Learning by Explaining to Oneself and to Others

Rolf Ploetzner[1], Pierre Dillenbourg[2], Michael Preier[1] and David Traum[3]

[1]University of Freiburg, Germany

[2]University of Geneva, Switzerland

[3]University of Maryland, U.S.A.

Authors' addresses:

Rolf Ploetzner, University of Freiburg, Department of Psychology

Niemensstr. 10, D-79085 Freiburg, Germany

Tel.: +49 / 761 / 203 2484

Fax: +49 / 761 / 203 2496

E-Mail: ploetz@psychologie.uni-freiburg.de

Pierre Dillenbourg, University of Geneva, TECFA

9, Route de Drize, CH-1227 Carouge, Switzerland

Tel.: +41 / 22 705 96 93

Fax: +41 / 22 343 89 24

E-Mail: Pierre.Dillenbourg@tecfa.unige.ch

Michael Preier, University of Freiburg, Department of Psychology

Niemensstr. 10, D-79085 Freiburg, Germany

Tel.: +49 / 761 / 203 2484

Fax: +49 / 761 / 203 2496

E-Mail: preier@mibm.ruf.uni-freiburg.de

David Traum, UMIACS, University of Maryland,

College Park, Maryland 20742, U.S.A.

Tel.: +1 / 301 / 405 1139

Fax: +1 / 301 / 405 6707

E-Mail: traum@cs.umd.edu

Running head: LEARNING BY EXPLAINING

Abstract

One important source for the acquisition of knowledge, especially of declarative knowledge, is the construction, transmission and comprehension of explanations. Two distinctive settings in which explanations are constructed are *self-explanation*, in which a single individual explains to himself i*nteractive explanation* in which several individuals mutually and interactively explain to each other. While providing explanations to oneself might lead to the construction of additional knowledge by the explaining individual, providing explanations to each other involves forms of collaborative learning. To better understand the differences between and potential benefits of explaining to oneself and explaining to others, we compare empirical research on both kinds of explanation. The comparison is guided by three main goals . The first goal is to describe the learning that takes place during the construction of self-explanations. The second goal is to reveal how the construction of explanations as well as the learning changes if one moves from a self-explanation setting to a more interactive explanation setting. With respect to interactive explanation settings, we particularly focus on learning by teaching. The third goal is to uncover differences in the learning between those individuals who construct explanations and those who receive them. Essentially, since learning by teaching has repeatedly been found to promote learning, this gave rise to the assumption that the beneficial effects of such arrangements are due to the successive construction of explanations. In particular, it has been hypothesized that explaining to others might be even more advantageous than explaining to oneself. However, so far empirical evidence does not strongly support this hypothesis.

Learning by Explaining to Oneself and to Others

Introduction

When two individuals collaborate, they often have to justify themselves to each other, to explain what they are doing and why they are doing it. Intuitively, these efforts should be related to the learning that is frequently observed during collaboration. For instance, these justifications may lead individuals to make assumptions explicit which would otherwise remain tacit. When an individual explains to a second individual, learning might take place by both individuals:

- The amount of learning by the individual who provides explanations seems to be related to the cognitive activities necessary for constructing and presenting explanations (e.g., Bargh & Schul, 1980; Webb, 1989).
- The amount of learning by the individual who receives explanations seems to be related to variables such as how relevant, understandable and elaborated the explanations are (e.g. Webb, 1989).

In this chapter, we focus on learning due to constructing and providing explanations to oneself or to others. With respect to the settings in which explanations might be constructed, five different levels of interactivity can be distinguished:

1. Explaining to oneself: During the attempt to understand something (e.g., instructional material), an individual might try to explain it to him or herself. While self-explanations may frequently be expressed silently, only self-explanations expressed aloud can been studied experimentally.
2. Explaining to a passive and anonymous listener: An individual might explain to somebody he/she does not know and who just listens. In such a setting it might be investigated, for example, whether the explainer better monitors the construction of

explanations as compared to a self-explanation setting.

3. Explaining to a passive listener: An individual might explain to somebody he/she knows and who just listens. In such a setting it might be examined whether individuals tailor the construction of explanations to specific listeners.
4. Explaining to somebody who responds in a constrained way: An individual might explain to somebody who responds to his/her explanations in a constrained way. For instance, the individual who receives the explanations might only indicate his/her understanding or non-understanding. In such a setting it might be scrutinized how the listener's responses affect the construction of explanations.
5. Mutually explainantion: Two individuals might mutually explain to each other without any imposed constraints. In this case, explanation is no longer something that is exclusively directed from one individual to a second, but rather corresponds to a process in which two individuals attempt to negotiate and, at least partially, share their understanding of the domain under consideration.

Though only mutual explanation, as described at Level 5, accounts for the complexity of many interactions in collaborative problem solving and learning, the investigation of how explanations are constructed at Level 1 to 4 may nevertheless help to disentangle various processes which might take place almost simultaneously at Level 5. Furthermore, fully unconstrained interactions might not necessarily result in the construction of appropriate explanations. Educators, for instance, very often constrain interactions to make them more structured and beneficial (e.g., Aronson, Bridgeman & Geffner, 1978; for an overview see Knight & Bohlmeyer, 1990).

To better understand the differences between and potential benefits of explaining to oneself and explaining to others, we compare empirical research on both kinds of explanation at the different levels of interactivity described above. The comparison is guided by three main

goals, as follows. The first goal is to describe the learning that takes place during the construction of self-explanations. The second goal is to reveal how the construction of explanations as well as the learning changes if one moves from a self-explanation setting to a more interactive setting. With respect to more interactive explanation settings, we particularly focus on learning by teaching. The third goal is to uncover differences in the learning between those individuals who construct explanations and those who receive them.

