APPLIED COGNITIVE PSYCHOLOGY, VOL. 10, S33-S49 (1996)
Note: The contents of this Acrobat version are the same that the one on the journal, only pagination has changed.

Constructing Self-Explanations and Scaffolded Explanations in Tutoring

MICHELENE T. H. CHI
Learning Research and Development Center, University of Pittsburgh

SUMMARY

One-on-one tutoring is a form of instruction that requires interaction between a tutor and a tutee. The effectiveness of tutoring is examined from the perspectives of the tutor's actions, the tutee's actions, and successive interactions. A tutor's actions that may not lead to successive interactions consist of asking an initiating question, providing feedback, and asking a comprehension-gauging question. It is suggested that these types of actions can lead to the learning of an ideal template of solution procedures for solving problems, but may not lead to deep understanding. A tutee's actions are postulated as self-explaining, in response to either tutors' questioning, tutors' prompting, or tutors' scaffolding. Finally, an interaction is defined as a sequence of tutor actions, such as scaffolding, which elicits a successive series of exchanges between the tutor and tutee, so that they collaboratively construct a response. An exercise in a detailed protocol analysis of a case study of a student being tutored in solving a mechanics problem is presented. It shows what misconceptions the student exhibited, whether these misconceptions were removed, and what actions triggered the learning. The results of this case study support the suggestion that tutor actions that prompt for co-construction (which includes self-explanation) may be the most beneficial in producing deep learning, in the sense of removing misconceptions.

Human tutoring is a more effective means of instruction than classroom teaching, mastery learning, computer-aided or programmed instruction and computer tutors. To understand what tutoring is, we might consider what it is not. Tutoring is not listening to a lecture, reading a manual or textbook, solving a problem solo, studying an example by oneself, observing a model or modelling (i.e., observing an expert performing, Chi and Bjork, 1991*)*, or practicing. Thus, tutoring is not learning by oneself. Rather, tutoring refers to the interactive portion of instruction. It consists of a continuous stream of exchanges between a tutor and a tutee.

A typical dialogue pattern of tutoring can be thought of as consisting of the following five broad steps, henceforth referred to as the tutoring frame (adapted from Graesser, Person and Magliano, 1995):

(1) Tutor asks an initiating question
(2) Student provides a preliminary answer
(3) Tutor gives feedback on answer
(4) Tutor scaffolds to improve or elaborate the answer in a successive series of exchanges (taking 5-10 turns, Graesser *et al.,* 1995)
(5) Tutor assesses student's understanding of answer

A portion of this paper was reported at the University of Memphis Conference on Reasoning, September, 1995. This research was supported by the Spencer Foundation and the Office of Naval Research (Contract N00014- 94-2-0674). The opinions expressed in this paper do not necessarily reflect the positions of the sponsoring agencies, and no official endorsement should be inferred. The author is grateful for comments provided by Bill Baggett. Reprints can be obtained by writing to the author at the Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara Street. Pittsburgh, PA 15260, USA. The author may also be contacted via email at CHI @vms.cis.pitt.edu.

The goal of this paper is to (a) consider the contribution of the tutor, the tutee, and their interaction, as sources of learning, and (b) present a case study to assess what specific knowledge is learned in the context of solving mechanics problems, and then determine whether it is the tutor's actions, the tutee's actions, or their interactions, that contributed to the learning. The discussion below examines tutoring effectiveness from the perspective of the tutor's actions, the tutee's actions, and their interactions.

