What mediates the self-explanation effect? Knowledge gaps, schemas or analogies?*

Kurt VanLehn
Learning Research and Development Center
University of Pittsburgh
Pittsburgh, PA 15260
VanLehn@cs.pitt.edu
(412)624-7458

Randolph M. Jones
Artificial Intelligence Laboratory
University of Michigan
Ann Arbor, MI 48109-2110
rjones@engin.umich.edu
(313)764-0683

Abstract

Several studies have found that learning is more effective when students explain examples to themselves. Although these studies show that learning and self-explanation co-occur, they do not reveal why. Three explanations have been proposed and computational models have been built for each. The *gap-filling* explanation is that self-explanation causes subjects to detect and fill gaps in their domain knowledge. The *schema formation* explanation is that self-explanation causes the learner to abstract general solution procedures and associate each with a general description of the problems it applies to. The *analogical enhancement* explanation is that self-explanation cause a richer elaboration of the example, which facilitates later use of the example for analogical problem solving. We claim that, in one study at least, gap filling accounts for most of the self-explanation effect.

Introduction

Chi et al. (1989) found that students learn more when they explain instructional material to themselves. Chi et al. took protocols of students learning college physics by studying worked example problems. Some subjects simply read the solutions with hardly a pause, while others explained each solution line by deriving it from physics principles and preceding lines. The students who self-explained the example solutions learned more. This co-occurrence, called the *self-explanation effect*, has now been observed when students study other problem solving task domains (Pirolli & Bielaczyc, 1989; Ferguson-Hessler & de Jong, 1990) and when students study declarative subject matter, such as descriptions of the human circulatory system (Chi et al., 1991; Pressley et al., 1992). Training experiments indicate that subjects can be taught, or even just prompted, to self-explain, and when they do, their learning rate increases (Chi et al., 1991; Bielacsyc & Recker, 1991). Although the self-explanation effect appears ubiquitous and educationally important, the studies cited so far only demonstrate that more effective learning co-occurs with self-explanation. They do not reveal why.

In analyzing the data from the original Chi et al. (1989) study, we came to believe that self-explanation caused students to uncover gaps in their knowledge and fill them. When students did not explain the examples to themselves, they seem not to discover their ignorance, so their knowledge gaps persisted and caused errors. This hypothesis, here called *gap-filling*, is consistent with impasse-driven learning (Vanlehn, Jones & Chi, 1991; Newell, 1990), failure-driven learning (Schank, 1986) and other learning mechanisms.

*This research was supported by grant N00014-88-K-0086 from the Cognitive Sciences Division of the Office of Naval Research. We thank Micki Chi and Joel Martin for their comments on the first draft.

However, gap-filling is not the only possible explanation for the phenomena. Pirolli and Anderson (1985) suggested that studying an example creates a richer, more elaborated understanding of the example, which increases the effectiveness of analogical problem solving by increasing the chances that the student will both retrieve a useful example and adapt it successfully to solve the problem at hand. Reimann (in press) suggested that studying an example creates a problem schema: a moderately general plan for solving problems of this type. Pirolli and Reimann are both building computational models of the self-explanation effect.

The computational sufficiency of the gap-filling hypothesis has already been demonstrated by implementing a computational model, called Cascade, that learns by filling gaps while studying examples and solving problems. When Cascade is directed to thoroughly explain physics examples, it learns more and exhibits several other behaviors characteristic of self-explainers (VanLehn, Jones & Chi, 1991). When an individual student's protocol is fit by directing Cascade to self-explain exactly the same example lines that the student self-explained, Cascade behaves much like the student during problem solving (VanLehn & Jones, 1993). However, these computational studies show only that gap-filling suffices to model the self-explanation effect. It could be that other types of learning, such as schema acquisition and enhanced analogical problem solving, might be just as good at modeling the data. These hypothesized processes are not mutually exclusive. All three or any combination could concurrently mediate the self-explanation effect.

The analyses presented below indicate that, in this study at least, self-explanation enhances learning mostly by filling knowledge gaps, a little by improving schema acquisition, and hardly at all by facilitating analogical problem solving.

The Chi et al. study

The protocols for these analyses came from the Chi et al. (1989) study of physics learning. The subjects, 9 college students, reviewed basic mathematical and kinematics material until they could pass criterion tests on it. They studied chapter 5 from Halliday and Resnick (1981), a popular physics textbook, which introduces Newtonian mechanics and gives a brief introduction to solving physics problems. The subjects then talked aloud as they studied 3 worked example problems from Halliday and Resnick, then solved 19 problems.