Because our comparison is based on a review of various empirical studies and therefore on various perspectives on what makes up explanations, we do not attempt to provide a concise definition of explanatory activities. Instead, we refer to many activities as explanations which in other contexts might be referred to as elaborations or argumentations, for example.

The rest of the chapter is organized as follows. In the next section, we consider self-explanations and describe empirical investigations of their relation to individual problem solving. In addition, we describe *Cascade*, a cognitive simulation program, which models a possible process by which self-explanations can be constructed during individual problem solving and how they might lead to the acquisition of new knowledge. Next, we consider empirical investigations of explanations in collaborative problem solving and how they might facilitate learning. We conclude with a discussion of (missing) differences in the learning between self-explanation settings and more interactive settings .

Explaining to Oneself

Revealing Self-Explanations

How do students learn when they read textbooks, study examples and solve problems? With this question in mind, Chi, Bassok, Lewis, Reimann and Glaser (1989) investigated nine high school students as they worked on a common college physics textbook (Halliday & Resnick, 1981). Within this investigation, the students carried out the following steps. They read the first three chapters of the textbook which were on various physics concepts such as time, distance, velocity and acceleration in the context of motion in one dimension. Afterwards, the students read the textbook's chapter on Newtonian mechanics. The different chapters of the textbook had to be studied until a criterion test had been passed.

After reading the textbook's chapters, the students studied three worked-out examples and solved 25 problems on Newtonian mechanics. A worked-out example consisted of a problem along with a solution that was demonstrated for the students. Figure 1 shows one of the worked-out examples used in the study by Chi et al. (1989). Typically, a worked-out example does not specify all the information that underlies the problem's solution. How does a student know, for instance, that the forces shown in Figure 1 are all the forces acting on the block?

While the students worked on the examples and problems, they were asked to verbalize everything that came to their mind. These verbalizations were tape recorded and later transcribed. Before and after reading the textbook's chapters and working on the examples and problems, the students had to take a pre- and posttest, respectively.

Insert Figure 1 about here

On the basis of the students' scores on problem solving, they were divided post hoc into two groups: the four students with the highest scores and the four students with the lowest scores. The data of the student closest to the median was omitted. Since the high and the low scoring students performed on average equally well on the pre-test, the high scoring students seemed to have learned more than the low scoring students. To gain more insight in what distinguishes the high scoring students from the low scoring students, the students' verbalizations were classified according to various classes of statements by means of a

protocol analysis. The main differences that were revealed by the protocol analysis are:

- the high scoring students tried more often than the low scoring students to explain the different solution steps in the worked-out examples to themselves,
- during example studying and problem solving, the high scoring students were more accurate at assessing their own understanding than the low scoring students and
- during problem solving, the high scoring students referred back to the examples less often than the low scoring subjects. In particular, while the high scoring students took advantage of only selected aspects of the examples, the low scoring students frequently re-read the complete examples.

Chi et al. (1989) named the students' attempts to explain the solution steps in the examples to themselves "self-explanations". Table 1 presents various examples of self-explanations as observed in the study by Chi et al. (1989). By means of self-explanations, students refined or generalized the conditions under which solution steps were taken, explicated the meaning of solution steps or extrapolated the consequences of solution steps, for example (see also Chi & VanLehn, 1991). Thus, on the basis of self-explanations, students explain parts of the application domain to themselves and possibly improve their understanding of the domain.

Insert Table 1 about here

The findings of Chi et al. (1989) and Chi and VanLehn (1991) suggest that self-explanations are constructive cognitive activities which potentially lead to the modification of already available pre-knowledge and the construction of new knowledge. However, could understanding be improved by systematically initiating the construction of self-explanations? Chi, de Leeuw, Chiu and LaVancher (1994) performed an empirical investigation to attempt to answer this question.

Within the investigation of Chi et al. (1994), two groups of students read a slightly shortened section on the human circulatory system out of a common biology textbook (Towle, 1989). Before and after reading the textbook's section, the students had to take a pre- and posttest, respectively. These tests were made up of four different sets of items:

1. items referring directly to single pieces of information out of the textbook,
2. items requiring different pieces of information out of the textbook to be integrated,
3. items demanding the induction of new information not provided in the textbook and
4. items requiring the application of information not in the textbook to answer questions about human health.

One group of students, the experimental group, was asked to read each sentence of the textbook's section out loud and to explain what it meant to them. The students were instructed to identify new information, to describe how it relates to what they already knew and to explain whether it gives them new insights or raises new questions. If required, an experimenter reminded the students to explain what they read. The other group of students, the control group, was not asked to explain the textbook's section. Instead, in order to keep the time spent on the section approximately constant in both groups, they were asked to read the section twice.

From the pretest to the posttest, the number of correct answers produced by the experimental group increased significantly more than the number of correct answers produced by the control group. This was especially true with respect to items out of the third and fourth set of test items. These test items required inferences to be drawn and new information to be related to one's own pre-knowledge. It thus appears that understanding can be (partially) improved by systematically initiating the construction of self-explanations.

It might be questioned, however, whether the students' activities in the studies of Chi et

al. (1989) and Chi et al. (1994) are appropriately denoted as "self-explanations". Essentially, self-explanations as observed in the study of Chi et al. (1989) and as initiated in the study of Chi et al. (1994) are explanations to an experimenter who is (repeatedly) asking for them, though in a constrained way. It remains an open question whether explanations directed (exclusively) to ourselves and explanations directed to others are the same. For instance, are explanations directed to ourselves as complete and consistent as explanations directed to somebody else? Or do we frequently construct only partial self-explanations because nobody else needs to understand them? In order to mentally prepare for important events such as examinations, it is often helpful to imagine the situation to come and, in particular, to imagine how one would argue in this situation. This suggests that we might adapt our explanations even when the listener is merely imagined.