TUTOR'S ROLE AND ACTIONS: THE INSTRUCTIONAL HYPOTHESIS

Tutoring effectiveness has traditionally been attributed to the role of the tutor, even if the tutor is unskilled. The implicit assumption has been that tutoring requires some kind of skill, so there has been a quest to find the appropriate skills or strategies that tutors use. This assumption is not at all surprising given that the tutoring frame shows a preponderance of the tutor's actions: the tutor seems to set the agenda completely (i.e., the tutor crafts the appropriate next question, selects the next problem, analogy, example, or counter-example, directs the conversation and basically controls the entire interchange). Therefore, the natural question that arises about the effectiveness of tutoring has always been, ‘What do tutors do that is so effective’ (Merrill, Reiser, Ranney and Trafton, 1992, p. 279). Thus, tutoring researchers attribute the tutor with these responsible actions, much as educators attribute the teachers with such similar responsible actions. Below, three of the five tutor's actions from the tutoring frame are elaborated to see what a tutee might learn from such actions. The nature of these three actions are viewed as more discrete and less conducive for providing extended interaction. Thus, any learning that may result from these three actions can be attributed primarily to the responsibility of the tutor.

Tutor asks an initiating question (step 1)

Tutors initiate most of the exchanges in tutoring dialogues. According to Graesser *et al.*'s (1995) data, only 5-10% of the subtopics were initiated by the students. The questions tutors ask to initiate a dialogue tend to be consistent with their curriculum scripts, that is, a set of subtopics and example problems that are consistent with the standard materials that have to be covered. For example, if the topic to be covered is factorial design, a tutor might start the dialogue by asking ‘Now what is a factorial design?’ (Graesser *et al.*, 1995). One can also think of a curriculum script at a more micro level as an ideal template (such as a standard solution procedure for a problem) that the tutors wish the students to learn. If tutors' questions are derived from a template, then one would predict that students would learn the template. The literature from intelligent tutoring systems seems to support this conclusion. Thus, the tutors' action of asking an initiating question can be broadly construed to be effective at promoting learning of an ideal template or materials contained in a curriculum script.

Tutor gives feedback (step 3)

One can differentiate between three kinds of feedback on the basis of their content: corrective feedback, didactic explanations, and suggestive feedback. If students' preliminary answers are incorrect (such as using the wrong equation or substituting a wrong value in an equation, or retrieving an incorrect fact), then a tutor can give feedback to correct it directly, such as saying the chosen equation is the wrong one. Merrill, Reiser, and Landes (1992) reported that tutors are very quick to respond with corrective feedback when students make a syntactic error (such as an algebra mistake). This kind of corrective feedback is critical in problem solving in that it effectively reduces searches down the wrong paths of a search space (Anderson, Boyle and Reiser, 1985). Corrective feedback of this kind is largely responsible for the success in intelligent tutoring systems. Thus, providing corrective feedback is a type of tutor action that is successful at enabling a student to learn the ideal template. (A subtype of corrective feedback is reinforcing feedback: that is, feedback that reinforces that what the tutee did was correct.)

Tutor's feedback can also take the form of long didactic explanations. Graesser *et al.*'s (1995*)* data showed that when students made errors, the tutors provided lengthy explanations 25% of the time. However, there is no direct evidence showing that students learn from long didactic explanations, either in the context of a classroom or in the context of tutoring.

If a student's preliminary answer is incomplete rather than blatantly incorrect, then tutors are likely to give suggestive feedback, which is more indirect in that they simply alert the tutee that there is a problem, without telling the tutee exactly what the problem is. Such suggestive feedback can take the form of hints, pauses, a raised intonation in the voice, a non-specific query (such as, ‘Is that so?’), a specific contentrelated question, or providing some new information (such as, ‘Have you thought of _________?’) in order to redirect the tutee's thinking. Because suggestive feedback typically launches a successive series of exchanges with the student, this kind of action will be considered to be a kind of scaffolding (step 4), to be discussed later in the third (interactive) hypothesis for tutoring effectiveness.

Tutor assesses student's understanding (step 5)

This action consists mostly of the tutors asking the students whether the students think they understood the answer. That is, it is a comprehension-gauging question, to be differentiated from assessing whether the students actually did understand the answer. Tutors rarely test the students' understanding. Although tutors often ask comprehension-gauging questions, such questions are fairly useless because students do not generally know whether or not they understand (Chi, Bassok, Lewis, Reimann and Glaser, 1989; Glenberg, Wilkinson and Epstein, 1982; Person, Graesser, Magliano and Kreuz, 1994).

