2-14-2018

Varying effects of subgoal labeled expository text in programming, chemistry, and statistics

Lauren Margulieux

Richard Catrambone

Laura M. Schaeffer

Varying Effects of Subgoal Labeled Expository Text in Programming, Chemistry, and Statistics

Lauren E. Margulieux[1], Richard Catrambone[2], and Laura M. Schaeffer[2]

[1]Georgia State University

[2]Georgia Institute of Technology

Keywords: worked examples; expository text; discipline based education research; STEM education; instructional design; subgoal learning

Author Note

Correspondence concerning this article should be addressed to Lauren Margulieux, Learning Technologies Division, Georgia State University, Atlanta, GA 30302-3978. Email: lmargulieux@gsu.edu. ORCID iD: 0000-0002-8800-2398.

Acknowledgments

This research was supported by the American Psychological Foundation's and Council of Graduate Departments of Psychology's Graduate Research Scholarship. The authors would like to thank John Sweller and Mark Guzdial for their feedback. We also thank Gerin Williams for her help collecting data.

Abstract

Originally intended as a replication study, this study discusses differences in problem solving performance among different domains caused by the same instructional intervention. The learning sciences acknowledges similarities in the learners' cognitive architecture that allow interventions to apply across domains, but it also argues that each domain has characteristics that might affect how interventions impact learning. The present study uses an instructional design technique that had previously improved learners' problem solving performance in programming: subgoal labeled expository text and subgoal labeled worked examples. It intended to replicate this effect for solving problems in statistics and chemistry. However, each of the experiments in the three domains had a different pattern of results for problem solving performance. While the subgoal labeled worked example consistently improved performance, the subgoal labeled expository text, which interacted with subgoal labeled worked examples in programming, had an additive effect with subgoal labeled worked examples in chemistry and no effect in statistics. Differences in patterns of results are believed to be due to complexity of the content to be learned, especially in terms of mapping problem solving procedures to solving problems, and the familiarity of tools used to solve problems in the domain. Subgoal labeled expository text was effective only when students learned more complex content and used unfamiliar problem solving tools.

Keywords: STEM education; subgoal learning; worked examples; expository text.

Varying Effects of Subgoal Labeled Expository Text in Programming, Chemistry, and Statistics

The learning sciences, and discipline based educational research in particular, has started to identify similarities and differences in teaching and learning practices in science, technology, engineering, and mathematics (STEM) domains. This community argues, correctly, that some aspects of STEM domains overlap (e.g., Kirschner, Verschaffel, Star, & Van Dooren, 2017), but each has features that make it unique, which matter when teaching in that domain (e.g., Mishra & Koehler, 2006). For example, many STEM disciplines are similar in that they use formulas to solve problems and need to train learners to match parts of a problem to part of a formula, making formula-matching a general problem solving skill (Singer, Nielsen, & Schwingruber, 2012). In contrast, STEM disciplines are different in that some are more rooted in the physical world, such as biology, some, such as algebra and computer science, focus on processes that are fundamentally abstract but can be exemplified through concrete examples, while others such as physics involve real objects that experience processes that are not directly observable even though the results might be. This difference matters because students in disciplines that have obvious connections to items that we interact with regularly are more prone to having misconceptions about how those objects operate (Chi, 2005) whereas students in more abstract disciplines can have more difficulty grasping concepts with which they have little experience (Booth, 1984; Kuchemann, 1981).

In light of differences like these, interventions that improve learning across different disciplines are valuable. Interventions that are effective across disciplines can provide insight into similarities among domains and how people learn them. They also have practical pedagogical value in providing approaches of wide utility. Similarly, finding interventions that are not effective across disciplines can provide insight into differences among domains and how those differences affect learning. Both types of results can contribute meaningfully to the learning sciences, making cross-disciplinary research beneficial to our community. The present study attempted to replicate the effect of a new intervention, subgoal labeled expository text, that was successful in improving learning in one domain, computer programming, but has yet to be tested in other domains.

Subgoal Learning and Subgoal Labeled Instructional Material

Subgoal labeled expository text is an extension of the subgoal learning framework, which focuses on teaching learners the subgoals of problem solving procedures to improve their retention and transfer (Catrambone, 1998). A subgoal is a functional component of a problem's solution. For example, consider the procedure described in Figure 1. To find the solution, which is your goal, you would divide 243 by 104 to calculate the average frequency of an event. The subgoal label, in this case "calculate average frequency," points out the function of this step to the learner. This explicit explanation of the function is necessary because novices tend to organize and encode procedural information based on surface features of the problem, such as frequency of a video being rented, rather than structural features of the problem, such as calculating the average frequency of an event (Bransford, Brown, & Cocking, 2000; Chi, Feltovich, & Glaser, 1981).

Subgoals repeat within a class of problems within a discipline, and learning to identify and complete subgoals helps learners solve novel problems better (Catrambone & Holyoak, 1990). For this reason, subgoal labels have been used extensively to help students learn from worked examples. Worked examples are example problems presented with their step-by-step solution. Subgoal labeled worked examples improve learning by highlighting the procedure used in examples (Atkinson & Derry, 2000; Atkinson, Catrambone, & Merrill, 2003; Margulieux, Catrambone, & Guzdial, 2016), providing organization for new information (Morrison, Margulieux, & Guzdial, 2016), and inducing self-explanation (Catrambone, 1998; Renkl & Atkinson, 2002; Morrison, Decker, & Margulieux, 2016).