A Cognitive Simulation Model of Constructing Self-Explanations

To better understand why and how self-explanations lead to the construction of new knowledge, VanLehn, Jones and Chi (1992) developed and implemented the cognitive simulation program Cascade as a knowledge-based system. In order to formalize the model's initial physics knowledge, they conducted a task analysis. The task analysis comprised two main steps. In the first step, the physics knowledge has been formalized in such a way that it was sufficient for solving all but 2 of the 25 problems which had been used in the study of Chi et al. (1989). In the second step, it was determined for each piece of physics knowledge whether it could be aligned with information given in the textbook chapters which had been used in the study of Chi et al. (1989). Surprisingly, only about half of the physics knowledge sufficient to solve the problems which had been posed by Chi et al. (1989) was found to be addressed in the textbook chapters. This knowledge made up Cascade's initial knowledge of physics.

On the basis of its initial physics knowledge, Cascade models two abilities: the explanation of worked-out examples and the solution of problems. A worked-out example is made up of a problem along with a solution. However, such a solution is hardly ever complete. Usually only the main solution steps are demonstrated by the example. The explanation of examples as well as the solution of problems is modeled in two different ways in order to account for the observed differences between the high scoring students and the low scoring students in the study of Chi et al. (1989).

The explanation of worked-out examples by high scoring students is modeled in the following way. The model attempts to derive each solution step provided within an example by means of the physics knowledge available to the model. Thus, in Cascade, the construction of an explanation corresponds to the construction of a formal proof. If a derivation can be found, then it is saved for later use during problem solving. If only a partial derivation can be found, then the model tries to complete it by taking advantage of general background knowledge such as commonsense knowledge as well as domain-specific and general heuristics. Whenever a derivation can successfully be completed by means of background knowledge, new physics knowledge is constructed which encodes the applied background knowledge in a form constrained according to the information given in the example. Thus, under these conditions Cascade realizes a form of explanation-based learning as it has been conceptualized by Mitchell, Keller and Kedar-Cabelli (1986).

If the derivation of a solution step cannot be completed by applying background knowledge, then it is concluded that such a solution step always has to occur within the respective context. Afterwards, new physics knowledge is constructed which encodes the solution step and the context in which it occurred in a generalized form.

The explanation of worked-out examples by low scoring students is modeled by merely saving each solution step as it is provided in the example. The solution of problems in Cascade

relies on the same mechanisms as the explanation of examples. The model attempts to derive a problem's solution by means of physics knowledge as well as general background knowledge available to the model. In particular, the construction of new physics knowledge might take place during the explanation of worked-out examples as well as during the solution of problems.

Problems can also be approached by referring back to the examples. In Cascade, the use of examples by high scoring students is modeled by taking advantage of the derivations which have been saved during the explanation of examples. In order to realize a solution step with respect to a posed problem, the physics knowledge which enabled the derivation of an analogous solution step in the example is (re-) applied. In contrast, the use of examples by low scoring students is modeled by the attempt to simply substitute the information given in the solution steps of the examples for the information given in the posed problems. In many cases, however, the problems posed in the study of Chi et al. (1989) cannot be solved by means of such simple substitution processes.

When Cascade simulates how high scoring students explain examples and solve problems, then the model successfully solves all 23 problems which have been taken into account during the task analysis. Furthermore, under this condition Cascade learns 23 new pieces of physics knowledge: 8 during self-explanation and 15 during problem solving. When Cascade simulates how low scoring students explain examples and solve problems, then only nine problems can be solved successfully. Under this condition the model learns only three new pieces of physics knowledge.

Cascade simulates how explaining to oneself can lead to the construction of new knowledge. In particular, the model suggests that learning from worked-out examples is most efficient if the solution steps provided within the examples are explained. During the attempt to explain the solution steps, missing knowledge can be identified and subsequently be constructed deductively or inductively by making use of domain-specific and general background knowledge. In many cases, this newly acquired knowledge is a necessary prerequisite for successful problem solving and further learning.

Further Research on Self-Explanations

The main findings of the study of Chi et al. (1989) have also been found by Ferguson-Hessler and de Jong (1990) in the application domain of electricity and magnetism and by Pirolli and Bielaczyc (1989) as well as Pirolli and Recker (1994) in the application domain of programming recursive functions in Lisp. However, by taking advantage of regression analyses, Pirolli and Recker (1994) demonstrated that the relationship between the amount of constructed explanations and the amount of newly acquired knowledge is not linear. While the construction of self-explanations seems to be related to learning, the learning seems to be more rapid on the basis of the first explanations constructed and less rapid on the basis of the last explanations constructed.

Pirolli and Recker (1994) assume that the longer the students work on a worked-out example, the fewer explanations are constructed which lead to the acquisition of new knowledge and the more superficial explanations are produced: "... after the first insightful elaborations are generated, there is a tendency to persevere on paraphrasing or embellishing those early insights without adding anything fundamentally new to the insight and at the expense of explaining other elements." (Pirolli & Recker, 1994, p. 266). Hence, there appears to be a trade-off where the construction of self-explanations becomes more effective by moving on to the next part of the instruction.

Like Chi et al. (1994), Bielaczyc, Pirolli and Brown (1995) investigated whether understanding in the application domain of Lisp-programming can be improved by systematically initiating the construction of self-explanations. Two groups of students were

investigated. Both groups worked on the same instructional material. However, after working on an introduction to Lisp-programming, one group was trained in constructing self-explanations and making use of self-regulation. This self-regulation training aimed at enabling the students to monitor their problem solving behavior and to identify and resolve comprehension failures. The other group was nottrained to take advantage of these learning strategies. Bielaczyc, Pirolli and Brown (1995) observed that the trained students made significantly more use of self-explanation and self-regulation strategies while working on the instruction than the untrained students. Even more important, this increase in applying learning strategies was also reflected by significantly better programming performances of the trained students.

