30$, $p = .18$, $f = .14$; interaction: $F < 1$). The level of prerequisite knowledge in Experiment 2 was comparable to that in Experiment 1.

Similar to Experiment 1, learning time was almost doubled in the molar compared to the modular solution-approach conditions ($F(1, 87) = 30.02$, $MSE = 26.16$, $p < .001$, $f = .45$). Moreover, prompting learners for self-explanations increased learning time substantially ($F(1, 87) = 64.23$, $MSE = 26.16$, $p < .001$, $f = .75$). There was no interaction between the two factors ($F < 1$). The longer learning times in the molar-example conditions were not exclusively due to a more frequent retrieval of examples as an analysis of the frequency of example retrieval revealed ($F(1, 87) = 2.33$, $MSE = 49.58$, $p = .13$, $f = .15$). Interestingly, prompting learners for self-explanations reduced the frequency by which learners (re-)viewed examples compared to the highly-elaborated conditions ($F(1, 87) = 9.20$, $MSE = 49.58$, $p = .003$, $f = .32$). In addition, a marginally significant interaction was observed for the frequency of example retrieval ($F(1, 87) = 2.77$, $MSE = 49.58$, $p < .10$, $f = .17$).

Problem-solving times were longer in the molar compared to the modular solution approach ($F(1, 87) = 14.73$, $MSE = 54.99$, $p < .001$, $f = .40$). Moreover, extended problem-solving times were observable for students who had been prompted to generate self-explanations ($F(1, 87) = 6.47$, $MSE = 54.99$, $p < .01$, $f = .25$). There was no interaction between the two factors ($F < 1$).

Regarding problem-solving performance, students solved more isomorphic problems correctly after learning with modular examples ($F(1, 87) = 10.51$, $MSE = 1281.42$, $p = .002$, $f = .34$), while the availability of self-explanation prompts did not have any impact on performance for isomorphic problems ($F < 1$). There was no interaction between the two factors ($F(1, 87) = 1.61$, $MSE = 1281.42$, $p = .21$, $f = .13$). However, the percentage of correctly generated self-explanations correlated positively with performance for isomorphic problems in the condition with molar examples ($r = .52$, $p = .01$). The respective correlation was not significant in the condition with modular examples ($r = -.04$, $p > .80$).

There was the same pattern of main effects for solving novel problems (Solution Approach: $F(1, 87) = 4.98$, $MSE = 845.56$, $p = .03$, $f = .23$; Prompts: $F < 1$). In addition, there was a significant interaction ($F(1, 87) = 4.42$, $MSE = 845.56$, $p = .04$, $f = .22$) revealing that self-explanation prompts – compared to highly-elaborated examples – did not improve transfer performance when learning with molar examples ($t(43) = 0.94$, $p = .36$, $d = 0.28$), while prompts led to deteriorations in performance in the modular-example condition ($t(44) = -2.25$ $p = .03$, $d = 0.66$). To put it the other way around: modular examples were still superior to molar examples with regard to transfer performance on novel problems – but only in the conditions without self-explanation prompts. There was no significant correlation between the number of correctly generated self-explanations and transfer performance on novel problems (condition with molar examples: $r = .31$, $p > .15$; condition with modular examples: $r = .02$, $p > .90$).

The cognitive load data supported part of the problem-solving data, as modular examples were rated more favourably than molar examples (task demands: $F(1, 77) = 5.91$, $MSE = 437.97$, $p = .02$, $f = .28$; effort: $F(1, 77) = 5.55$, $MSE = 371.88$, $p = .02$, $f = .27$; feeling of success: $F(1, 77) = 5.56$, $MSE = 392.91$, $p = .02$, $f = .27$). There were no main effects of prompts or interactions for the aforementioned variables (all Fs < 1). No significant differences could be observed for navigational effort (Solution Approach: $F(1, 77) = 1.40$, $MSE = 230.68$, $p = .24$, $f = .13$; Prompts: $F(1, 77) = 1.08$, $MSE = 230.68$, $p = .30$, $f = .12$; interaction: $F < 1$). The results for the subjective stress ratings seemed suited to explain the findings obtained regarding problem-solving performance for novel problems: There were no main effects of the solution approach for this variable ($F < 1$), but prompts slightly increased feelings of stress ($F(1, 77) = 3.27$, $MSE = 403.03$, $p = .07$, $f = .20$). A marginally significant interaction ($F(1, 77) = 3.26$, $MSE = 403.03$, $p = .08$, $f = .20$), which corresponded to a medium effect, revealed that the stress learners' experienced was unaffected by self-explanation prompts in the molar condition ($t(39) = -0.02$, $p = .98$, $d = 0.01$). However, stress was increased by prompts in the modular conditions ($t(38) = 2.66$, $p = .01$, $d = 0.84$). Thus, prompts were perceived as stressful and reduced performance for novel problems in this condition.

General Discussion

In prior work we have successfully demonstrated that modular examples – where solution procedures are broken down in smaller meaningful steps – are superior to molar examples that convey knowledge on problem categories together with category-specific solution recipes (Gerjets, Scheiter, & Catrambone, 2004). This superiority was found for a variety of performance and cognitive load measures and proved to be robust across a variety of factors such as the number of problem categories taught and learner background. These findings were replicated in Experiment 1 and 2 where learners in the modular example conditions took less time for learning, showed better problem-solving performance for isomorphic as well as for novel test problems, and reported lower cognitive demands imposed by the learning materials.

Initially, we had assumed that instructional explanations might foster better understanding of molar examples. Contrary to this expectation the results of Experiment 1 indicated that instructional explanations did not improve performance for participants who studied molar or modular examples although learners had the erroneous impression of being more successful when studying examples with a higher amount of instructional explanations. Learners also reported lower task demands in the more elaborated conditions.