In most domains, students receive both worked examples and expository text. Expository text is text that introduces and defines terminology and also describes problem solving procedures conceptually (Trafton & Reiser, 1993). For instance, the expository text in Figure 1 describes the procedure conceptually, whereas the worked example gives an exemplar of the procedure being used to solve a concrete problem. Expository text is valuable because learners who master the problem solving procedure conceptually can solve novel problems better than leaners who are given more specific instructions (Eiriksdottir & Catrambone, 2011). Because expository text is abstract, however, it can be difficult for students to grasp, especially when they have little knowledge in the field (Eiriksdottir & Catrambone, 2011; Fu & Gray, 2006). Therefore, the pairing of expository text and worked examples provides the learner with the conceptual problem solving procedure and concrete examples of this procedure being applied.

Just as adding subgoal labels to worked examples improved problem solving performance, adding subgoal labels to expository text could also improve problem solving. Subgoal labeled expository text could help students organize the information around functions, which is better than organizing it around calculation as they are prone to do (Atkinson, Catrambone, & Merrill, 2003), and adding signaling, which provides clues about which information is most important (Lemarié, Lorch, Eyrolle, & Virbel, 2008). Further, including subgoal labels in both expository text and worked examples could help students make connections between different representations of the same procedure, which makes both types of instructions more effective (McGee & Reis, 2012).

Prior Study to be Replicated

In line with these predictions, Margulieux and Catrambone (2016) found that subgoal labels improved learning for a programming task most if they were placed in the expository text *and* worked examples. In that study, participants were taught to create applications (apps) for Android devices with expository text, which conceptually described the process of creating an app, and a worked example, which showed the exact steps taken to create an app. The study manipulated two aspects of the instructional materials. The expository text either had subgoal labels or not, and the worked example either had subgoal labels or not. Therefore, participants received subgoal labels in both, one or the other, or none of the instructional materials. After instruction, participants' problem solving performance was measured by asking them to solve novel problems using the procedure that they had learned. Margulieux and Catrambone (2016) ran the experiment twice; one experiment allowed participants to reference the text and example during problem solving and the other experiment did not.

In both experiments, Margulieux and Catrambone (2016) found that participants who received subgoal labels in both the expository text and worked example performed statistically better than those in the other three conditions. For participants who were allowed to use the text and example while solving novel problems, the only group that performed statistically better than the others received subgoal labels in both the expository text and worked example. The remaining three groups that received subgoal labels in only the expository text, only the worked example, or in neither performed equally (Margulieux & Catrambone, 2016). For participants who were not allowed to use instructional materials while solving novel problems, there were three levels of performance. Those who received subgoal labels in both materials performed statistically better than those who received subgoal labels in only the example. Both of these group performed statistically better than participants who received subgoal labels in only the expository text or received no subgoal labels, meaning that subgoal labeled expository text had no effect unless it was paired with the subgoal labeled worked example.

The results of both experiments suggest that the effect of subgoal labeled expository text interacts with the effect of subgoal labeled examples rather than having an additive benefit. To explain why participants who received subgoal labeled expository instructions did not perform better that those who received no subgoal labels, Margulieux and Catrambone (2016) argued that receiving subgoal labels in expository text alone did not provide learners information that they could apply to problem solving. Thus, subgoal labeled text alone did not improve problem solving performance. To explain why participants who received both the subgoal labeled expository text and subgoal labeled worked example performed better than all other groups, they argued that receiving subgoal labels in both types of materials improved performance by showing connections between conceptual expository text and concrete worked examples. Showing these connections allowed learners to integrate the two types of information better, making the abstract information in expository text more accessible and the concrete information in the worked example more transferable. For these reasons, they concluded that learners who receive subgoal labels in both types of instruction would perform better than learners who do not.

Present Study

The present study extended this research by exploring the effect of subgoal labeled expository text and its interactions with subgoal labeled worked examples in STEM domains other than programming. One of the domains was a math discipline, statistics. It was chosen for two reasons. On the theoretical side, statistics is a domain about which many people have misconceptions because some of the formulas, particularly around probability, can initially seem counterintuitive (Lesser, 1998). On the experimental side, this domain was used for much of the original subgoal labeled worked example research (e.g., Catrambone, 1998); therefore, we know that the effect of subgoal labeled worked examples have a stable effect, allowing us to isolate the effect of subgoal labeled expository text. The second domain was a science discipline that has not included subgoal labels before, chemistry. It was chosen because the problem solving procedure had many features that were similar to the statistics procedure, using a mathematical equation to solve a problem, but also had some features that were unique to chemistry: molecule notation and balancing each side of the equation to conserve mass. Features of these domains and discussion of how they are similar or different from the original programming research and each other are discussed in the experiment descriptions.

The present study used methods similar to those described in Margulieux and Catrambone (2016). In a laboratory experiment, participants were given abstract expository text that explained how to solve problems and concrete worked examples that demonstrated a problem being solved. Depending on their randomly assigned condition, participants received subgoal labels in both, one or the other, or none of these instructional materials. Subgoals of the procedure were determined using the Task Analysis by Problem Solving (TAPS) procedure and consultation with subject-matter experts (Catrambone, 2011) as is typical in recent subgoal learning research (e.g., Margulieux & Catrambone, 2016).