Renkl (1997a) noticed that in the study of Chi et al. (1994) the group of students who were asked to explain the textbook's section spent significantly more time on the instructional material than the group which was not asked to do so. Thus, it could be that the time spent on the instructional material accounts for the observed differences in learning between the two groups. In order to examine this possibility, Renkl (1997a) investigated 36 students as they explained aloud worked-out examples in the application domain of elementary probability theory. Independent of how many examples were studied, the students had 25 minutes time to work on the examples. Before and after studying the examples, the students had to work on a pre- and posttest, respectively.

By means of protocol and regression analyses, Renkl (1997a) found that the amount of constructed explanations correlates positively and significantly with the students' scores on the posttest. This result coheres with the findings of Chi et al. (1989) and Chi et al. (1994). The main result of Renkl (1997a) is, however, that different students predominantly and consistently constructed different kinds of explanations. For instance, while several students predominantly tried to explain the principles which were underlying the examples, other students predominantly tried to explain how the solution steps which were provided in the examples can be achieved. Thus, the concept of self-explanation appears to be less homogenous than is suggested by the findings of Chi et al. (1989) and it seems that different kinds of self-explanations need to be distinguished.

Summary

Research on explaining to oneself suggests that self-explanations make up constructive cognitive activities which frequently lead to the acquisition of new knowledge. During the construction of self-explanations, learning seems to take place due to the identification of missing knowledge which would be required in order to complete the self-explanations. Such identified knowledge gaps might subsequently be filled by taking advantage of deductive and/or inductive learning mechanisms. Thus, to request someone to (self-) explain corresponds to asking them to try to understand. Thereby, different kinds of self-explanations might take into account different aspects of the scrutinized application domain and different self-explainers might systematically prefer different kinds of self-explanations.

Explaining to Others

If explaining to oneself can lead to the acquisition of new knowledge, then explaining to somebody else might have the same beneficial consequences. The idea that the construction of explanations might lead to the acquisition of new knowledge is hardly new; it forms the foundation of the ancient "socratic teaching" method. In socratic teaching, the teacher does not teach by direct exposition of the instructional material, but guides the student's own explorations by successively posing questions.

As Webb (1989) points out, explaining to others potentially offers even more opportunities for learning than explaining to oneself. Learning might not only take place due to

one's own identification of missing knowledge, but also because the receiver of the explanation identifies further missing information, points out inconsistencies, requires further clarification or confronts the explainer with alternative points of view. In order to resolve these discrepancies, the explainer might search for further information, deduce and induce new information or restructure already available information and thus further learn about the domain under consideration. Schwartz (1995) has demonstrated, for instance, that students acquire more abstract knowledge during collaborative problem solving than during individual problem solving. Schwartz assumes that explanations constructed during collaborative problem solving frequently need to bridge different viewpoints and thus need to be more abstract than is required for each viewpoint alone.

Explaining to others takes place in almost all situations in which at least two individuals collaborate and communicate. Thus, the construction of explanations is at the heart of collaborative problem solving and learning. However, as Webb (1989) notices, during collaborative problem solving and learning, not all explanations lead to the desired learning effects. Only students who construct highly elaborated explanations seem to learn from them. This finding is in accord with the findings of research on self-explanation: while the construction of explanations which involve inferences seems to enhance students' knowledge, the construction of explanations which only rephrase what is already known seems to have only minor learning effects.

The main research on explaining to others has been conducted in learning by teaching settings. Within these settings, two or more students or a student and a tutor construct explanations to teach each other.

Reciprocal Teaching

One important finding of educational as well as psychological research is that students with deficient problem solving and learning abilities frequently behave rather passively during instruction. If these students could be encouraged to actively participate in instruction, would their understanding improve? With this question in mind, Palincsar and Brown (1984) investigated 24 seventh-grade students as they attempted to acquire various strategies for processing texts such as locating comprehension failures, questioning, clarifying, summarizing and predicting.

Palincsar and Brown (1984) formed two experimental and two control groups. The two experimental groups received instruction. In order to encourage the students of one experimental group to actively participate in instruction, Palincsar and Brown (1984) developed an instructional procedure called “reciprocal teaching”. The second experimental group received instruction by making use of a more traditional method termed “locating information”. Both instructional methods are described below. In order to determine the effects of repeated testing, one control group received all of the daily assessments but no instruction. The second control group participated only in the pre- and posttest.

In reciprocal teaching, an experienced tutor and an inexperienced student take turns discussing a text. The tutor selects a text to be read and indicates whether it is his/her or the student’s turn to teach the text. Afterwards, the tutor as well as the student read the text individually and silently. Thereafter, the teacher (i.e., the tutor or the student) raises questions, points out difficult sections of the text, offers clarifications and explanations, formulates summaries and/or makes predictions about the future content of the text.

Initially, the tutor models the mentioned comprehension activities and the student acts as a rather passive observer. In the beginning, for instance, the student might have severe difficulties when his/her turn to teach comes. However, gradually the student becomes more able to assume his/her role as a teacher and to employ the various comprehension strategies by

him-/herself. As a consequence, the tutor behaves less as a mentor and more like a partner. Finally, both tutor and student pose questions and offer explanations to each other as well as engage in negotiations to reach mutual agreement.

According to Palincsar and Brown (1984), the locating information method is frequently used by school teachers to help students in answering questions about a text they have just read. Essentially, the students are shown how to locate information in texts in order to answer specific questions. If information from several text sections needs to be combined, then the teacher demonstrates how to do this. The teacher reinforces correct answers of the students and guides the students back into the text when incorrect answers have been provided. If necessary, the teacher mentions even the lines where the required information can be found.

