Changes in Self-Explanation while Learning Vector Arithmetic

Troy D. Kelley (tkelley@arl.mil)
Human Factors & Applied Cognition Program
George Mason University
Fairfax, VA 22030

Irvin R. Katz (ikatz@gmu.edu)
Human Factors/Applied Cognition Program
George Mason University
Fairfax, VA 22030

Abstract

Verbal elaboration of a worked example has been shown to be helpful to learners before attempting to solve similar problems. This has been termed as the self-explanation effect. (Chi, Bassok, Lewis, Reimann & Glaser, 1989). This study examined how self-explanation changes before and after sequential problem solving rounds. We found that changes in self-explanation within an individual may affect individual performance across a series of problem solving episodes. Also, some participants appear to use the worked-out example as a self-generated feedback (SGF) mechanism to help with their problem solving rounds, while other participants do not. Locations or points in a worked-out example where self-explanation (elaboration) is most likely to occur for students with higher performance scores versus those with lower performance scores, is discussed. The implications of these differences for the design of a computational cognitive model are also addressed.

Introduction

Learning from examples has been shown to be an important aid in the learning process (VanLehn, 1986, 1996). Using examples to provide a basis for learning has also been shown to be the preferred way of learning by novices (Anderson, Farell, & Sauers, 1984; Pirolli & Anderson, 1985; Recker & Pirolli, 1995). However, most research has been conducted on worked-out examples which were presented before problem solving episodes (Chi, Bassok, Lewis, Reimann & Glaser, 1989). Chi et. al.'s (1989) original study was limited since the worked examples were only presented prior to problem solving rounds, which did not allow for an examination of the changes in self explanation as learning progressed. In this study, we will look at how learning might proceed if worked-out examples are presented following problem solving episodes instead of prior to problem solving.

In the seminal work on learning from examples (Chi, Bassok, Lewis, Reimann & Glaser, 1989), a *self-explanation effect* was found in effective learners who could, among other things, use a worked-out example to elaborate upon broader principles which they had previously acquired while studying text. The authors also found that effective learners monitored their own performance and knowledge base better than ineffective learners; which has been confirmed by some researchers (Ferguson-Hessler & DeJong, 1990), but questioned by others (Renkl, 1997).

We were interested in how self explanations might change as subjects examined worked examples after problem solving rounds. One hypothesis might be that subjects will have a strategy of using the worked-out example which follows the problem solving differently than they used earlier worked examples. Given that subjects have had the experience of attempting to solve earlier problems, they may choose to use the latter worked examples as a feedback mechanism to the previous problem solving rounds. We characterized this as self generated feedback (SGF). This feedback mechanism would allow subjects to use a strategy of analyzing their previous problem solving rounds (from memory) in order to improve subsequent problem solving. However, other subjects may decide not to use the latter worked examples as a feedback mechanism, instead they may concentrate all of their learning efforts on the first worked example. In which case, these subjects would not show any signs of SGF.

Our research agenda addresses: 1) How do self-explanations change as performance improves? 2) Do subjects use self-explanation strategically and can these strategies be detected? 3) Where in the worked example is a subject most likely to engage in self-explanation behavior; are these locations stable across different worked examples?

We wanted to examine the possibility that subjects might have an identifiable strategic use for the different worked examples. One strategic use could be that subjects would use the later worked examples as a feedback mechanism to their earlier problem solving episodes. If this were occurring, this would change the nature of the latter self-explanation statements. Subjects would begin to make statements which referenced earlier gaps in their knowledge. For example, a subject might say, "Oh, now I see how to use that equation, that is not what I was doing before."

We hypothesized that if participants were using the second or latter worked examples as a SGF mechanism, then they would show fewer self-explanation statements than learners who relied heavily on the first worked-out example. Also, they would utter fewer words on the first worked example and instead concentrate their efforts on the latter worked examples. However, if a subject were relying heavily on the first worked example to establish a foundation for subsequent problem solving, and not relying on latter worked-examples to provide SGF to their problem solving episodes; then the change in explanation statements should be significant. Also, the number of utterances while studying the first worked example would be relatively high. We also hypothesized that if a subject were using a delayed SGF strategy, then their performance would be worse than a subject who was not using the strategy.