As a possible explanation we suggest that the instructional explanations were superfluous for both solution approaches, but for different reasons. As initially assumed, learners studying modular examples seemed to have sufficient cognitive resources at their disposal to engage in effective self-explanatory activities on their own and thus might not need instructional explanations. Learners with molar examples, however, refrained from thoroughly processing the elaborations possibly because they suffered from illusions of understanding (Renkl, 1999, 2002). While we had thought that these illusions of understanding might be overcome by prompting learners to generate these explanations on their own, Experiment 2 proved otherwise. Even worse, self-explanation prompts in the modular-example conditions led to a performance decrement. Initially we had expected that due to the low level of intrinsic cognitive load induced by this approach, learners might have sufficient cognitive resources available and might thus benefit from self-explanation prompts. It seems that while the assumption concerning their experienced cognitive load is true, the consequences drawn from that are not: Learners in the modular-examples condition were forced to generate self-explanations (and were given feedback for these explanations) for materials that they – according to the results of Experiment 1 – had already sufficiently understood. As Halabi and Tuovinen (2002, p. 36) have noted “feedback on well understood concepts is superfluous, or redundant, and usually interferes with learning”. Thus, our findings are consistent with the redundancy effect according to which repeatedly processing the same information might hinder rather than help learning (Kalyuga, Chandler, & Sweller, 1999). In line with this interpretation, learners provided with modular examples reported higher stress when prompted for self-explanations.

Moreover, there may have been problems with the design of the prompts and of the feedback (i.e., the instructional explanations) given to learners. For instance, Atkinson and Renkl (2001) did not observe any beneficial effects of instructional explanations, which were located on a different page and thus not integrated into the worked examples. Similarly, in our second experiment the feedback learners received for their self-explanations was presented on a separate page, because we wanted learners to notice that now a different activity than simply reading the example was expected of them. While we also integrated the feedback into the worked examples in a next step, we may have nevertheless caused a split-attention by this design. A possible alternative would be to present the feedback in a pop-up window on the same page as the example in a subsequent experiment.

A second design issue concerns the fact that learners were forced to type in their answers rather than select from a list of alternative answers. Unless the system is able to recognize natural language, no adaptive feedback can be provided to learners in case free responses are given to the prompts. Thus, learners in our case were forced to process the feedback irrespective of whether their self-explanation had been wrong or right. We had deliberately opted for this design compared to a multiple-choice format for at least three reasons. First, if learners are presented with wrong alternatives, they might remember these wrong answers later on without being able to tell whether they were wrong or not. Second, selecting from a predefined set of alternatives requires less cognitive activity than formulating one's own answer; thus, advantages of self-explanation prompts due to the generation effect (Lovett, 1992) might diminish. Third, in a prior experiment (Gerjets, Scheiter, & Schuh, 2005) we had not found any positive effects of a multiple-choice prompting method. However, a multiple-choice format has the distinct advantage of allowing for adaptive feedback depending on the answer given, which may explain the positive results achieved by this prompting method in the work by Atkinson, Renkl, and Merrill (2003) or Conati and VanLehn (2000). A multiple-choice format still seems to be a more effective method than refraining from giving any feedback to self-explanations presented in natural language, as has been done in a study by Hausmann and Chi (2002).

A third design issue refers to the number of prompts presented to learners. Participants in our experiment received 12 prompts in the molar condition, and 13 in the modular condition. In the current experiment we had opted to present prompts for all solution steps of each of the four examples. In both solution approaches, this may have caused some redundancy across examples as a very similar reasoning had to be applied to respond to the prompts irrespective of the problem category illustrated by a specific example. Moreover, for the modular solution approach there may have been redundancy within each example, as the reasoning for each solution step is very similar due to the reiterative structure of this approach. This may also explain that learners felt stressed (or annoyed) by having to respond to very similar prompts over and over again. An alternative to presenting prompts for all solution steps and examples would be to reduce the number of prompts in the example sequence. This can be done either by prompting only for explanations for one or two solution steps within each example or by using a fading approach, in which the number of prompts is increased or decreased within the example sequence (see for a similar approach to foster anticipative reasoning Renkl, Atkinson, & Maier, 2000; Renkl, Atkinson, Maier, & Staley, 2002). In order to overcome this problem of too many prompts, Stark (1999) has suggested that it would be best to allow learners to decide whether to produce self-explanations or not. However, in our own experiments (Gerjets et al., 2005; Scheiter, Gerjets, Catrambone, & Vollmann, in prep.) this learner control caused students to not use the prompts and the feedback option.

The fact that the instructional elaborations that were either provided by the system (Experiment 1) or were to be generated by students (Experiment 2) did not improve performance might also suggest that these explanations were of low quality and thus did not aid learning. However, we assume that the instructional explanations were helpful to students, but that students needed additional instructional support to cope with the challenges imposed by novel problems. Thus, students seem to be able to understand the rationale of solution procedures to an extent that allows them to solve isomorphic problems once the information relevant to solving the example problem is already available. For instance, when a learner is told which structural features are part of the problem, he or she seems to be able to explain why this problem is characterized by these structural features. This assumption is supported by the fact that in both solution approaches more than 80 percent of the self-explanations were correct. However, students might still be unable to identify the structural features or variable values by themselves. To use the terms introduced by Renkl (1997), learners might already be good principle-based explainers, but they might still lack the ability to anticipate solution steps.

According to this reasoning, a substantial decrease in intrinsic cognitive load might already be sufficient to allow students a basic understanding of the materials – so that the type of elaborations prompted in our experiments yields no further performance improvements. In this case, one has to guarantee that learners are prompted to apply cognitive processes that go beyond their initial understanding and are thus not redundant to what they already know. Thus, the question arises of whether the far transfer performance of students learning with modular examples can better be served by supporting their problem-solving abilities rather than their conceptual understanding. For instance, the provision of completion problems (Paas, 1992; Van Merriënboer, 1990) might be more appropriate for these students to improve their knowledge than eliciting principle-based self-explanations. Similarly, Stark (1999) "forced" learners to anticipate solution steps. Stark omitted text and inserted blanks into the solution steps of worked examples. The learners' task was to try to name what was missing. After attempting to fill in the blanks, the students received feedback on the correctness of their responses. Stark found that compared to studying complete examples, studying these incomplete examples fostered explanations and reduced ineffective self-explanations, such as rereading or paraphrasing. As a consequence, incomplete examples enhanced the transfer of learned solution methods.