Three schemas (solution procedures) were required for solving the problems and understanding the examples. All the examples and 16 of the problems require the *force* schema, which has 4 steps: decide which objects to focus on (the "bodies"), find all the forces acting on the bodies, instantiate Newton's law ($F = ma$) and other relevant principles, then solve the resulting system of equations. One problem must be solved with the *scalar equation* schema, which applies means-ends analysis to a set of scalar equations (as in solving high-school algebra word problems). Two problems require using the *kinematic* schema, which interrelates acceleration, velocity, distance and time. In all the problem solving analyses, we coded only the 16 problems that used the force schema, as that was the topic of chapter 5. The kinematics and scalar equation schemas were taught earlier.

From the protocols taken during example studying it was easy to tell which of the subjects preferred to self-explain the solutions, and which preferred to read them casually. Using the protocol coding of Chi et al. (1989), 3 subjects were clearly high self-explainers because they uttered 75, 55 and 41 self-explanations each. The other 6 subjects uttered 19 or fewer self-explanations each, so we call them the low self-explainers.

Just as Chi et al. found, with a slightly different method, the high self-explainers learned more than the low self-explainers. While solving the 19 problems, the high self-explainers got an average of 16.4 problems correct, while the low self-explainers got an average of 10.5 problems correct, a significant difference ($p < .05$, $t(7) = 3.14$). Because all subjects had similar pre-test scores (Chi & VanLehn, 1991), the high self-explainers must have learned more physics than the low self-explainers. This result shows only that self-explanation co-occurs with more effective learning. It doesn't say how that learning was accomplished.

An error analysis

A direct way to find out the source of the self-explanation effect is to classify the errors of both high and low self-explainers, and see what types of errors were reduced by self-explanation. This should indicate which learning processes were engaged by self-explanation.

Working from the protocols, 69 errors were identified and classified (see VanLehn & Jones, in prep.) using the following categories.

- *Gap errors.* The error was caused by lack of knowledge of a physics principle or concept. For instance, some subjects did not know that when an object is supported by a surface, the surface exerts a force on the object (called the normal force).

- *Inappropriate analogies.* Subjects were classified as using analogical problem solving if they opened the textbook to an example or explicitly referred to an example during problem solving. Errors were caused both by referring to inappropriate examples and using appropriate examples in an incorrect manner.

- *Inappropriate schemas.* Although the force schema was required for correct solution of all problems, some subjects used the scalar equation schema on particularly simple problems. For instance, one problem asked for the force exerted by a 160-pound fireman on a pole as he slide down it at 10 feet per second. Some subjects immediately invoked $F = ma$ because F is sought and a is given. They subgoaled to get the mass, m, from the given weight. The force schema would begin by choosing the fireman as the body, then noting that there are two forces acting on the body, a frictional one (whose magnitude is sought) and a gravitational one (whose magnitude is given).

- *Mathematical errors.*

- *Miscellaneous errors.*

Table 1: Mean errors per subject

Error type	high SE	low SE	t(7)
Gap errors	*0.0	*4.5	2.22
Inappropriate schemas	0.7	1.7	1.53
Inappropriate analogies	1.3	1.8	0.55
Math errors	0.3	0.7	0.88
Miscellaneous errors	1.0	1.3	0.41
Totals	*2.7	10.2	2.87

Table 1 shows the mean errors for each error category. For all classifications, the high self-explainers had fewer errors than the low self-explainers. However, the difference was significant only for one category, gap errors ($p < .05$ one-tailed). Moreover, this category accounts for most (60%) of the differences in error rates between the high and low self-explainers. These results strongly suggests that self-explanation encourages some kind of gap-filling process. There are weak trends suggesting that schema learning and analogical enhancement might also be encouraged by self-explanation, so we will examine those hypotheses more thoroughly in the next two sections.

Schema acquisition and selection

Although there are many definitions of "schema," we use the term to mean the general procedure that textbooks teach for solving classes of equations. Chapter 5 of Halliday and Resnick explicitly teaches a 4-step procedure for solving force problems. Another approach (schema) for solving physics problems, which has often been noted in novice behavior, is to treat $F = ma$, $w = mg$ and other vectoral equations as if they were scalar equations.