Finally, we wanted to identify specific points in the worked example which were likely to trigger a self-explanation episode. This was particularly critical for building any models of the self-explanation process over time. As our model proceeded through the worked example, the model's actions should correlate with what the average subject does at each line of the example. If certain lines in the worked example are more likely to trigger explanation events by the participants, then the model should also respond with explanation events in the same places of the worked example. Furthermore, it was necessary to see if the worked examples differed in the places where explanation statements were likely to occur, as the subjects progressed through the experiment.

Method

Participants

Participants were seven high school juniors and three college students. The high school students were recruited from the same physics class which, a few months earlier, had covered vector arithmetic concepts simpler than those covered in the present experiment. The college students (freshman and two seniors) had no prior physics training beyond a high school course. All were paid for their participation.

Materials and Design

Participants completed four self-explanation tasks conducted at regular intervals throughout the experiment. Each task consisted of the participant talking aloud while studying one of four worked-out examples. All participants were given the same tasks to perform in the same order.

Problem solving tasks.

Participants completed four rounds of problem solving, three sets of 10 problems and a final round of 8 problems. All the problems were similar to the worked examples in that they described two or more vectors (in the context of the story) and asked the participants to find the magnitude and direction of the resultant. Unlike the worked examples, the problem statement was included no diagram.

Procedure

Students participated in individual sessions lasting 2.5 3 hours. To complete the experiment, students generally attended 5 sessions, although some students needed fewer sessions. In the first session, students studied a textbook chapter on vector arithmetic and completed the first self-explanation task. For the self-explanation task, students were asked to, "Study this example as if you were studying for a test. Try to understand why each solution step was taken and why the solution correctly answers the question." Students had access to a calculator while performing all tasks. During the study, students were asked to "talk aloud" (Ericsson & Simon, 1993), providing concurrent verbal protocols.

The remaining sessions included both self-explanation tasks and problem solving. In the second session, students completed 10 vector arithmetic problems (referred to as round 1 or R1), without any feedback on the correctness of their responses. The work done by the students during the problem solving episodes was not analyzed in great detail for this paper. We were primarily concerned with the work done by the students while they where studying the examples. Information about the problem solving rounds is presented here for clarity and completeness.

In the third and fourth sessions, students completed a self-explanation task (SE2 and SE3) followed by solving of another set of 10 vector arithmetic problems. In the final sessions, students completed a set of 8 vector arithmetic problems, some of which were easier and some of which were more difficult than problems completed during the previous sessions. After this fourth round of problem solving (R4), students completed the final self-explanation task (SE4). Students were then debriefed.

Results

Background

This study proceeded in two major phases and consequently the results will be presented in two major sections: replication and changes in self-explanation. The replication section addresses the major findings of the Chi et al. (1989) study. The "changes in self-explanation" section includes discussions on the observed changes in self-explanation over time. There are also subsections which include the strategies participants exhibited while using the worked examples and the computational cognitive model which was based on expert performance while studying one worked example.

Replication

Replication of the Chi et al (1989) study was conducted to determine the validity and generalizability of our data and to determine if there were any inconsistencies with the original work of Chi et al. (1989). However, a few problems were encountered. Most of the data which Chi et al. (1989) analyze was split into good versus poor performers, of which they seemed to have a clear delineation. We conducted a similar split (a median split) however, we had only one subject performing above 50 percent.