This kind of reasoning has a variety of different consequences for further research: First, we need to investigate whether incomplete problems are a more appropriate means to foster transfer performance, because they elicit anticipative reasoning rather than principle-based explanations. Second, the question arises whether the effectiveness of the instructional methods investigated in this paper are moderated by prior knowledge. Renkl (1997) demonstrated that anticipative reasoners compared to principle-based explainers possessed a higher level of prior knowledge, which might be a prerequisite for anticipative reasoning. Accordingly, the instructional explanations and self-explanations might be more suited for learners with very low levels of prior knowledge, who lack the necessary conceptual understanding. On the other hand, completion problems might be more appropriate for learners with higher levels of prior knowledge. Third, if modular examples support the acquisition of some initial conceptual understanding, we need to investigate ways of combining the two solution approaches. In particular, it is worthwhile studying whether modular examples convey the conceptual understanding that the molar approach requires, before learners can benefit from the computational-friendly solution procedure enabled by the use of complex formulas.

References

Aleven, V., & Koedinger, K. R. (2000). Limitations of student control: Do students know when they need help? In G. Gauthier, C. Frasson, & K. VanLehn (Eds.), *Proceedings of the 5th International Conference on Intelligent Tutoring Systems* (pp. 292-303). Berlin: Springer.

Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science, 26*, 147-179.

Atkinson, R. K., Catrambone, R., & Merrill, M. M. (2003). Aiding transfer in statistics: Examining the use of conceptual equations and elaborations during subgoal learning. *Journal of Educational Psychology, 95*, 762-773.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. W. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70*, 181-214.

Atkinson, R. K., & Renkl, A. (2001). *The provision of instructional explanations in example-based learning: Analysis from a cognitive-load perspective*. Paper presented at the 9th EARLI conference. Fribourg (CH), August 2001.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Combining fading with prompting fosters learning. *Journal of Educational Psychology, 95,* 774-783.

Catrambone, R. (1994). Improving examples to improve transfer to novel problems. *Memory & Cognition, 22*, 606-615.

Catrambone, R. (1998). The subgoal learning model: Creating better examples to improve transfer to novel problems. *Journal of Experimental Psychology: General, 127*, 355-376.

Chi, M. T. H. (2000). Self-explaining expository tests: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161-238). Mahwah, NJ: Erlbaum.

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conati, C., & VanLehn, K. (2000). Toward computer-based support of meta-cognitive skills: a computational framework to coach self-explanation. *International Journal of Artificial Intelligence in Education, 11*, 398-415.

Derry, S. J. (1989). Strategy and expertise in solving word problems. In C. B. McCormick, G. Miller, & M. Pressley (Eds.), *Cognitive strategy research: From basic research to educational applications* (pp. 269-302). New York: Springer.

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., & Kleinbeck, S. (2004). Instructional examples in hypertext-based learning and problem solving: Comparing transformational and derivational approaches to example design. In H. M. Niegemann, R. Brünken, & D. Leutner (Eds.), *Instructional design for multimedia learning* (pp. 165-179). Muenster: Waxmann.

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*, 23-38.

Gerjets, P., Scheiter, K., & Tack, W. H. (2000). Resource-adaptive selection of strategies in learning from worked-out examples. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings from the 22nd Annual Conference from the Cognitive Science Society* (pp. 166-171). Mahwah, NJ: Erlbaum.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.

Halabi, A. K., & Tuovinen, J. E. (2002). Designing better CBL than face-to-face tutoring by incorporating teacher-like feedback and guidance in accounting. In A. McDougall, J. S. Murnane, & D. Chambers (Eds.), *Computers in Education 2001 - Australian Topics* (ACS Conferences in Research and Practice in Information Technology, Vol. 8, pp. 35-38). Sydney, New South Wales: Australian Computer Society.

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.

Hausmann, R. G. M., & Chi, M. T. H. (2002). Can a computer interface support self-explaining? *Cognitive Technology, 7,* 4-14.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*, 352-371 .

Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research, 47*, 211-232.

Lovett, M. C. (1992). Learning by problem solving versus by examples: the benefits of generating and receiving information. In *Proceedings of the 14th Annual Conference of the Cognitive Science Society* (pp. 956-961). Hillsdale, NJ: Erlbaum.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*, 429-434.

Pirolli, P., & Recker, M. (1994). Learning strategies and transfer in the domain of programming. *Cognition and Instruction, 12*, 235-275.

Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology, 88*, 144-161.

Reed, S. K. (1993). A schema-based theory of transfer. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 39-67). Norwood, NJ: Ablex.

Reed, S. K. (1999). *Word problems*. Mahwah, NJ: Erlbaum.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21*, 1-29.

Renkl, A. (1999). Learning mathematics from worked-out examples: Analyzing and fostering self-explanations. *European Journal of Psychology of Education, 14*, 477-488.

Renkl, A. (2002). Learning from worked-out examples: Instructional explanations supplement self-explanations. *Learning and Instruction, 12*, 529-556.

Renkl, A., Atkinson, R. K., & Maier, U. H. (2000). From studying examples to solving problems: Fading worked-out solution steps helps learning. In L. Gleitman & A. K. Joshi (Eds.), *Proceeding of the 22nd Annual Conference of the Cognitive Science Society* (p. 393-398). Mahwah, NJ: Erlbaum.

Renkl, A., Atkinson, R. K., Maier, U. H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *Journal of Experimental Education, 70*, 293-315.

Scheiter, K., & Gerjets, P. (2005). When less is sometimes more: Optimal learning conditions are required for schema acquisition from multiple examples. In B. G. Bara, L. Barsalou,

& M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 1943-1948). Mahwah, NJ: Erlbaum.

Scheiter, K., Gerjets, P., Catrambone, R., & Vollmann, B. (in prep.). *Strategies of information utilization in example-based hypermedia environments.*

Scheiter, K., Gerjets, P., & Schuh, J. (2004). The impact of example comparisons on schema acquisition: Do learners really need multiple examples? In Y. B. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera (Eds.), *Proceedings of the 6th International Conference of the Learning Sciences* (pp. 457-464). Mahwah, NJ: Erlbaum.

Schworm, S., & Renkl, A. (in press). Computer-supported example-based learning: When instructional explanations reduce self-explanations. *Computers & Education*.

Stark, R. (1999). *Lernen mit Lösungsbeispielen. Der Einfluß unvollständiger Lösungsschritte auf Beispielelaboration, Motivation und Lernerfolg* [Learning by worked-out examples. The impact of incomplete solution steps on example elaboration, motivation, and learning outcomes]. Bern, CH: Huber.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295-312.

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

VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 527-579). Cambridge, MA: MIT Press.

Van Merriënboer, J. J. G. (1990). Strategies for programming instruction in high school: Program completion vs. program generation. *Journal of Educational Computing Research 6*, 265-287.