It was hypothesized that the pattern of results in the new domains would replicate those found in the programming domain. That is, subgoal labeled expository texts would interact with subgoal labeled worked examples to produce the best problem solving performance. Labeled expository text without labeled worked examples was not expected to improve problem solving because learners were not expected to know how to apply the information in labeled expository text to solving problems. Labeled worked examples without labeled expository text were expected to improve problem solving performance as they have in the past (i.e., when learners are not able to reference the worked examples during later problem solving), but they were not expected to improve performance as much as labeled examples and text combined.

Experiment 1

In this experiment, participants were taught to solve problems using the Poisson distribution, which is used to find the probability of an event happening over a given time period (see Figure 1). The problem solving procedure had three subgoals: identify the event and interval from the problem, calculate the average frequency of the event, and compute the probability with the formula. Each subgoal was used once in each problem, and typically included only one sub-step, meaning that the problems were short. Furthermore, mapping between the problem statement and problem solution was typically straightforward. Identifying the event was sometimes counterintuitive, but after that the steps of the procedure were executed in the same way for every problem. These features make using the Poisson distribution relatively simple when compared to the app programming procedure described earlier. Due to these differences in complexity and length of worked examples, the statistics instructions had three short examples unlike the programming instructions, which had one long example. In both cases, participants saw multiple instances of each subgoal.

Method

Design. The experiment had four between-subject conditions based on two independent variables: the format of the worked examples (subgoal labeled or unlabeled) and the format of expository text (subgoal labeled or unlabeled). This design means that participants could either have subgoal labels in both text and example, either text or example, or neither. Participants were randomly assigned to conditions with 30 participants per condition. The dependent variables were problem solving performance and time on task.

Participants. Participants were 120 students from a mid-sized university who received class credit for participation. Participants must not have taken more than one statistics course in secondary school or college, and if they had taken a statistics course, it must not have been within the past year. These restrictions were intended to reduce prior knowledge of the subject matter.

Consistent with the student population, 42% of participants were female. The mean age was 19.1, and average years in college was 1.8. Average GPA was 3.3 (out of 4), though this number might be unreliable because 52% of participants were in their first semester of college and did not have a GPA yet. Less than half of participants, 43%, had taken one statistics classes beforehand, but prior experience did not significantly correlate with performance or time on task, $r = .01$, $p = .95$, and $r = .02$, $p = .87$, respectively.

Procedure. Sessions were between 30 and 45 minutes depending on how quickly participants completed the tasks. Throughout the session, experimenters provided help with only administrative questions (e.g., "Could I have a different calculator?").

During the instructional phase, participants received paper-based instructional materials for solving problems using the Poisson distribution. The materials included expository text and three worked examples (see Figure 1). Participants were also required to complete three practice problems, which gave them an opportunity to practice using the procedure. The exercises were similar to the problems that were worked in the examples but with different contexts and numerical values; therefore, these problems required near transfer. For instance, the worked example would be for the problem, "Over a period of time at a certain video store, 243 people rented 104 different videos. Use the Poisson distribution to determine the probability that a randomly chosen video was rented exactly 4 times," and the practice problem would read, "A number of celebrities were asked how many commercials they made of the last year. The 40 celebrities made a total of 142 commercials. Use the Poisson distribution to determine the probability that a randomly chosen celebrity made exactly 4 commercials." The correct answer for each problem was included on the sheet to allow participants to check their work. Participants had up to 20 minutes to use the materials in whichever way they wanted to learn the procedure.

During the assessment phase, participants completed problem solving tasks. The problem solving tasks asked participants to solve novel problems using the procedure that they had learned. The assessment problems were different from the examples and practice problems that they had seen, but with the same subgoals. For instance, one assessment problem was, "Malingerers Life currently insures 10 men aged 62. The average probability of a man aged 62 dying within the next year is .30. On average, 3 of the insured men will die in a coming year, but there are variations in any given year. Use the Poisson distribution to determine the probability that Malingerers Life will have to pay exactly 3 claims on those 10 policies during the coming year?"

Participants were given five problems and up to 25 minutes to work on them. Similar to an exam, participants were not given unlimited time to work on problems. During the assessment, participants did not have access to the instructional materials. Participants were informed about this restriction at the beginning of the session to encourage effortful learning. Participants did have access to the formula for the Poisson distribution. At the end of each problem, participants were asked to rate how difficult it was on a scale from "1 – not difficult at all" to "7 – very difficult." The amount of time participants spent using the instructional materials and solving the problems was recorded by the experimenter.

Performance on problem solving tasks was scored based on the number of correct steps (one point per step) that participants made toward the solution. This scoring scheme allows participants to receive points for using components of the procedure correctly even if they made a mistake and did not get the correct answer (i.e., partial credit). Computational errors were not counted against participants. For example, if participants used the correct numbers for a step but misplaced the decimal point, they would receive credit for that step. For this study, procedural accuracy was more important than numerical accuracy of the solutions. Due to the objective nature of the scoring, a second rater was not needed. The maximum score for this task was 17.