Chi et al. (1989) first analyze their data determining a count of the number of phrases made by good and poor students during the problem solving episodes. Instead of using a line count, we used word count by each subject while they were studying the example. Because some of the verbal protocol lines were long, whereas other lines were single words or phrases, we feel a word count might be more accurate than a line count. They find the line count to be significant "(142 lines versus 21 lines, $t(6) = 1.97$, $p < .05$)." We found that on average the good students uttered more words than the poor performing subjects (1119.33 versus 586.75, $t(5) = 1.96$, $p > .05$) performance on the first set of ten problems for 7 subjects in the experiment. A fairly high Pearson's correlation coefficient of ($r(6) = .42$, $p > .05$) was obtained but, due to the small sample size, this was not significant.

Next, Chi et al. (1989), found that good students produced significantly more explanations that related to the content of the problem than did the poor students (15.3 versus 2.8). Our data support this result (22.6 versus 12.5). Chi et al. (1989) go on to analyze the number of times the good and poor students refer to the worked-out examples during problem solving rounds. They found the good students referred to the example less often than did the poor students. We looked at the two best performing subjects (with performance scores i.e. questions answered correctly of 80 and 50) in comparison with the two worst performing subjects (with performance scores of 20 and 30) and found that the best performing students refer to the example fewer times than did the worst performing subjects. Furthermore, we found that the amount of explanations while studying an example was correlated to the subsequent performance of the subject during the following problem solving rounds ($r(6) = .65$).

Our data was consistent with the results presented by Chi et al. (1989), with the exception of one area - the amount of negative versus positive monitoring statements uttered by participants. Positive monitoring statements included statements such as: "OK, I understand this", while negative monitoring statements consisted of statements such as, "What does this mean? I don't understand." Chi et al. (1989) found significant results on the negative monitoring variable. Poor performers averaged 1.1 negative monitoring statements while good performers averaged 9.3 negative monitoring statements. This is where our data is inconsistent with Chi et al.'s (1989) original findings. We found, if anything, that negative monitoring and subsequent performance seemed to be slightly inversely related, however the result was non-significant ($r(6) = -.39$, $p > .05$). Good students had an average of 3.3 negative monitoring statements while poor subjects had an average of 4.0 negative monitoring statements. Renkl (1997) also found no relation between the amount of negative monitoring statements during the study of a worked-out example subsequent problem solving performance.

Changes in self-explanation

As an extension to the Chi et al. (1989) data, we were interested in three major points. 1) How do self-explanations change as performance improves during problem solving? 2) Do subjects have a specific strategy of using the latter worked examples as a feedback mechanism to the earlier problem solving rounds and does this affect their subsequent performance on the problem solving rounds? 3) Where in the worked examples is a subject most likely to engage in self-explanation behavior, and how does this likelihood to explain change across problem solving rounds?

Nine out of ten subjects showed an decrease in the amount of explanation statements for the second worked-out example. The total amount of explanation statements for 10 participants for the first worked-out example (SE1) was 163, or an average of 16.3 per subject. For SE2 and SE4 the explanation statements dropped to 96 (9.6 per subject) statements for SE2 and 76 (7.6 per subject) explanation statements for SE4. This change in explanation statements yielded a significant sign test of ($X2(1) = 6.4$, $p < .025$). A Wilcoxon matched-pairs signed-ranks test showed a significant difference between SE1 and SE2 on explanation ($t(10) = 10$, $p < .05$) as well as between SE1 and SE4 on explanation ($t(10) = 0$, $p < .005$). A significant difference was also found for overall word count. The Wilcoxon matched-pairs signed-ranks test yielded a significant effect for number of words for SE1 compared with SE2 ($t(7) = 0$, $p < .001$).

In general, while the overall trend in explanation statements decreases across the different worked-out examples as problem solving continues, learning is still continuing as evidenced by the improved performance of each subject across the rounds. Out of ten participants, the average improvement in score from the first problem solving round to the last problem solving round was 77 percent. The most improved subject (JE08) went from a score of 20 on the first round to a score of 100 by the last round of problem solving.