Van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load of a learner's mind: Instructional design for complex learning. *Educational Psychologist, 38,*

5–13.

Table 1

Molar and Modular Example Formats Used for Experimentation

<table>
<tr><td colspan="2" align="center">100M-SPRINT EXAMPLE

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?</td></tr>
<tr><th>MOLAR EXAMPLE FORMAT</th><th>MODULAR EXAMPLE FORMAT</th></tr>
<tr><td>IDENTIFY TASK FEATURES
This problem is a permutation-without-replacement problem. Problems of this type have two important features: First, the order of selection is important, and second, there is no replacement of selected elements. We are not interested only in finding out just which 3 of the 7 sprinters win medals, rather we want to know specifically which sprinter wins which medal. Therefore, the order of selection matters. A sprinter can win at most only one medal. Thus, this problem is a problem without replacement. That is, after a sprinter wins a medal he is not eligible for being selected again.

APPLY FORMULA
For this type of problem the following formula should be applied: A = n!/(n-k)! with n being the number of all sprinters and k being the number of sprinters that have to be correctly guessed.

INSERT VALUES
In the given example there are 7 sprinters to choose from. This is the set of elements for selection (n = 7). As we want to find out the probability of correctly guessing the winner of the gold, the silver, and the bronze medals, 3 sprinters out of these 7 sprinters have to be selected. Therefore, the number of selected sprinters equals k = 3. Inserting these values into the formula for permutation without replacement yields 7! / (7- 3)! = 210 possible permutations.

CALCULATE PROBABILITY
In order to calculate the probability of correctly guessing the winner of each of the three medals, divide 1 (the particular permutation we are interested in) by the number of possible permutations. Thus, the probability of getting this permutation (the winner of each of the three medals) equals 1/210.</td>
<td>FIND 1ST EVENT PROBABILITY
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
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 100m-sprint are eligible to receive the silver medal. Thus, the probability of correctly guessing the winner of the silver medal is 1/6.

FIND 3RD EVENT PROBABILITY
In order to find the third 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 bronze medal. The pool of possible choices is reduced to 5 because only the remaining 5 sprinters participating in the 100m-sprint are eligible to receive the bronze medal. Thus, the probability of correctly guessing the winner of the bronze medal is 1/5.

CALCULATE THE OVERALL PROBABILITY
The overall probability is calculated by multiplying all individual event probabilities. Thus, the overall probability of correctly guessing the winner of each of the three medals is 1/7 * 1/6 * 1/5 = 1/210.</td></tr>
</table>

Note. In experimental conditions with highly-elaborated examples the solution procedure contained all information stated in the relevant table column. Conditions with low-elaborated examples contained only the mathematical information printed in bold. Conditions with medium-elaborated examples received a verbal description of the mathematical information but without further explanatory justifications.

Table 2

Results of Experiment 1: Means (and standard deviations)

Solution approach	**Molar**			**Modular**		
Level of elaboration	Low	Medium	High	Low	Medium	High
Pretest (% correct)	66.67	72.20	66.89	68.18	73.94	67.92
	(20.81)	(17.74)	(21.58)	(26.51)	(13.25)	(25.13)
Learning time (min)	13.53	17.44	17.13	7.24	7.10	9.93
	(4.90)	(9.52)	5.66)	(3.40)	(3.40)	(4.05)
Frequency of example retrieval	22.00	15.65	15.29	15.78	10.80	12.00
	(12.37)	(8.07)	(6.40)	(9.55)	(1.66)	(4.06)
Problem-solving time (min)	33.04	35.71	36.60	33.88	28.40	33.04
	(13.17)	(14.91)	(9.88)	(8.06)	(6.51)	(14.74)
Isomorphic problems (% correct)	44.44	43.14	33.33	64.81	66.67	64.71
	(37.09)	(36.83)	(26.15)	(35.19)	(33.33)	(34.30)
Novel problems (% correct)	14.44	14.71	9.52	37.96	36.67	39.22
	(23.46)	(21.96)	(19.30)	(31.73)	(26.87)	(34.33)

Task demands	64.33	62.35	55.36	56.39	42.00	40.00
	(15.22)	(19.69)	(24.92)	(21.88)	(24.63)	(28.99)
Effort	55.67	45.00	42.14	41.11	29.00	25.59
	(23.82)	(23.45)	(29.20)	(26.76)	(23.99)	(24.99)
Navigation demands	26.67	21.18	20.71	11.39	15.00	12.65
	(23.20)	(24.40)	(23.85)	(9.97)	(16.48)	(15.12)
Feeling of success	50.00	63.53	65.36	55.83	71.67	76.18
	(24.64)	(22.96)	(16.69)	(22.05)	(12.77)	(20.58)
Stress	39.67	38.83	44.29	37.50	28.67	30.00
	(30.50)	(27.59)	(26.37)	(26.36)	26.96)	(30.57)

Table 3

Results of Experiment 2: Means (and standard deviations)

Solution approach	**Molar**		**Modular**	
Self-explanation prompts	Without	With	Without	With
Pretest (% correct)	71.90	67.59	73.16	67.27
	(19.22)	(19.92)	(14.79)	(17.61)
Learning time (min)	9.50	19.13	4.63	12.23
	(5.64)	(6.31)	(2.31)	(5.10)
Frequency of example retrieval	15.86	8.91	11.14	9.12
	(13.11)	(2.52)	(4.36)	(2.77)
Problem-solving time (min)	25.57	30.28	20.33	23.55
	(7.73)	(10.03)	(5.16)	(5.71)
Isomorphic problems (% correct)	45.45	47.83	79.37	62.67
	(41.84)	(38.70)	(28.82)	(32.38)
Novel problems (% correct)	22.73	31.88	49.21	32.67
	(28.43)	(36.55)	(28.61)	(21.24)

Task demands	40.25	42.64	28.50	31.75
	(22.39)	(20.95)	(23.23)	(16.49)
Effort	35.25	35.95	25.00	26.00
	(23.71)	(16.63)	(19.53)	(16.51)
Navigation demands	13.25	15.00	7.50	12.75
	(20.66)	(17.32)	(10.07)	(9.52)
Feeling of success	76.75	68.81	81.50	76.00
	(24.20)	(22.13)	(16.79)	(14.47)
Stress	22.25	22.38	11.50	27.75
	(25.17)	(15.46)	(12.99)	(24.09)