All subjects used both schemas at least once. One subject used the scalar equation schema only on one problem, problem 14, which is so

Table 2: Number of problems solved

Method	high SE (N=3)	low SE (N=6)
Scalar equation schema	4.5	13
Force schema	3.5	3
Analogical problem solving	1	1
Unclassifiable	0	1
Total	9	18

simple that the scalar equation schema is arguably the correct one to use with it. The other subjects used the scalar equatin schema more than once. Apparently, subjects already knew the scalar equation schema (from high-school algebra, perhaps) and had no trouble acquiring the force schema from the text of chapter 5.

However, some errors were caused by subjects using the scalar equation schema on problems that required the force schema. Immediately following problem 14 were two other problems (one was the fireman problem described earlier) that require the force schema for a correct solution even though they appear just as simple as problem 14. Problem 19 was also an apparently simple problem that required the force schema. We coded the type of method used by subjects on these 3 problems (see VanLehn & Jones, in prep. for details). In addition to employing one of the two schemas, some subjects used analogical problem solving and one subject used a method that we could not classify. Table 2 shows the resulting distribution of methods for the high and low self-explainers. Although the high self-explainers chose the force schema more of often than the low solvers, the difference was not signicant ($\chi(3) = 4.88$). This indicates that self-explanation does not co-occur with better skill at selecting schemas.

Moreover, no subject used the scalar equation schema on any problems other than these 3 and problem 14. Even if there were a tendency for low self-explainers to be worse at schema selection, their poor selections on these few 3 problems could not account for the relatively large difference in scores between the high and low self-explainers. It appears that skill at selecting schemas is not responsible for the large performance differences caused by self-explanation.

Analogical problem solving

Several investigators (e.g., Pirolli & Anderson, 1985) have suggested that self-explanation of an example creates a more elaborated understanding and memory for the example, and this causes students to more often retrieve appropriate examples and apply them correctly. To test this, we focused on 8 subjects who routinely tried to retrieve an example before solving a problem, and on 12 problems that were isomorphic to example problems. We classified each of the 96 ($8 \cdot 12$) retrieval attempts as successful or not. Of the 24 retrieval attempts by 2 high self-explainers, 92% were successful. Of the 72 retrieval attempts by 6 low self-explainers, 83% were successful. This is not a significant difference ($t(5)=0.53$), because all subjects were highly successful (a ceiling effect). Restricting the analysis to the 2 problems that were most likely to cause retrieval errors also showed no significant differences between high and low self-explainers ($t(5)=0.85$).

In order to determine whether self-explanation facilitates analogical application, we examined all cases where subjects explicitly referred to an example and tried to use it to help them solve their current problem. We classified an analogical application as successful if the subject correctly mapped objects in the example to objects in the problem, and correctly imported relationships (usually equations) from the example to the problem. Of the 22 analogical applications by high self-explainers, 82% were successful. Of the 55 analogical applications by low self-explainers, 80% were successful. This different was not significant ($\chi(5) = .033$), so we conclude that self-explanation did not help student apply analogous examples to solve problems. Restricting the analysis to the 2 problems most likely to cause application errors also showed no significant differences ($t(5)=0.0$).

The trends were consistent with hypothesis that self-explanation helps analogical problem solving, and the trends could reach significant if more subjects were studied. However, the small difference between the two groups' success at analogical problem solving could not explain the large difference in their learning rates. It appears that the high self-explainers are no better at analogical problem solving than the low self-explainers.

What happens to gaps

According to the gap-filling hypothesis, incomplete instruction should cause incomplete knowledge, which in turn should cause errors until the gap is filled by learning. To test this prediction, we located 9 gaps in the textbook and examined every place in the protocols where these 9 pieces of knowledge were relevant. For each of the 9 subjects, we classified each of the 9 gaps into one of the following patterns (the numbers in parentheses are the number of cases of each pattern).

- (6) The subject never used the knowledge correctly, and these omissions caused errors.
- (17) At the first place where the missing knowledge could be used, the subject's protocol showed clear signs of learning followed by a correct usage of the knowledge. The knowledge was used correctly on most subsequent occasions.
- (14) At the first place where the missing knowledge could be used, the subject used it correctly but did not show verbal signs of learning. The knowledge was used correctly on most subsequent occasions.
- (44) The subject avoided all places where the knowledge could be used. During example studying, the subject did not self-explain lines where the knowledge could be used. During problem solving, the subject used analogical problem solving to avoid opportunities for using the knowledge.