All students in the study of Palincsar and Brown (1984) had to work on pre- and post tests which assessed the students’ text comprehension. In addition, the students in the two experimental groups as well as in the first control group had to work each day on comprehension questions. The students were given feedback on a daily basis. For instance, they were shown diagrams which visualized the percentage of correct answers to the comprehension questions as well as cumulative records for each week. Overall, the students in the two experimental groups received 20 days of instruction.

The main findings of Palincsar and Brown (1984) can be summarized as follows:

- the students who participated in reciprocal teaching performed much better on the assessments of text comprehension than the students who participated in locating information,
- the students who participated in reciprocal teaching gradually improved in assuming their role as a teacher, their questions and summaries were increasingly expressed in their own words and focused more and more on the main ideas of the text (cf. Table 2),
- the performance on the daily assessments of text comprehension of those students who

participated in reciprocal teaching improved in most of the cases from 30% to 80% correctness within 12 days and classroom comprehension scores rose from 20% to 60% and

- the students who participated in locating information performed hardly better on the assessments of text comprehension than the students who received no instruction at all but all assessments.

Insert Table 2 about here

The effects found by Palincsar and Brown (1984) were durable and rather general. Further assessments of text comprehension showed almost no decline in the level of performance for a period of eight weeks. With respect to the classroom, several students who had participated in reciprocal teaching reached or even surpassed the average level of performance for their age. As a consequence of observing models and teaching, students seem thus to acquire more complete and presumably better organized knowledge.

Palincsar and Brown (1984) assume that two reasons are responsible for the success of reciprocal teaching. The first reason is that reciprocal teaching involves extensive modeling of the activities to be taught. The second reason is that reciprocal teaching forces the student to actively locate comprehension failures, formulate questions, judge answers to these questions and construct explanations.

Further Research on Learning by Teaching

In an early investigation, Cloward (1967) already demonstrated that students can substantially improve their performance in the application domain of text comprehension by means of teaching other students. A group of 240 tenth and eleventh-grade students with

deficiencies in text reading and comprehension tutored a group of fourth and fifth-grade students who, in comparison to their class mates, also had difficulties in reading and comprehending texts.

Initially, the tutors received training in various teaching activities.[1] Afterwards, each tutor guided one fourth or fifth-grade student for a period of about five months. In parallel, small groups of tutors met once a week with a supervising teacher in order to discuss problems encountered during the tutorial sessions. A control group of fourth and fifth-grade students with deficiencies in text reading and comprehension received no tutorial support. Before and after the five month of tutoring, both groups of students had to take a pre- and posttest on text reading and comprehension, respectively.

Cloward (1967) found that not only the fourth- and fifth-grade students' ability to read and comprehend texts improved significantly from the pre- to the posttest, but also the tutors'. Surprisingly, the tutors gained even more than the students who were tutored. Comparable findings are reported by others, including Fantuzzo, Riggio, Connelly and Dimeff (1989) and Kafai and Harel (1991a, 1991b); for an overview see Goodlad & Hirst, 1989. In accordance with the results of the study of Palincsar and Brown (1984), these findings also suggest that students acquire more complete and presumably better organized knowledge by means of teaching. However, what are the reasons for the beneficial effects of teaching? Does the learning mainly take place during preparing to teach, during presenting the instructional material or during responding to questions?

Bargh and Schul (1980) investigated whether students learn while preparing to teach or whether they learn while presenting the material and responding to questions. In a first experiment, after taking a pretest, one group of undergraduate students studied a text for a fixed period of time to learn it themselves. A second group of undergraduate students studied the same text for the same period of time. Both groups of students were paid according to their scores on a posttest. However, the second group of students were told that after the study period they would have to teach the contents of the text to other students. In addition, they were told that the students who received the teaching would take the posttest in their stead and the students' scores would count as their own. Those students who expected to teach gained significantly more than those students who studied only for themselves. Comparable findings are reported by Benware and Deci (1984), for example.

Dayer (1996) observed comparable effects of developing educational software. During the development of courseware one has not only to design the presentation of information about the application domain but also the presentation of feedback and thus to anticipate potential errors of those who utilize the courseware. In a study conducted by Dayer (1996), twelve pairs of fourth graders developed simple programs on how plurals of compound names in french are formed. The programs have been developed by means of a graphical authoring tool. Before and after the study the students had to work on a pre- and posttest.

Dayer (1996) compared the programs developed by the five pairs with a low gain from the pre- to the posttest with the programs developed by the seven pairs with a high gain. Compared to the pairs with a low gain between pre- and posttest, the pairs with a high gain more frequently included feedback that did not show the correct solution but mentioned the grammatical category of the word to be changed. For example, one pair of students included the feedback "carry is a verb and must hence not be in agreement." Feedback on the grammatical category relates a specific case and a general rule such as "if the first word is a verb, then it does not take the plural mark."

These results are in accord with Webb's (1989) findings that only elaborated explanations support learning. They might also indicate that the anticipation of the explainee's behavior plays a role equivalent to the reaction to the explainee's actual behavior. However, because the study by Dayer (1996) was of rather preliminary nature this hypothesis needs to be

confirmed in the future by more comprehensive studies.

In a further experiment, Bargh and Schul (1980) scrutinized whether students learn from presenting instructional material and from responding to questions. Three groups of students were investigated. All three groups studied a text as they worked on a problem solving task for a fixed period of time. One group of students studied alone and silently. The second group of students also studied alone, but were asked to verbalize everything that came to their minds. The third group of students each taught another student while studying. Before and after problem solving and studying, all three groups of students had to take a pre- and posttest, respectively. Bargh and Schul (1980) found no significant differences between the investigated groups with respect to both problem solving and text studying.