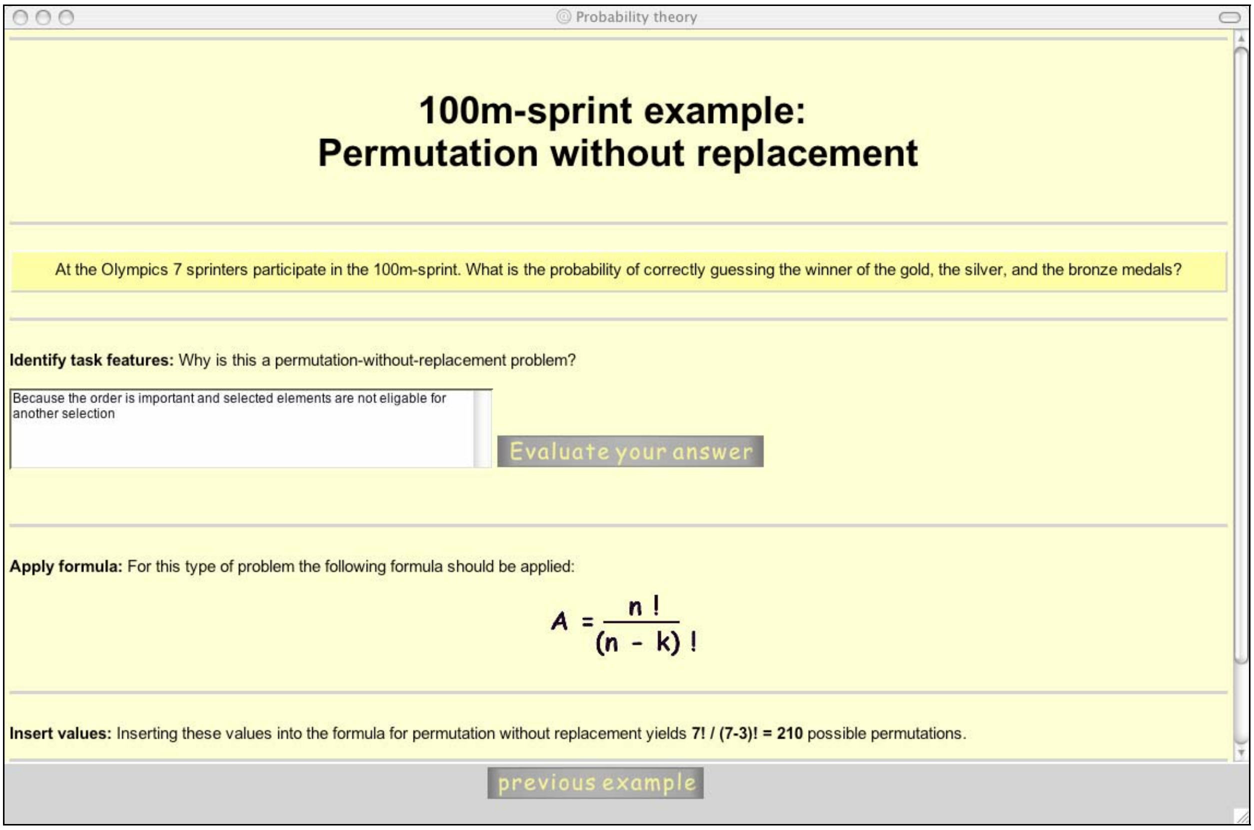
Figure Captions

Figure 1. Self-explanation prompts, feedback, and instructional explanations for molar examples (sample).

Figure 2. Self-explanation prompts, feedback, and instructional explanations for modular examples (sample).

Figure 3. Illustration of the prompting procedure.

Figure 1


Probability theory
100m-sprint example:
Permutation without replacement
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?
Identify task features: Why is this a permutation-without-replacement problem?
Because the order is important and selected elements are not eligable for another selection
Evaluate your answer
Apply formula: For this type of problem the following formula should be applied:
A = n! / (n - k)!
Insert values: Inserting these values into the formula for permutation without replacement yields 7! / (7-3)! = 210 possible permutations.
previous example