The cases in the first three categories support the gap-filling hypothesis. Although the 14 cases in the fourth category could be interpreted as disconfirming the hypothesis, it is more likely that the students had learned the knowledge before reading the text or that they learned the knowledge at the first place where it could be used but gave no signs of that learning in their protocol. Although the naturalistic design of this study does not allow disconfirmation of the gap-filling hypothesis, the hypothesis receives some support from the fact that there were 17 clear cases of learning at exactly the placed predicted by the hypothesis, and that there were errors at many of the predicted places.

The biggest surprise from this analysis was that subjects often avoided the places where the missing knowledge could be used, most frequently by employing analogical problem solving instead of trying to solve the problem from first principles. This suggests that analogical problem solving tends to preserve ignorance. If subjects had not been allowed to refer to the examples, perhaps they would have discovered the gaps in their understanding and found ways to fill them. This interpretation of the data, which has strong pedagogical implications, is explored in another paper (Vanlehn & Jones, in press).

Discussion

The primary research question was to determine the type of learning that causes self-explanation to improve post-test scores. By analyzing errors, we discovered that approximately 60% of the difference in post-test error rates between high and low self-explainers is due to knowledge gaps. Knowledge gaps tend to remain and cause errors among low self-explainers, whereas they tend to disappear among high self-explainers.

Presumably, self-explanation causes students to discover their ignorance and do something about it. This assumption was checked by protocol analysis of all the places where a gap could be detected. We found 17 clear cases of a gap being filled by learning, and another 14 cases where the gap was either filled by learning or did not exist because the student learned the target knowledge before the study. As predicted, 11 of the 17 clear cases of learning occurred with the high self-explainers.

Another analysis checked whether analogical problem solving was aided by self-explanation. The results indicate, somewhat surprisingly, that high and low self-explainers were equally good at finding appropriate examples and at applying analogies to those examples to help them solve problems. Thus, it is unlikely that differential success at analogical problem solving can account for the self-explanation effect.

Self-explanation apparently has no effect on whether students acquire the force schema, which is not unreasonable given that it is taught quite explicitly in the text. Moreover, self-explanation did not significantly improve the students' skill at deciding whether to use the force schema or another, simpler schema. Neither the acquisition of schemas nor the acquisition of skill at selecting schemas contributed much to the self-explanation effect.

References

Bielaczyc, K., & Recker, M. M. (1991). Learning to learn: The implications of strategy instruction in computer programming. In L. Birnbaum (Ed.), *The International Conference on the Learning Sciences* (pp. 39–44). Charlottesville, VA: Association for the Advancement of Computing in Education.

Chi, M. T. H., de Leeuw, N., Chiu, M., & LaVancher, C. (1991). *The use of self-explanations as a learning tool.* Manuscript submitted for publication.

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*, 13, 145-182.

Chi, M.T. H. & VanLehn, K. (1991) The content of physics self-explanations. *Journal of the Learning Sciences*, 1(1), 69-106.

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

Halliday, D. & Resnick, R. (1981). *Fundamental of physics, Second edition.* New York: Wiley.

Newell, A. (1990). *Unified Theories of Cognition.* Cambridge, MA: Harvard.

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.

Pirolli, P. & Bielaczyc, K. (1989). Empirical analyses of self-explanation and transfer in learning to program. In *Proceedings of the 11th Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Lawrence Erlbaum.

Pressley, M., Wood, E., Woloshyn, V.E., Martin, V., King, A. & Menki, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational psychologist, 27*, 91-109.

Reimann, P. (in press). Modeling active, hypothesis-driven learning from examples. In E. De Corte, M. Linn, H. Mandl, & L. Verschaffel (Eds.), *Computer-based learning environments and problem solving.* Berlin: Springer.

Schank, R.C. (1986). *Explanation patterns: Understanding mechanically and creatively.* Hillsdale, NJ: Lawrence Erlbaum Assoc.

VanLehn, K., Jones, R.M. & Chi, M.T.H. (1991). A model of the self-explanation effect. *Journal of the Learning Sciences*, 1(1), 69-106.

VanLehn, K. & Jones, R.M. (1993). Learning by explaining examples to oneself: A computational model. In S. Chipman & A. L. Meyrowitz (Eds.) *Foundations of knowledge acquisition: Cognitive models of complex learning.* Boston: Kluwer Academic Publishing.

VanLehn, K. & Jones, R.M. (in preparation). How self-explanation helps learning and analogical problem solving hurts it.

VanLehn, K. & Jones, R.M. (in press). Better learners use analogical problem solving sparingly. In P. Utgoff (Ed.), *Proceedings of the 1993 International Machine Learning Conference.* Los Altos, CA: Morgan-Kaufman.