Results

Problem Solving Performance. Effect sizes for the following results are reported in both est. ω^2 and *f*. Est. ω^2 describes the proportion of variance accounted for by the intervention, which is similar to η^2 but slightly more conservative. For example, an est. ω^2 of .10 would mean that 10% of the variance in scores can be attributed to the intervention. An est. ω^2 of .06 in social science is considered a medium-sized effect (Cohen, 1969). The other effect size, *f*, describes the difference between groups using the standard deviation as the unit of measurement. This statistic is similar to Cohen's *d*, but it is for ANOVA rather than *t*-tests. For example, an *f* of .25 would mean that the difference between groups is half (*f* value times two; Cohen, 1988) of the standard deviation within groups. In social science, an *f* of .25 is considered a medium-sized effect (Cohen, 1969). In this paper, *f* is reported to describe the magnitude of differences between groups only when they are statistically significant.

Consistent with the programming results and previous literature, participants who received subgoal labels in the example (M = 9.5, SD = 3.8) performed better than those who did not (M = 7.3, SD = 4.8), F (1, 116) = 8.00, MSE = 17.6, p = .006, est. ω^2 = .06, f = .26. No difference was found between participants who received or did not receive subgoal labels in text (average M = 8.4, SD = 4.3, F (1, 116) = .29, p = .61, est. ω^2 = .06. In contrast to the programming results, there was no interaction between text design and example design, F (1, 116) = .10, p = .75, est. ω^2 = .05. This pattern of results suggest that only subgoal labeled worked examples improved problem solving performance (see Figure 2).

Time on Task. Participants had up to 20 minutes to study the instructional materials and complete the exercise problems, but most did not use the full time (M = 14.1 minutes, SD = 4.1). No difference was found between participants who received subgoal labels or not in examples, F (1, 116) = .16, MSE = 16.9, p = .69, est. $\omega^2 < .01$, or in the text, F (1, 116) = .18, p = .67, est. $\omega^2 < .01$. No interaction was found either, F (1, 116) = 2.75, p = .10, est. ω^2 = .02. These results suggest that participants spent approximately the same amount of time learning the procedure regardless of their assigned condition.

Participants had up to 25 minutes to complete the problem solving tasks, and about a third used the entire time, 34% (M = 22.1 minutes, SD = 3.6). No difference was found between participants who received subgoal labels or not in examples, F (1, 116) = .71, MSE = 13.5, p = .40, est. ω^2 = .01, or in the text, F (1, 116) = 1.00, p = .33, est. ω^2 = .01. No interaction was found, F (1, 116) = .11, p = .68, est. $\omega^2 < .01$. These results suggest that all participants spent approximately the same amount of time solving problems. Time on task was not correlated with performance, r = .07, p = .48.

Unlike in the previous research in programming (Margulieux & Catrambone, 2016), the current research done with statistics found that subgoal labeled expository text had no effect on problem solving performance, either by itself or as part of an interaction. The lack of effect led the researchers to hypothesize that the statistics procedure was too simple and straightforward to make subgoal labeled expository text helpful to learners. In this procedure, and unlike in programming, it was easy to map the steps described in the expository text to problem solving because there were few ways in which the procedure could be applied to problem solving, making the problem solving process straightforward. To examine whether the complexity of the procedure impacted the effect of subgoal labeled text, the intervention was tested again in a chemistry procedure that was more complex.

Experiment 2

Experiment 2 explored the effect of subgoal labeled expository text and examples (see Figure 3) in the chemistry domain. In this experiment, participants were taught to solve problems using reaction stoichiometry. This procedure was somewhat similar to the statistics procedure. It had three subgoals: convert molecules from given units to moles, find moles of molecule B required to support reaction for molecule A, and convert molecules from moles to desired unit. Like in statistics, each subgoal was completed once per problem and required the learner to find the correct values from the problem statement to apply to the equation to complete the solution. In both procedures, learners needed to map from the problem statement to the mathematical equation, which can be challenging in chemistry problems because the element you need might be part of a molecule that complicates the equation (e.g., the element Fe in the molecule Fe_2O_3 comes two at a time). The chemistry procedure added complexity over the statistics procedure because the equation was different for each problem, based on the chemical reaction, and because the equation had to be balanced at each step to ensure conservation of mass. Therefore, achieving each subgoal was less straightforward than in the statistics procedure.

Method

The method for Experiment 2 was the same as for Experiment 1 (i.e., in sample size, selection of participants, procedure, and design) though sessions took slightly longer, between 40 and 50 minutes. Participants must not have taken more than two chemistry courses, and they must not have taken a chemistry course in the past year. These restrictions were intended to minimize effects of prior experience.

Thirty-four percent of participants were female, the mean age was 19.5, and average years in college was 2.3. Average GPA was 3.4 (out of 4), though this number might be unreliable because 42% of participants were in their first semester of college and did not have a GPA yet. Most participants had taken one to two chemistry classes, typically in secondary school (M = 1.4), but prior experience did not significantly correlate with performance or time on task, $r = .10$, $p = .25$, and $r = .05$, $p = .57$, respectively.

To assess their learning, participants were given five problem solving tasks that were different from the examples and exercise problems that they had seen during the instructional period but were solved using the same subgoals as in Experiment 1. They had 25 minutes to complete these problems. Performance on problem solving tasks was scored based on the number of correct steps (i.e., one point per step) that participants made toward the solution. Computational errors were not counted in the scoring. The maximum score for this task was 15.