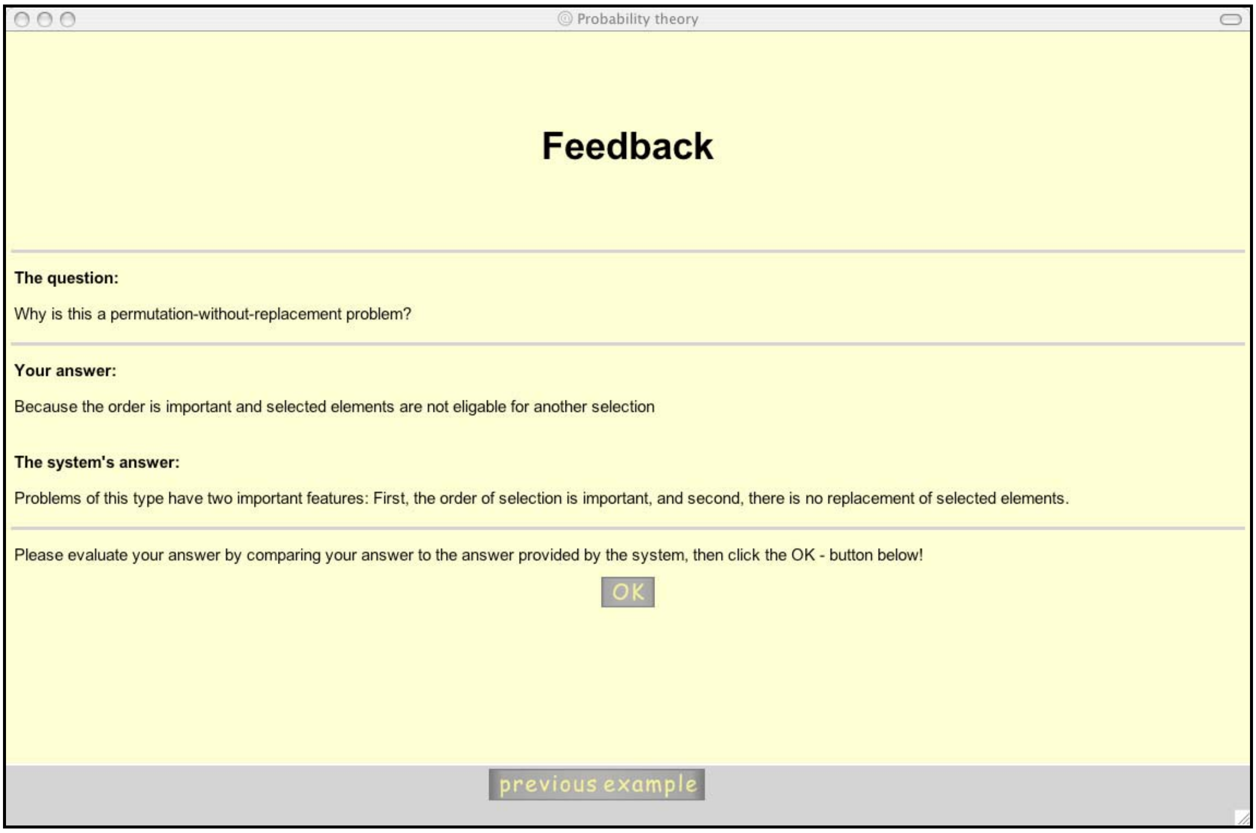

Probability theory
Feedback
The question:
Why is this a permutation-without-replacement problem?
Your answer:
Because the order is important and selected elements are not eligable for another selection
The system's answer:
Problems of this type have two important features: First, the order of selection is important, and second, there is no replacement of selected elements.
Please evaluate your answer by comparing your answer to the answer provided by the system, then click the OK - button below!
OK
previous example

Figure 1 (cont.)

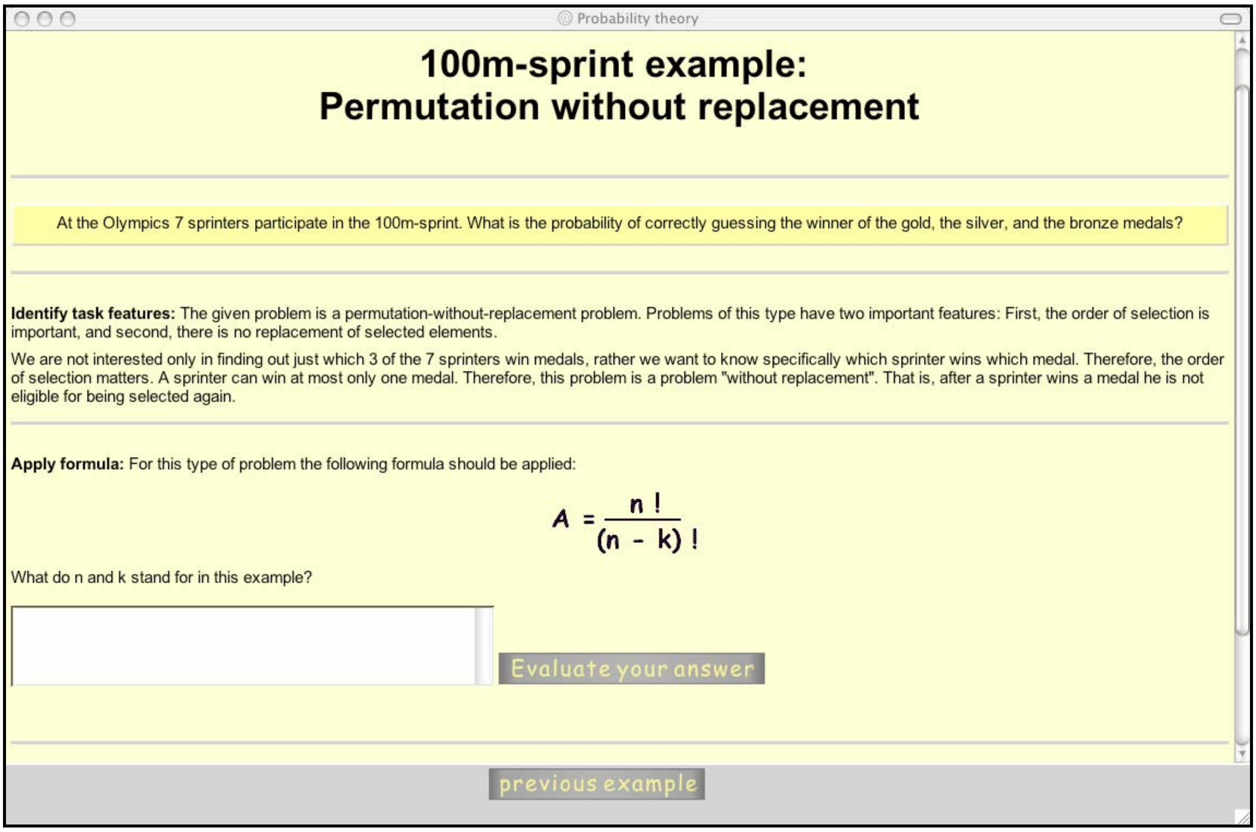
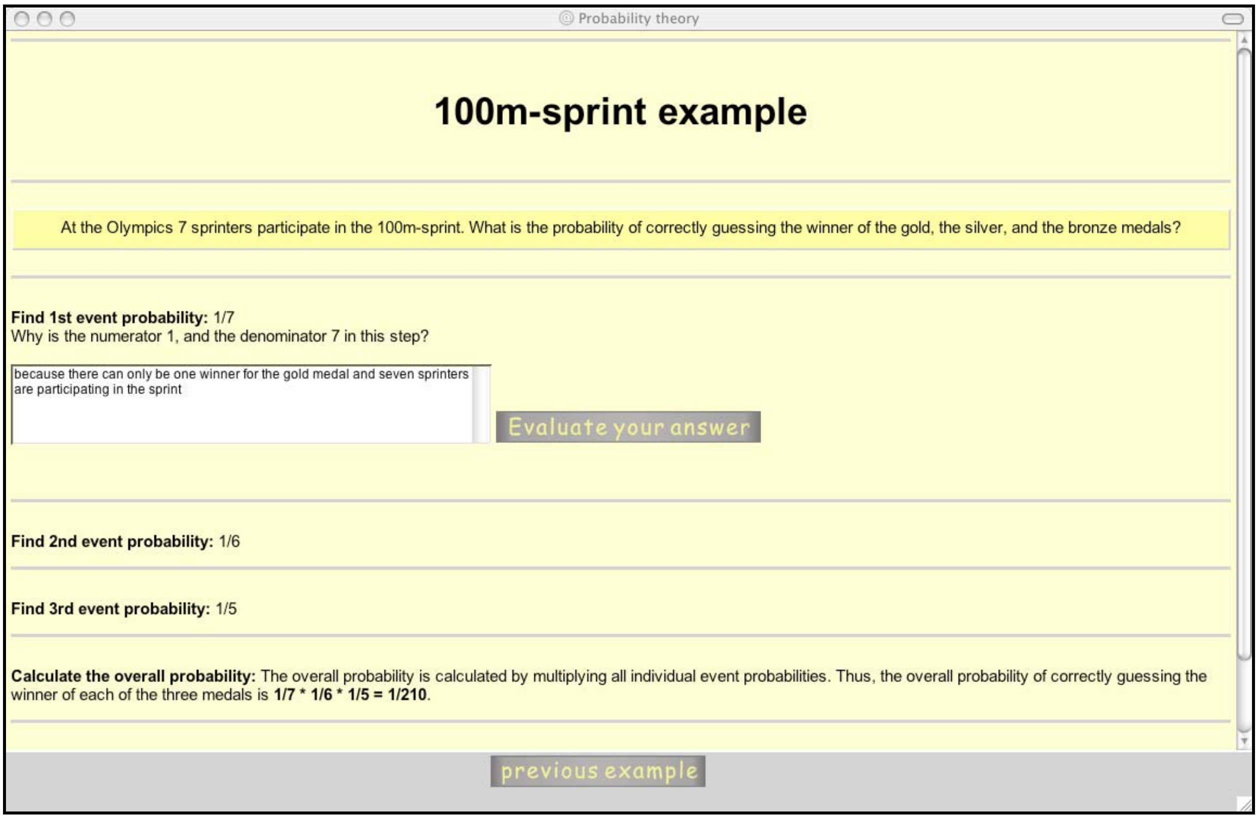