Across all the participants, the percentage of explanation statements, in relation to their total number of statements (which would include the extra categories of "read", "monitor" and "other") did not change across the self-explanation rounds. The percentage of all statements which were explanation statements, for 10 subjects, for SE1, SE2 and SE4 was 27%, 26%, and 27% respectively. However, there were some small and relatively consistent differences when comparing good versus poor participants. On SE1, good performing subjects had 30 percent explanation statements while poor performing subjects had 26 percent. By SE2, the gap widened slightly with the good students having 30 percent explanation statements while the poor students had 24 percent explanation statements. Finally, on SE4, this difference was still fairly consistent with the good students having 30 percent explanation statements and the poor subjects having 25 percent explanation statements.

Self-explanation strategies

In general, some participants appeared to use the first worked example (SE1) to provide a solid foundation for their subsequent problem solving rounds, which we termed the upfront strategy. This was evidenced by an apparent decrease in word count from SE1 to SE2. These learners appeared to expend less effort while examining the second worked example, as compared to the first, and consequently had a reduction in word count. However, other participants used SE2 as more of a SFG mechanism, relying on it to fill in any gaps they may have had in their knowledge which they may have still had after the first round of problem solving. We termed this the catch-up strategy. Again, this was evidenced by the increase in word count from SE1 to SE2. We hypothesized that if participants were using SE2 as a SGF mechanism, then they would show less of a reduction in word count than a subject who relied heavily on SE1. However, if someone is relying heavily on the first worked example to establish a foundation for subsequent problem solving, and not relying on SE2 to provide SGF to their problem solving episodes; then the change in word count from SE1 to SE2 should be large. More importantly, the number of explanation statements should show the same change in direction as was hypothesized for the word counts.

We found that changes in word count from the first worked-out example (SE1) to the second worked-out example (SE2) were correlated with performance in the hypothesized direction ($r(6) = .65$, $p > .05$). However, the more sensitive count of explanation statements, and using more participants, produced a small negative correlation in the opposite direction ($r(9) = \ .13$). Those subjects we had categorized as using the upfront strategy, based on their change in explanation statements, had a total combined score 47.5 questions answered correctly. Those subjects which were categorized as using the catch-up strategy, based on their change in explanation statements, had a total combined score of 51.3 questions answered correctly. So one would have to conclude that even though the changes in word counts were occurring in the hypothesized direction, changes in explanation statements, which must be considered a better indicator, were not occurring in the hypothesized direction.

The subjects which had the smallest absolute changes in word counts from SE1 to SE2 were subjects KB07, JE08, MT16 and MT11 respectively. These protocols were searched for examples of possible SGF examples, and examples were found for subjects KB07, JE08 and subject MT11. These appeared to be instances where the participant was referring back to the previous problem solving rounds while studying a later worked-out example. For this analysis, we concentrated specifically on the latter worked examples after the problem solving episodes (SE2 and SE4) and looked for any statements which made references to earlier problem solving episodes.

While examining the fourth worked example, on lines 30 to 32, subject KB-07 makes these statements:

30) Okay, add 'em up, you get Rx and get Ry
31) Woa, Woa, Woa, Oh.. so that's where you put in
32) but it's still positive 1.8 ft.

This subject is examining where certain positive and negative values came from and appears to realize at what part in the process of solving the equation that the values are actually necessary.

While examining the second worked example, on lines 77-84, subject KB-07 realizes:

77) X squared
78) Rx, Ry squared
79) So we're looking
80) Oh!!
81) So we're looking for this too
82) So this would be equal to 6.25, wait.

Participant KB-07 has realized that part of the process she had previously used did not include a necessary step. Hence the exclamation, "oh, so we are looking for this too".

Participant JE08, on lines 54 to 56 makes the statement:

54) The direction of R may be found using an inverse trigonometric function such as arctan
55) Now here is where I get lost.

Subject JE08 knows from previous problem solving rounds that there is a gap in her declarative knowledge which still has not be resolved, even by the time she gets to the second worked example.