Results

Problem Solving Performance. Consistent with the programming and statistics results, participants who received subgoal labels in the examples ($M = 11.4$, $SD = 2.6$) performed better than those who did not ($M = 9.6$, $SD = 4.3$), F (1, 116) = 10.35, MSE = 12.2, $p = .002$, est. $\omega^2 = .07$, $f = .29$. Unlike in the programming and statistics results, participants who received subgoal labels in the text ($M = 11.4$, $SD = 2.8$) performed better than those who did not ($M = 9.9$, $SD = 4.2$), F (1, 116) = 6.18, $p = .014$, est. $\omega^2 = .04$, $f = .23$. No interaction between the example and procedure design was found, F (1, 116) = 1.27, $p = .26$, est. $\omega^2 = .02$. Because the subgoal labeled expository text improved performance whether it was paired with subgoal labeled worked examples or not, these results suggests that subgoal labels in the expository text are not necessarily always ineffective without subgoal labeled examples, as they were in the programming experiment. Instead, subgoal labeled expository text and examples might improve problem solving performance in an additive way in some cases (see Figure 4).

Time on Task. Participants had up to 20 minutes to study the instructional materials and complete the exercise problems, but most did not use the full time (M = 14.7 minutes, SD = 4.0). No difference was found between participants who did or did not receive subgoal labels in examples, F (1, 116) = 2.57, MSE = 15.8, p = .11, est. ω^2 = .02, or text, F (1, 116) = .189, p = .67, est. ω^2 < .01. No interaction was found either, F (1, 116) = .166, p = .68, est. ω^2 < .01. These results suggest that participants spent approximately the same amount of time learning the procedure regardless of their assigned condition.

Participants also had up to 25 minutes to complete the problem solving tasks, and again most did not use the full time (M = 15.1 minutes, SD = 4.7). For the assessment, participants who received subgoal labels in the examples (M = 14.1, SD = 4.6) completed the tasks faster than those who did not (M = 16.2, SD = 4.6), F (1, 116) = 6.10, MSE = 20.7, p = .015, est. ω^2 = .06, f = .23. No difference was found between participants who received subgoal labeled text or not (average M = 15.1, SD = 4.7), F (1, 116) = 2.89, p = .09, est. ω^2 = .02, nor was there an interaction, F (1, 116) = .79, p = .38, est. ω^2 = .01. Time on task was not correlated with performance, r = .06, p = .52.

Discussion

For both procedures that were studied, subgoal labeled worked examples had a consistent effect: improved problem solving performance. These results align with the original Margulieux and Catrambone (2016) study and other previous research (e.g., Catrambone, 1998; Margulieux et al., 2016), suggesting that subgoal labels reliably help students learn from worked examples more effectively. For subgoal labeled expository text, however, the effect differed in both domains, and neither matched the pattern of results found in programming. The differences among procedures and results are shown in Table 1. Please notice that the procedures are in different domains in these experiments, but procedures within domains could have important differences as well (Kirschner et al., 2017). For instance, procedures in computational biology would be different from experimental biology, and work at the cellular, organism, or system level would be different.

The primary limitation of this work is that the results were not the hypothesized results. Therefore, an attempt to explain the results in a way that generalizes beyond the current data sets would not be scientifically sound. In light of that, the following discussion is a post hoc speculation about the underlying causes of the results. The proffered explanations of the results would need to be operationalized and scientifically tested in future research to hold weight.

The subgoal labeled expository text improved problem solving performance for the medium and highest complexity procedures and had no effect for the lowest complexity procedure. Margulieux and Catrambone (2016) argued that subgoal labels helped learners integrate the abstract information in the expository text and concrete information in the worked example. It could be the case that this help is not needed for simpler procedures. Specifically, when the procedure for solving problems is straightforward, meaning that there is not much variability in how the procedure is applied to problems, learners are more easily able to map abstract instructions to solving concrete problems. Therefore, the learners likely do not need as much help implementing procedures described in abstract expository text to solve new problems. When the method of applying the procedure to the problems is less obvious, however, subgoal labeled expository text seems to help improve performance. Because the current study was set up as a replication study, it does not provide adequate evidence for this hypothesis. This hypothesis could be tested by experimentally manipulating the complexity of the procedure between participants in addition to the subgoal manipulations.

Another possible explanation is that for simpler procedures, students might not need expository text to gain a conceptual understanding of the procedure. Earlier work by LeFevre and Dixon (1986) and Zhu and Simon (1987) found that students prefer to learn from worked examples and many of them completely ignore expository text. Students ignore expository text because they find it difficult to apply to concrete problems (Fu & Gray, 2006). VanLehn, Jones, and Chi (1992) found that learners only look to expository text to resolve problem solving impasses, which they are less likely to encounter in more simple procedures. Based on these student practices, which the authors would argue are still true today, subgoal labeled expository text might have had no effect on performance for more simple procedures because the participants did not use it. In this study, total time spent studying instructional materials was collected, but that measurement does not distinguish between time spent on text and examples; therefore, the current paper cannot provide evidence for this hypothesis. It could be directly tested with an additional experimental condition that had worked examples only with no expository text. Understanding how complexity affects instructional interventions could help us focus our efforts on learning environments in which instructional supports will be most effective for students.