Figure 2


Probability theory
100m-sprint example
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
Why is the numerator 1, and the denominator 7 in this step?
because there can only be one winner for the gold medal and seven sprinters are participating in the sprint
Evaluate your answer
Find 2nd event probability: 1/6
Find 3rd event probability: 1/5
Calculate the overall probability: The overall probability is calculated by multiplying all individual event probabilities. Thus, the overall probability of correctly guessing the winner of each of the three medals is 1/7 * 1/6 * 1/5 = 1/210.
previous example

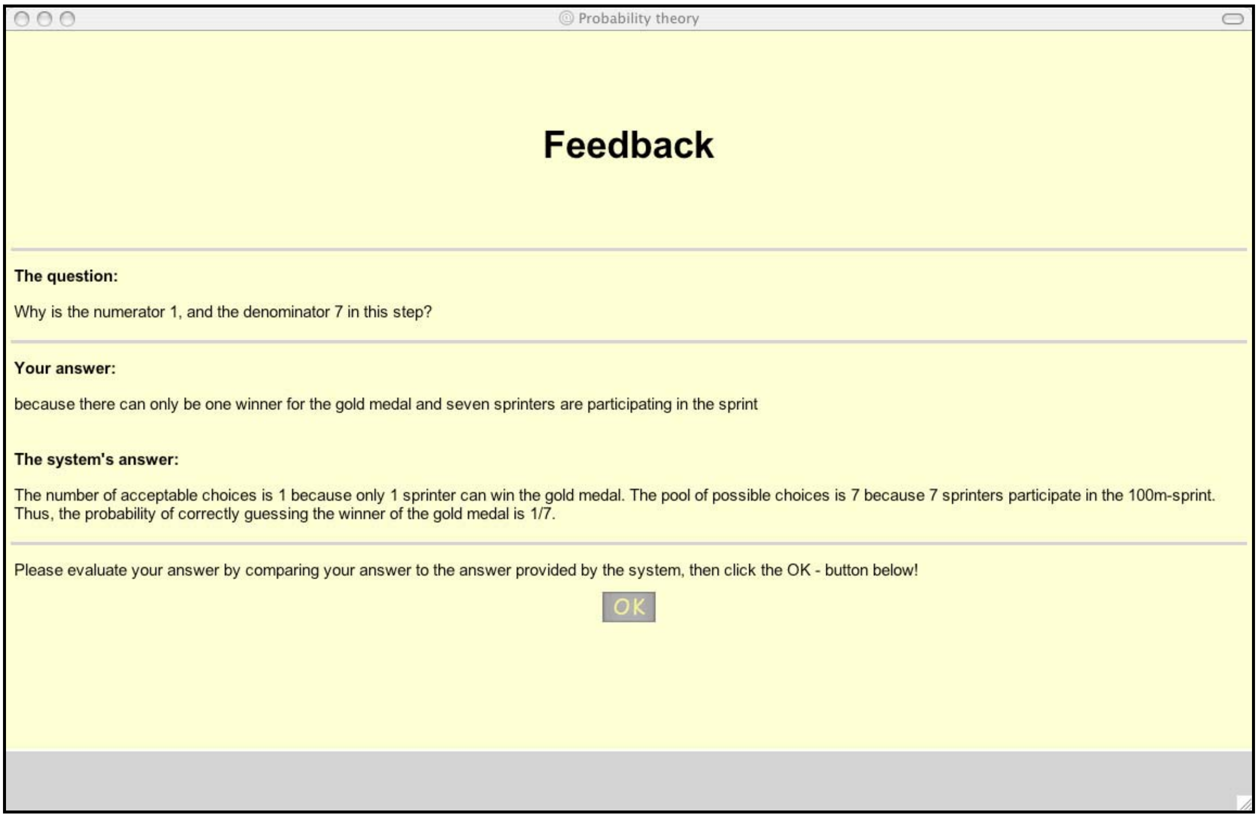

Probability theory
Feedback
The question:
Why is the numerator 1, and the denominator 7 in this step?
Your answer:
because there can only be one winner for the gold medal and seven sprinters are participating in the sprint
The system's answer:
The number of acceptable choices is 1 because only 1 sprinter 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.
Please evaluate your answer by comparing your answer to the answer provided by the system, then click the OK - button below!
OK

Figure 2 (cont.)