While examining the second worked example, on lines 49 to 53, subject MT-11 makes these statements:

49) Ok, they find out where this was
50) So,..... and use this right angle
51) arc tan
52) adjacent over
53) hypotenuse ... Ah, it doesn't matter

The subject has realized that the example provides an alternative way to approach the problem from what the subject had previously been doing. The subject realizes, from studying the example, that a step the subject had previously taken while solving the problem "doesn't matter", and the example shows how the step can be eliminated.

So while the explanation statements did not decrease in the anticipated direction to show evidence of possible SGF for those subjects who we categorized as using the catch-up strategy, there was some indication within the protocols that SGF was taking place.

Model data

Results from our analysis will be used to build a cognitive model of self-explanation behavior. An expert level model of a subject solving vector arithmetic tasks has already been developed. The model solves a vector arithmetic problem by using the first worked-out example (SE1) as a guide. The model assumes expert performance in that the model knows what each next step is, and the model knows in what order to do each step. The model has four basic decision points as it precedes through the worked-out example. The model uses logical evaluations at each of the four steps to determine the information needed by the model to find any unknown variables, then it proceeds to the next step. The progression is linear, through the worked-out example, toward a solution.

To further analyze the data, and to further help us develop our cognitive model, we were interested in where in the worked-example participants were elaborating or doing self-explanation. Most participants proceeded in a linear fashion through the worked example. As a subject reached each line of the example, the number of explanation statements that occurred at that line were totaled. The highest points of explanation occur at lines 6, 7, 12, 14, 17 and 21. The four major decision points of the previously described expert model occur at lines 6, 7, (one decision) 11, 12, (one decision) 17 and 21. So the model seems to be making critical decisions at the appropriate points in the worked-out example. These points appear to be occurring primarily at mathematical areas of the worked example (formulas) and not textual (written) sections.

The total number of explanation statements for seven subjects were also totaled for each example (not just SE1 on which the model was based upon) and the totals were then compared. As with SE1, subjects appear to be concentrating on the formulas of the worked examples much more than the textual components of the worked examples. Also, the subjects studying SE2 and SE4 tended to do a great deal of explaining near the end of the round. For SE2, which had 10 lines, the largest amount of explanation statements occurred at lines 10, 7 and 6 receptively. For SE4, which had 21 lines, the largest number of explanation statements occurred at lines 11, 14, and 15 (which were tied) and lines 21, 10 and 2 (which were also tied). A direct comparison of SE1 with SE4 was possible since these worked examples were highly similar to each other. Generally, in percentage terms, the amount of explanation statements decreases from SE1 to SE4. But also, there was a tendency for explanation statements to occur later in the worked example for SE4 as opposed to SE1.

While the current version of the model can account for the areas where a participant is most likely to self-explain, it does not account for the tendency of subjects to do most of their explaining near the end of the worked example. While the model does show and increase in explanation statements near the end of the example, it is not proportional to the amount shown by the subjects.

Further development of the model needs to take place in two specific areas. First, in order to account for our empirical data, the model needs to show explanation statements at the very end of the worked example. Secondly, model does not simulate what we have called self-generated feedback, and does not account for the utterances we identified in our protocols as SGF. SGF is an important change in self-explanation behavior that does not occur during the first SE round but rather it is more likely to occur in the latter rounds. This change in self-explanation behavior must be addressed by our model of the self-explanation effect.

Discussion

It is clear that the self-explanation effect is a powerful phenomena in the study of examples. Replication of the self-explanation effect has been conducted by other researchers (Renkl, 1997). The majority of our data shows a clear indication of the self-explanation effect as it was first defined by Chi et al. (1987). However, we did find differences from the Chi et al. (1987) study in the amount of negative monitoring statements and performance during problem solving rounds. Our data is consistent with that of the later Renkl (1997) study, therefore it would be difficult for us to conclude that increased negative monitoring is one of the underlying features of the self-explanation effect.