In addition to the complexity of the procedure, the tools used to solve problems might have impacted the effect of subgoal labeled expository text. In programming, subgoal labeled text did not improve problem solving performance unless it was paired with subgoal labeled worked examples. In contrast, subgoal labeled text in chemistry improved problem solving performance regardless of whether or not it was paired with subgoal labeled examples. To complete the programming procedure, participants had to learn to use the tool (i.e., the programming interface) and the procedure to solve problems. In this case, participants performed better only when subgoal labels appeared in both expository text and example. In this case also, the worked example provided information about how to use the interface because it showed the creation of an app with the interface, whereas the expository text only conceptually described the process of making an app. Therefore, subgoal labels in only expository text could be difficult to translate to problem solving within the interface.

In contrast to the programming procedure, the chemistry participants used familiar tools (i.e., calculator, pencil, and paper) to solve problems, and a subgoal labeled text improved performance over unlabeled text. In this case, learners likely needed less help to apply abstract instructions to problems, making subgoal labels in expository text effective by themselves. This hypothesis could be tested by repeating the experiment in statistics and chemistry using different tools. For example, statistics could be taught using Roman numerals, or chemistry could be taught using a drag-and-drop computer interface. The effects of tools on learning is particularly important for understanding the impact of educational technology and predicting when it will be effective.

Subgoal labeled expository text improved performance on problem solving using complex procedures (i.e., the programming and chemistry procedures). These are both procedures with which learners are more likely to have difficulty. In addition, these improvements affected learners equally, regardless of prior experience, and they did not come at the cost of increased time on task. However, subgoal labeled expository text did not improve problem solving performance for a relatively simple procedure (i.e., the statistics procedure), making improvements to expository text for that procedure unnecessary. This study concludes that subgoal labeled expository text can help learners perform better on problem solving tasks, but the features of the procedure determine how subgoal labeled text would affect performance.

References

Atkinson, R. K., Catrambone, R., & Merrill, M. M. (2003). Aiding Transfer in Statistics: Examining the Use of Conceptually Oriented Equations and Elaborations During Subgoal Learning. *Journal of Educational Psychology, 95*(4), 762.

Atkinson, R. K., & Derry, S. (2000). Computer-based examples designed to encourage optimal example processing: A study examining the impact of sequentially presented, subgoal-oriented worked examples. Proceedings of 2000 International Conference of the Learning Sciences, 132-133.

Booth, L. R. (1988). Children’s difficulties in beginning algebra. *The ideas of algebra, K-12*, 20-32.

Bransford, J. D., Brown, A. L., & Cocking, R. R. 2000. *How people learn: Brain, mind, experience, and school.*

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General, 127*, 355-376.

Catrambone, R. (2011). Task Analysis by Problem Solving (TAPS): Uncovering expert knowledge to develop high-quality instructional materials and training. Paper presented at the 2011 Learning and Technology Symposium (Columbus, GA, June).

Catrambone, R., & Holyoak, K. (1990). Learning subgoals and methods for solving probability problems. *Memory & Cognition, 18*(6), 593-603.

Chi, M. T. H. (2005). Commonsense Conceptions of Emergent Processes: Why Some Misconceptions Are Robust. *Journal of the Learning Sciences, 14*(2), 161- 199.

Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*(2), 121-152.

Cohen, J. (1969). Statistical power analysis for the behavioural sciences. New York: Academic Press.

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

Eiriksdottir, E., & Catrambone, R. (2011). Procedural instructions, principles, and examples: how to structure instructions for procedural tasks to enhance performance, learning, and transfer. *Human Factors, 53*(6), 749-770.

Fu, W., & Gray, W. D. (2006). Suboptimal tradeoffs in information seeking. *Cognitive Psychology, 52*, 195-242.

Kirschner, P. A., Verschaffel, L., Star, J., & Van Dooren, W. (2017). There is more variation within than across domains: an interview with Paul A. Kirschner about applying cognitive psychology-based instructional design principles in mathematics teaching and learning. *ZDM, 49*(4), 637-643.

Küchemann, D. (1981). Cognitive demand of secondary school mathematics items. *Educational Studies in Mathematics*, *12*(3), 301-316.

LeFevre, J., & Dixon, P. (1986). Do written instructions need examples? *Cognition and Instruction, 3*, 1-30.

Lemarié, J., Lorch Jr, R. F., Eyrolle, H., & Virbel, J. (2008). SARA: A text-based and reader-based theory of signaling. *Educational Psychologist, 43*(1), 27-48.

Lesser, L. (1998). Countering indifference using counterintuitive examples. *Teaching statistics, 20*(1), 10-12.

Margulieux, L. E., & Catrambone, R. (2016). Improving problem solving with subgoal labels in procedural instructions and worked examples. *Learning and Instruction*, *42*, 58-71. doi: 10.1016/j.learninstruc.2015.12.002

Margulieux, L. E., Catrambone, R., & Guzdial, M. (2016). Employing subgoals in computer programming education. *Computer Science Education*, *26*(1), 44-67. doi: 10.1080/08993408.2016.1144429

McGee, P., & Reis, A. (2012). Blended Course Design: A Synthesis of Best Practices. *Journal of Asynchronous Learning Networks*, *16*(4), 7-22.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017-1054.