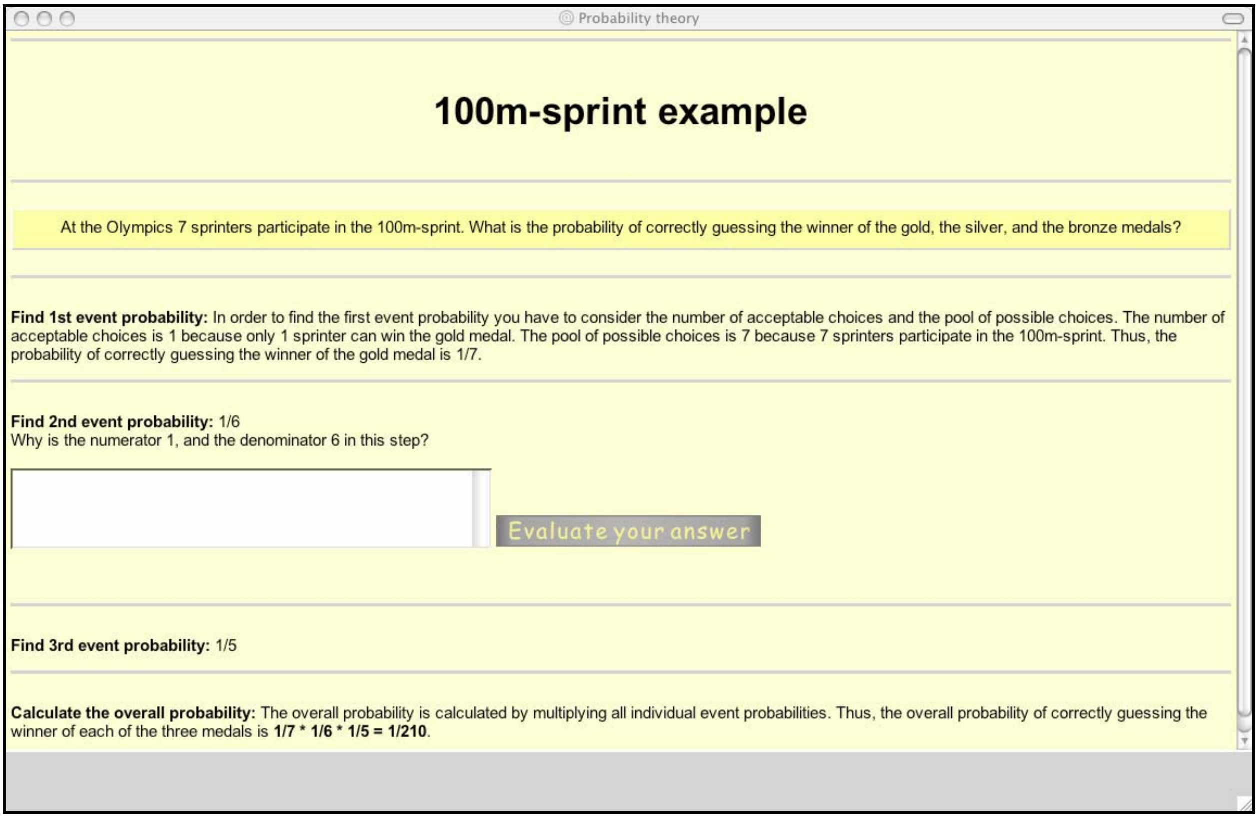
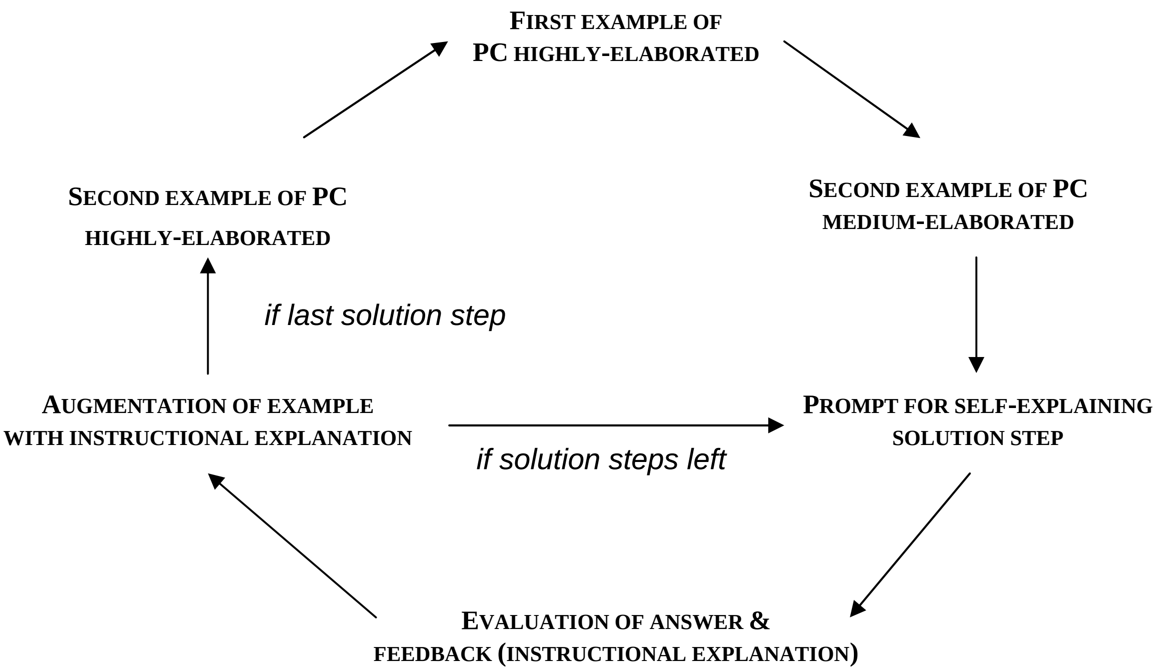


Figure 3


FIRST EXAMPLE OF
PC HIGHLY-ELABORATED
SECOND EXAMPLE OF PC
MEDIUM-ELABORATED
PROMPT FOR SELF-EXPLAINING
SOLUTION STEP
EVALUATION OF ANSWER &
FEEDBACK (INSTRUCTIONAL EXPLANATION)
AUGMENTATION OF EXAMPLE
WITH INSTRUCTIONAL EXPLANATION
if solution steps left
if last solution step
SECOND EXAMPLE OF PC
HIGHLY-ELABORATED