Together, these findings suggest that the construction of explanations during teaching in combination with subsequent feedback might be less beneficial than one might expect. In order to further wade into the question of whether explanations facilitate learning, Renkl (1995, 1996) investigated two groups of undergraduate students. Initially, both groups of students read an introductory text on elementary probability theory and studied several worked-out examples for a fixed period of time. Afterwards, dyads were formed with one student out of each group. While one student had been asked to explain several new worked-out examples to the other student for a fixed period of time, the other student had been asked to listen. However, the listening student was allowed to pose questions and provide short comments to the explaining student. Before and after studying and then explaining and listening, both groups of students had to take a pre- and posttest, respectively.

In accordance with the results of the study of Bargh and Schul (1980), Renkl (1995, 1996) found that the explaining students did not learn more than the listening students. To further pinpoint how questions affect learning by explaining, Renkl (1997b) investigated two further groups of undergraduate students. Again, both groups of students read an introductory text on elementary probability theory and studied several worked-out examples for a fixed period of time. Afterwards, both groups of students were asked to explain several new worked-out examples to another person for a fixed period of time. However, this time the person who received the explanations was given specific instruction as to the type of allowable response. In one group, the instructed person posed semi-standardized questions to the explaining students. These questions always had the form "what if ...". In the other group, the instructed person behaved rather passively and provided only neutral acknowledgements such as "hm" and "I see". Before and after studying as well as explaining, both groups of students had to take a pre- and posttest.

Renkl (1997b) found that the students who were systematically asked for further information and clarification did not learn more than the students who received simple content-free acknowledgements. Preier (1996) investigated several groups of undergraduate students as they (1) only read a text on the functioning of steam engines, (2) explained the text to themselves, (3) explained the text to other students without receiving feedback from the listeners and (4) explained the text to other students while receiving feedback. Again, no differences were found between students who self-explained and students who explained to others or between students who explained with or without feedback and students who listened. However, since various aspects were confounded in this study, only preliminary conclusions should be drawn from it.

The categories examined by Preier (1996) are related to the different levels of interactivity as described in the introductory part of this chapter. While no effects were observable in the study by Preier (1996), there is still some reason to believe that different kinds of interactivity can have effects, epsecially for the receivers of explanations. In a study on task performance rather than learning per-se, Clark and Schober (1989) demonstrated that

"addressees" perform better on a directed task than "overhearers". Using a tangram matching task, pioneered by Krauss and Weinheimer (1964), in which a director explains how to reconstruct a figure before him using component pieces, two experimental groups tried to perform the task. The first group, the addressees, were allowed to ask questions and interact with the director, giving both positive and negative feedback of their current understanding and task performance. The second group, the overhearers, were allowed to listen to the exchanges of the others, but were not allowed to interact with them. The addressees were not aware that the overhearers were also trying to perform the task.

The results showed that addressees performed significantly better on the task (85% - 98% correct performance; $\underline{F}(1, 18) = 10.83$, $\underline{p} < .005$). Clark and Schober (1989) conclude from these results that the social role of interaction plays a central role in the cognitive process of learning and understanding (see also Chapter 3). Still, it remains to see how well this kind of performance related learning carries over to conceptual learning as investigated in other studies.

Summary

In research on collaborative problem solving and learning, it has repeatedly been demonstrated that teaching frequently leads to more complete and presumably better organized knowledge. Since one obvious activity during teaching is explaining to others and responding to questions, this gave raise to the assumption that the beneficial effects of teaching are due to the successive construction of explanations (e.g., Webb, 1989). During the construction of explanations, learning might take place due to one's own identification of missing knowledge, but also because the receiver of the explanation identifies further missing information, points out inconsistencies within the explanation, requires further clarification or confronts the explainer with alternative points of view. However, recently conducted research which aimed at clarifying the role of explanations for learning by teaching does not support this hypothesis. Instead, it has been demonstrated that preparing for teaching, i.e., studying the instructional material individually, results in substantial learning.

Why have no substantial differences between
explaining to oneself and explaining to others been observed?

Although, intuitively, explaining to others seems to offer even more opportunities for learning than explaining to oneself, so far, no substantial differences between explaining to others and explaining to oneself have been observed with respect to learning effects. In the following, we discuss why it might be the case that until now no such differences have been observed.

We hypothesize that the empirical research conducted so far did not focus specifically enough on the question of whether explaining to oneself or explaining to others leads to more learning. Especially, it appears that possible differences between the investigated settings have not been maximized, but rather have been blurred. In the research realized so far, various aspects blurred possible differences:

- Self-explanations as observed, for instance, by Chi et al. (1989) were not exclusively directed to oneself, but were (mainly?) directed to an experimenter who listened to them.
- Explanations directed to oneself but listened to by an experimenter have possibly been constructed more carefully than explanations directed to peers, because of an experimenter's status.
- In settings for explaining to oneself, students were possibly more systematically prompted to explain than in settings for explaining to others.
- In research as conducted by Chi et al. (1989) and Chi et al. (1994), students were explicitly asked to construct self-explanations which are beyond simple paraphrases, but

relate different kinds of information and give new insight. Conversely, the research on explaining to others, did not closely monitor which kinds of explanations were constructed by the students.
- On one hand, it might be the case that successful self-explainers imagined interactions with fictitious explainees and thus imagined explaining and reacting to others. On the other hand, in research on explaining to others it is often not controlled to which extent an explainer takes into account the reactions of the student who receives the explanations.

It could be that controlling empirical aspects such as those mentioned above would allow one to observe more substantial differences between the learning effects of explaining to oneself and explaining to others. However, it could also be that the mechanisms involved in explaining to oneself and explaining to others are not distinct enough or that the effects of these mechanisms are confounded:

- While during self-explaining a student might check his own explanations for consistency and completeness, during explaining to others this task might be shared between the explainer and the explainee. Thus, one might hypothesize that explaining to others results in lower cognitive load and therefore in more learning. However, lower cognitive load does not necessarily lead to more individual learning. Perhaps it is exactly the self-monitoring that plays an important role in learning.
- If explaining to others is to be instructive and efficient, the explainer needs to take into account the explainee's knowledge and understanding (cf. Renkl, 1995, 1996): Does he/she already know what I am going to tell? How shall I explain this? Does he/she have a different point of view? Did he/she understand what I explained? Most probably, such monitoring leads to additional cognitive load on the explainers side. Again, it is unclear whether this additional cognitive load impedes or improves learning in the long run.