Morrison, B. B., Decker, A., & Margulieux, L. E. (2016). Learning loops: A replication study illuminates impact of HS courses. In Proceedings of the Twelfth Annual International Conference on International Computing Education Research (pp. 221-230). New York, NY: Association for Computing Machinery. doi: 10.1145/2960310.2960330

Morrison, B. B., Margulieux, L. E., & Guzdial, M. (2015). Subgoals, context, and worked examples in learning computing problem solving. Proceedings of the Eleventh Annual International Conference on International Computing Education Research (pp. 21-29). New York, NY: Association for Computing Machinery. doi: 10.1145/2787622.2787733

Renkl, A., & Atkinson, R. K. (2002). Learning from examples: Fostering self-explanations in computer-based learning environments. *Interactive Learning Environments*, *10*(2), 105-199.

Singer, S. R., Nielsen, N. R., & Schweingruber, H. A. (Eds.). (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. National Academies Press.

Trafton, J. G., & Reiser, B. J. (1993). *Studying examples and solving problems: Contributions to skill acquisition*. Technical report, Naval HCI Research Lab, Washington, DC, USA.

VanLehn, K., Jones, R., & Chi, M. T. H. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences, 2*, 1-59. http://dx.doi.org/10.1207/s15327809jls0201_1

Zhu, X., & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction, 4*(3), 137-166.

Table 1: Differences among features of procedures and effects of subgoal labeled instructional materials in different disciplines.

Discipline	Features of Problem Solving Procedure	Effect of Subgoal Labeled Materials
Programming	High complexity – six subgoals (half straightforward, half not) repeated multiple times per problem	Subgoal labeled expository text combined with subgoal labeled worked example improved problem solving performance more than subgoal labeled example alone. Subgoal labeled text alone had no effect.
Statistics	Low complexity – three straightforward subgoals used once per problem	Subgoal labeled expository text did not affect performance. Subgoal labeled worked examples improve performance.
Chemistry	Mid complexity – three somewhat-straightforward subgoals used once per problem	Subgoal labeled expository text improved performance as did subgoal labeled worked examples. There was an additive effect of subgoal labeled text and subgoal labeled examples.

a) Partial Subgoal Labeled Expository Text

Calculate average frequency

Unless the average frequency is given to you, you'll need to calculate it. To do this find the total frequency of an event (i.e., the total number of times an event occurs) or calculate it by adding simple or weighted frequencies. Once you have the total frequency, divide it by the number of intervals.

Compute probability

Once you have the average frequency, use it and the number of times an event occurs in the formula to compute the probability.

b) Partial Unlabeled Expository Text

Unless the average frequency is given to you, you'll need to calculate it. To do this…

Once you have the average frequency, use it and the number of times an event occurs…

c) Partial Subgoal Labeled Worked Example

Problem:

Over a period of time at a certain video store, 243 people rented 104 different videos. Use the Poisson distribution to determine the probability that a randomly chosen video was rented exactly 4 times.

Calculate average frequency

$$\lambda = \frac{\text{total frequency}}{\text{number of intervals}} = \frac{243}{104}$$

Compute probability

$$P(X = 4) = \frac{e^{-2.34}\, 2.34^4}{4!} = .12$$

d) Partial Unlabeled Worked Example

$$\lambda = \frac{\text{total frequency}}{\text{number of intervals}} = \frac{243}{104}$$

$$P(X = 4) = \frac{e^{-2.34}\, 2.34^4}{4!} = .12$$

Figure 1: Expository text (a, b) and worked examples (c, d), either subgoal labeled (a, c) or unlabeled (b,d), describing and demonstrating the procedure used to solve problems with the Poisson distribution. The only difference is the presence of subgoal labels.