Beyond replication, we were interested in examining how self-explanations change as performance improves during problem solving. We found that self-explanation decreases as problem-solving performance increases. This would be expected if self-explanation was occurring in order to fill in gaps in their declarative knowledge. If subjects were using self-explanations to fill in gaps in their declarative knowledge base, then as their performance improved, then there should be less need to do any self-explanation.

Next, we addressed the question of whether subjects have a specific strategy of using the latter worked examples as a feedback mechanism to the earlier problem solving rounds and does this affect their subsequent performance on the problem solving rounds. What we found were examples in the protocols of subjects using the latter worked examples as a feedback mechanism to their earlier problem solving episodes. This occurred in subjects who had the smallest change in word count from SE1 to SE2. However, we could not find any reliable changes in performance from the subjects using this strategy.

Finally, we examined areas in the worked example where a subject was most likely to engage in self-explanation behavior, and how does this likelihood to explain change across problem solving rounds. We also found that subjects tended to do a great deal of explaining at the end of the worked-examples, especially for latter worked examples (SE2 and SE4). This was also supported by an analysis of the best performing subject (AM19), who seemed to do most of her explaining at the end of the example, while the worst performing subject (JE08) did not. We construed that participants will frequently reflect at the end of a problem solving episode, even when no gap in their declarative knowledge has been identified. Apparently a significant amount of learning could occur during these reflective, non-impasse periods, and this assumption is supported by the data of our best performing subject (AM 19), as well as our aggregate data. The fact that the best performing subject uses this strategy might reflect differences in good versus poor performance of participants. Perhaps good students tend to be more reflective after problem solving, while poor performing students do most of their problem solving only when they encounter a gap in their knowledge. However, the important assumption here is that gaps are not present when the learner has reached the end of the worked example. Obviously, a counter argument could be made that this has not been proven to be the case, and more research needs to be done to clarify this issue. However, if one accepts the assumption that a gap in knowledge is not likely to be identified after an example has been completed, then this is consistent with VanLehn (1992) findings that not all learning occurs at impasses.

Other models imply that there are strategy differences within participants who self-explain. The Cascade model (VanLehn, 1992) uses strategy differences to distinguish between good and poor learners by forcing the simulation of good learners to rederive an example's solution, while the simulation of poor learners never attempts any new derivations. Their model also found that this strategy caused the good learner model to acquire more knowledge while solving problems than the poor learner model. Our data indicates that participants continue to learn even while self-explanation behavior decreases, and it would seem that the Cascade model can account for this aspect of our data.

Our computational cognitive model simulated subjects' performance during the first worked-example. We found that our model of SE1 was consistent with our empirical data. The model engaged in self-explanation behavior in the same areas where participants were most likely to engage in self explanation behavior. However, our model does not account for the large number of self-explanation statements that occur at the end of the problem solving episodes. The model also needs to incorporate specific instances of SGF, which we had identified in the latter protocols of subjects (SE2 and SE4). These are two important aspects of the change in self-explanation behavior, which will be addressed in future versions of the model.

References

Anderson, J. R., Farrell, R., & Sauers, R. (1984). Learning to program in LISP. *Cognitive Science, 8*, 87-129.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 18*, 145-182.

Ferguson-Hessler, M. G. M., & DeJong, T. (1990). Studying physics texts: Differences in study processes between good and poor performers. *Cognition and Instruction, 7*, 41-54.

Pirolli, P., & Anderson, J. R. (1985). The role of learning from examples in the acquisition of recursive programming skills. *Canadian Journal of Psychology, 39*, 240-272.

Recker, M. M., & Pirolli, P (1995). Modelling individual differences in students' learning strategies. *The Journal of the Learning Sciences, 4*, 1-38.

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

VanLehn, K. (1986). Arthmetic procedures are induced from examples. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics.* Hillsdale, NJ: Erlbaum.

VanLehn, K. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences* 2(1), 1-59.