- During explaining to each other, not only does the explainer have a model of the explainee, but the explainee also has a model of the explainer. Messages, including feedback about previous explanations, can affect the models of the receiver, including the receiver's model of the sender's model of the receiver. Such embedded levels of modeling have been studied in linguistic pragmatics, for instance. Clark and Marshall (1981) give examples where deep nestings can be essential in picking out a correct referring expression. Cohen (1978) shows how even infinite nestings can be represented finitely and efficiently. Still, it is not known to what depth such models should be maintained, nor what the precise makeup of such a model is. For instance, Taylor and Carletta (1994) suggest that only two levels of nesting are sufficient for cooperative dialogue. Even if a person models another person sufficiently to continue the dialogue, this might still be merely at a shallow level of understanding, not enough to actually learn. It is possible that such modeling is only taken advantage of to superficially repair communications, for example, to rephrase an explanation without any further consequences for the knowledge which underlies these explanations.

Conclusions

Current research does not allow one to take a clear stance on the question of whether explaining to oneself or explaining to others is the more efficient way to learn. However, up to now almost no research directly addressed this question. More research aimed directly at this question is required to identify the learning effects of both kinds of explaining, and uncover the processes which lead to this learning,. Learning by explaining to oneself and by explaining to others needs to be empirically compared in more systematic ways. In this chapter, we discussed various factors which could be controlled in such comparison studies. Since both explaining to oneself and explaining to others make up constructive cognitive activities, it

might turn out, however, that this common characteristic is the one that essentially promotes learning and that it is less important in which of the two settings these activities are realized.

Acknowledgments

This research has been supported by the European Science Foundation (ESF) within the program "Learning in Humans and Machines".

References

Aronson, E., Bridgeman, D. L., & Geffner, R. (1978). The effects of a cooperative classroom structure on student behavior and attitudes. In D. Bar-Tal & L. Saxe (Eds.), Social psychology of education: Theory and research. Washington, D.C.: Wiley.

Bargh, J. A. & Schul, Y. (1980). On the cognitive benefits of teaching. Journal of Educational Psychology, 72, 593-604.

Benware, C. A., & Deci, E. L. (1984). Quality of learning with an active versus passive motivational set. American Educational Research Journal, 21, 755-765.

Bielaczyc, K., Pirolli, P. L., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. Cognition and Instruction, 13, 221-252.

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., de Leeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. Cognitive Science, 18, 439-477.

Chi, M. T. H. & VanLehn, K. (1991) The content of self-explanations. The Journal of the Learning Sciences, 1, 69-105.

Clark, H. H. & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Webber, & I. Sag (Eds.), Linguistic structure and discourse setting (pp. 10-63). Cambridge, MA: Cambridge University Press.

Clark, H. H. & Schober, M. (1989). Understanding by addressees and overhearers. Cognitive Psychology, 21, 211-232.

Cloward, R. D. (1967). Studies in tutoring. The Journal of Experimental Education, 36, 14-25.

Cohen, P. R. (1978). On knowing what to say: planning speech acts (Unpublished doctoral dissertation). Toronto: Department of Computer Science at the University of Toronto.

Dayer, P. (1996). Le tutorat virtuel: Apprendre en construisant un didacticiel (Unpublished graduate thesis). Geneva: Faculty of Education and Psychology at the University of Geneva.

Fantuzzo, J. W., Riggio, R. E., Connelly S., & Dimeff, L. A. (1989). Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: A component analysis. Journal of Educational Psychology, 81, 173-177.

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.

Goodlad, S. & Hirst, B. (1989). Peer tutoring - A guide to learning by teaching. London: Kogan.

Halliday, D. & Resnick, R. (1981). Fundamentals of physics (2nd ed.). New York, NY: Wiley.

Kafai, Y. & Harel, I. (1991a). Children learning through consulting: When mathematical ideas, knowledge of programming and design, and playful discourse are intertwined. In I. Harel & S. Papert (Eds.), Constructionism (pp. 111-140). Norwood, NJ: Ablex.

Kafai, Y. & Harel, I. (1991b). Learning through design and teaching: Exploring social and collaborative aspects of contructionism. In I. Harel & S. Papert (Eds.), Constructionism (pp. 85-110). Norwood, NJ: Ablex.

Knight, G. P. & Bohlmeyer, E. M. (1990). Cooperative learning and achievement: Methods for assessing causal mechanisms. In S. Sharan (Ed.), Cooperative learning - Theory and research (pp. 1-22). New York, NY: Praeger.

Krauss, R. M. & Weinheimer, S. (1964). Changes in reference phrases as a function of the frequency of usage in social interaction: A preliminary study. Psychonomic Science, 1, 113-114.

Mitchell, T. M., Keller, R. M., & Kedar-Cabelli S. T. (1986). Explanation-based generalization: A unifying view. Machine Learning, 1, 47-80.

Palincsar, A. S. & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. Cogntion and Instruction, 1, 117-175.

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

Preier, M. (1996). Lernen beim Erklären: Was lernt man selbst, was der Zuhörer? [Learning by explaining: What learns the explainer and what the listener?] (Unpublished graduate thesis). Freiburg: Department of Psychology at the University of Freiburg.

Renkl, A. (1995). Learning for later teaching: An exploration of mediational links between teaching expectancy and learning results. Learning and Instruction, 5, 21-36.