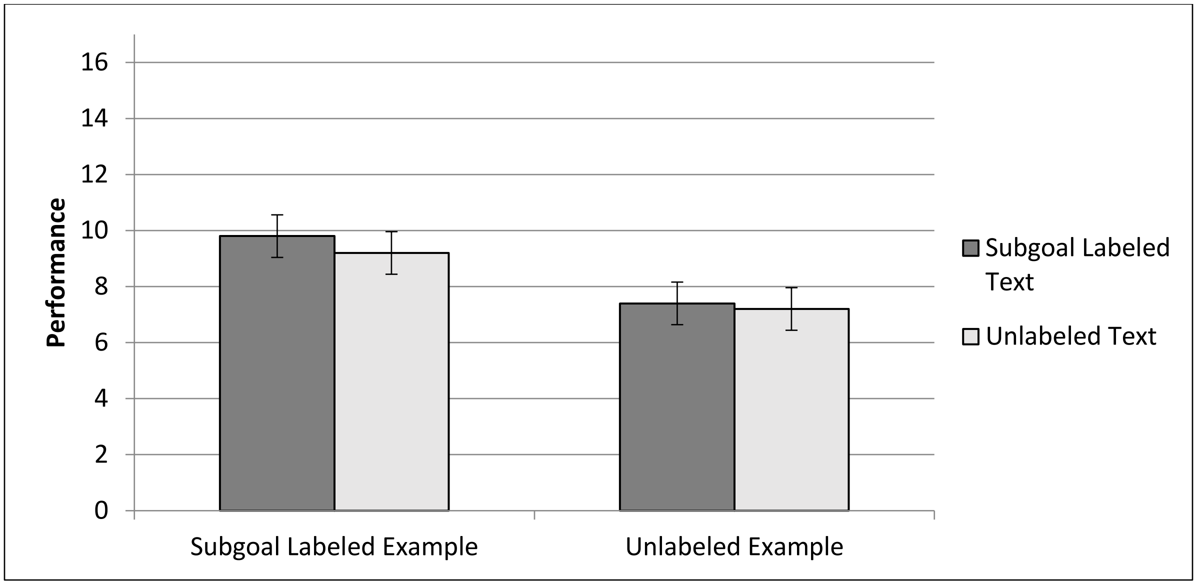

Figure 2: Performance on statistics problem solving tasks in Exp. 1. Error bars represent standard deviation.

a) Partial Subgoal Labeled Expository Text

Convert from units to moles

The first step is to convert the given mass of one substance (A) in grams into amount in moles by using its molar mass.

Multiply by mole ratio

The second step is to use the mole ratio derived from the coefficients in the balanced chemical equation to convert from the amount of one substance (A) into the amount in moles of the other substance (B).

b) Partial Unlabeled Expository Text

The first step is to convert the given mass of one substance (A) in grams into amount in moles by using its molar mass.

The second step is to use the mole ratio derived from the coefficients in the balanced chemical equation to convert…

c) Partial Subgoal Labeled Worked Example

Problem:

What mass of iron oxide, Fe_2O_3, present in iron ore is required to produce 10.0 g of iron, Fe, when it is reduced by carbon monoxide gas to metallic iron and carbon dioxide gas?

$$Fe_2O_3(s) + 3\,CO\,(g) \rightarrow 2\,Fe\,(s) + 3\,CO_2(g)$$

Convert from units to moles

$$\text{Amount of Fe (mol)} = \frac{10.0\text{g Fe}}{55.85\,\frac{\text{g}}{\text{mol}}} = \frac{10}{55.85}\,\text{mol Fe}$$

Multiply by mole ratio

$$\text{Amount of } Fe_2O_3 \text{(mol)} = \frac{10}{55.85}\,\text{mol Fe} * \frac{1\text{ mol } Fe_2O_3}{2\text{ mol Fe}}$$

d) Partial Unlabeled Worked Example

$$\text{Amount of Fe (mol)} = \frac{10.0\text{g Fe}}{55.85\,\frac{\text{g}}{\text{mol}}} = \frac{10}{55.85}\,\text{mol Fe}$$

$$\text{Amount of } Fe_2O_3 \text{(mol)} = \frac{10}{55.85}\,\text{mol Fe} * \frac{1\text{ mol } Fe_2O_3}{2\text{ mol Fe}}$$

Figure 3: Procedural instructions describing the procedure used to solve problems with reaction stoichiometry.

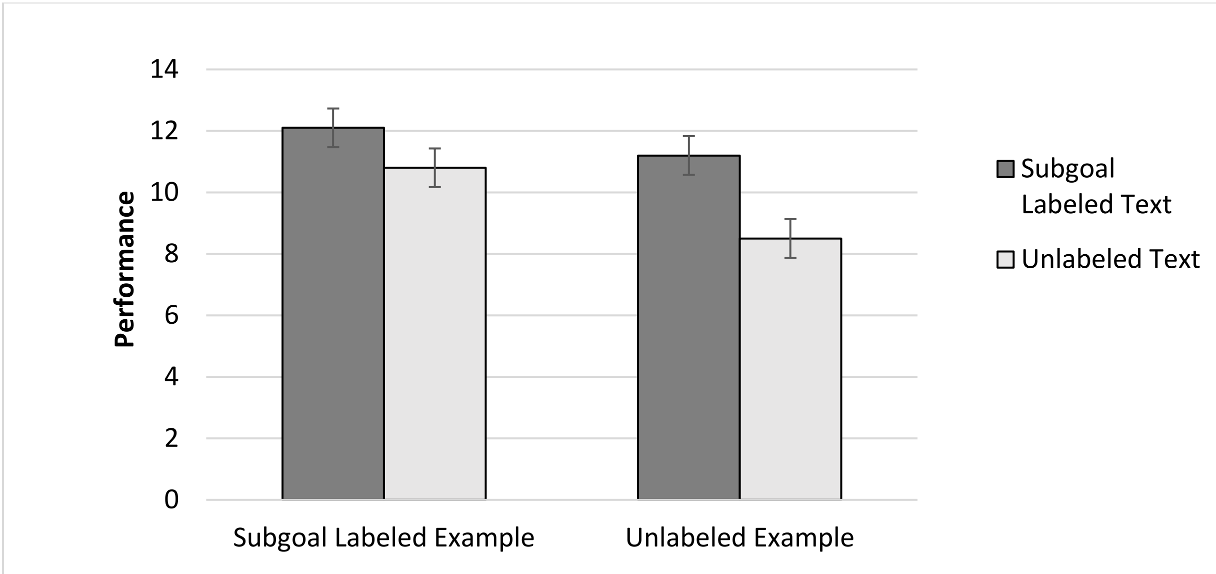

Figure 4: Performance on chemistry problem solving tasks in Exp. 2. Error bars represent standard deviation.