Renkl, A. (1996). Lernen durch Erklären - oder besser doch durch Zuhören? [Learning by explaining - or better by listening?] Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 28, 148-168.

Renkl, A. (1997a). Learning from worked-out examples: A study of interindividual differences. Cognitive Science, 21, 1-29.

Renkl, A. (1997b). Lernen durch Erklären: Was, wenn Rückfragen gestellt werden? [Learning by explaining: What happens if questions are posed?] Zeitschrift für Pädagogische Psychologie, 11, 41-51.

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. The Journal of the Learning Sciences, 4, 321-354.

Taylor, J. & Carletta, J. (1994). Why we can get away with simple belief models for cooperating agents. In A. Ram & K. Eiselt (Eds.), Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum Associates.

VanLehn, K., Jones, R. M., & Chi, M. T. H. (1992). A model of the self-explanation effect. The Journal of the Learning Sciences, 2, 1-59.

Webb, N. M. (1989). Peer interaction and learning in small groups. International Journal of Educational Research, 13, 21-39.

Footnotes

[1] Unfortunately, Cloward (1967) does not describe which teaching activities have been trained.

Table 1

Examples of self-explanations as observed in the study of Chi et al. (1989; see also Chi & VanLehn, 1991)

Self-explanations related to technical procedures:
Read (string problem in Figure 1): Choosing the x- and y-axes as shown, we can write this vector equation as three scalar equations.
Self-explanation: Ummmm, I guess always before when I thought of force vectors I just thought of them as going in a particular direction, I forgot about them having, having x-components and y-components and being, ummm, broken down into them.
Read (an inclined plane problem): With this choice of coordinates, only one force, m g, must be resolved into components in solving the problem.
Self-explanation: I see that because it is the only one that would not be on one of those axes.
Self-explanations related to physics principles:
Read (string problem in Figure 1): The body remains at rest under the action of the three forces.
Self-explanation: So the sum of forces should be zero.
Read (an inclined plane problem): Since the block is unaccelerated, we obtain $\mathbf{F} + \mathbf{N} + \mathbf{m}\ \mathbf{g} = \mathbf{0}$
Self-explanation: There seem to be the three forces involved also they all sum to zero somehow.
Self-explanations related to physics concepts:
Read (string problem in Figure 1): Consider the knot at the junction of the three strings to be the body.
Self-explanation: Uh, so they refer to the point as the body.
Read (string problem in Figure 1): The body remains at rest under the action of the three forces.
Self-explanation: I see. So the W will be the force and not the body.

Table 2

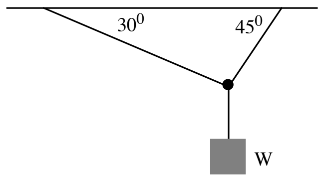
Examples of the acquisition of question asking as observed in the study of Palincsar and Brown (1984) during reciprocal teaching

First day (after reading a short text about snakes):	
Student:	What is found in the southeastern snakes, also the copperhead, rattlesnakes, vipers - they have. I am not doing this right.
Teacher:	All right. Do you want to know about the pit vipers?
Student:	Yeah.
Teacher:	What would be a good question about the pit vipers that starts with the word "why?"
Student:	---
Teacher:	How about, "Why are the snakes called pit vipers?"
Student:	Why do they want to know that they are called pit vipers?
Teacher:	Try it again.
Student:	Why do they, pit vipers in a pit?
Teacher:	How about, "Why do they call the snakes pit vipers?"
Student:	Why do they call the snakes pit vipers?
Teacher:	There you go! Good for you.
Eleventh day (after reading a short text about the Venus flytrap):	
Student:	What is the most interesting of the insect eating plants, and where do the plants live at?
Teacher:	Two excellent questions! The are both clear and important questions. Ask us one at a time now.

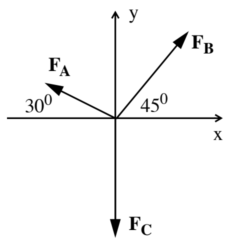
Figure Caption

Figure 1. A worked-out physics example as used in the study of Chi et al. (1989).

(a)



(b)



A block of weight W is suspended by strings (a). Consider the knot at the junction of the three strings to be the body. A free-body diagram shows all the forces acting on the knot (b). The strings are assumed to be weightless. The body remains at rest under the action of the three forces. Suppose we are given the magnitude of one of these forces. How can we find the magnitudes of the other forces?

$\mathbf{F_A}$, $\mathbf{F_B}$ and $\mathbf{F_C}$ are all the forces acting on the body. Since the body is unaccelerated, $\mathbf{F_A} + \mathbf{F_B} + \mathbf{F_C} = \mathbf{0}$. Choosing the x- and y-axes as shown, we can write this vector equation as three scalar equations:

$$F_{Ax} + F_{Bx} = 0 \text{ and}$$
$$F_{Ay} + F_{By} + F_{Cy} = 0.$$

The third scalar equation for the z-axis is simply

$$F_{Az} = F_{Bz} = F_{Cz} = 0.$$

That is, the vectors all lie in the x-y-plane so that they have no z-components. From the figure we see that

$$F_{Ax} = - F_A \cos 30^0 = - 0.866\, F_A,$$
$$F_{Ay} = F_A \sin 30^0 = 0.500\, F_A,$$
$$F_{Bx} = F_B \cos 45^0 = 0.707\, F_B,$$
$$F_{By} = F_B \sin 45^0 = 0.707\, F_B.$$

Also, $F_{Cy} = - F_C = - W$ because the string C merely serves to transmit the force on one end to the junction at its other end. Substituting these results into our original equations, we obtain

$$- 0.866\, F_A + 0.707\, F_B = 0,$$
$$0.500\, F_A + 0.707\, F_B - W = 0.$$

If we are given the magnitude of any one of these three forces, we can solve these equations for the other two.