In sum, of the three tutors' actions considered above, two of them clearly result in learning, but the learning may be of an ideal solution template or a curriculum script only. It is not clear whether, from these kind of tutoring actions, students can learn knowledge that transfers, nor is there any evidence that they learn to remove misconceptions. In fact, the evidence so far is quite convincing in showing that tutors do not diagnose the misconceptions that students do have (McArthur, Stasz and Zmuidzinas, 1990; Graesser et al., 1995). Instead, tutors seem quite competent at assessing what knowledge students are missing, using a curriculum script or an ideal solution template as a standard to determine what knowledge is missing (McArthur et al., 1990; VanLehn, Jones and Chi, 1992).

TUTEE'S ROLE AND ACTIONS: THE SELF-EXPLANATION HYPOTHESIS

A second source of tutoring effectiveness focuses on the action of the students, independent of the actions of the tutors. The tutees might be doing something more than they usually do in a classroom situation, or in a solitary studying situation. That is, in tutoring, students are placed in a situation in which they are given greater opportunities to generate answers or explanations to the tutors (steps 2 and 4). In fact, Graesser *et al.*'s (1995) data showed that tutors ask 104 questions per hour. Answering these questions may have the same beneficial effect as generating explanations to oneself (Chi *et al.*, 1989; Chi, de Leeuw, Chiu and La Vancher, 1994; Pirolli and Recker, 1994; Renkl, in press). Even though these answers may be fragmented, semi-coherent and distributed over many turns, they are similar to self-explanations in that self-explanations are also often fragmented small pieces of knowledge inferences. Thus, since the construction of self-explanations, directed by oneself without any guidance from another, is an effective means of learning, then it seems that the construction of explanations, elicited by others, may have the same beneficial effect and may account in part for the effectiveness of tutoring.

Besides self-explaining, tutees may be constructive in another way, by asking the tutors questions. Graesser *et al.* (1995) reported that tutees ask eight deep questions per hour, whereas students in a classroom setting ask only 0. 11 questions per hour, which include both deep and shallow questions. Thus, tutees clearly ask questions 72 times more often than students in classrooms. If a tutor responds to a student's questions, then such answers may enhance a tutee's learning, and receiving explanations from tutors to students' own questions may have beneficial effects as well, in contrast to receiving explanations from tutors or teachers that are not in response to students' specific questions, as was found in Webb's (1989) results. However, because the incidence of tutees asking questions is relatively small, it was not a salient part of the tutoring frame.

In general, tutoring provides many more opportunities for students to self-explain and to ask questions. The nature of both of these actions is that the students are participating in more active construction of knowledge (Bruer, 1993).

INTERACTION OF TUTOR AND TUTEE: THE CO-CONSTRUCTION OF SCAFFOLDED EXPLANATION HYPOTHESIS

The picture that emerges so far seems to suggest that some types of tutor actions (asking initiating questions, giving corrective feedback, assessing missing knowledge) enable learning of the curriculum script and/or ideal template, and may have minimal effect in promoting learning with deep understanding (that is, learning that transfers or learning that removes misconceptions). On the other hand, the tutees' actions, of having frequent opportunities to construct answers, may produce more profound learning, on the basis of the self-explanation work. The third perspective to consider is the interaction between the tutors and the tutees, resulting perhaps in some sort of co-construction of knowledge that facilitates learning. The tutors' scaffolding actions in step 4 of the tutoring frame seem to be the kind that encourage interactions. That is, scaffolding actions are the kind that elicit a response from the students, which then leads to additional scaffolding actions from the tutors. The fact that scaffolding can lead to numerous interchanges is consistent with typical tutoring protocols, which show a large number of conversational turns. As Graesser *et al.* (1995) reported, when their tutors asked a question, the median number of turns in producing an answer is between 5 and 10. This frequent turn-taking is consistent with the view that the tutors and the students are perhaps collaboratively constructing knowledge (or an answer), and this may be a salient aspect of learning from tutoring.

Evidence for the benefit of interaction

An exchange implies that there is an interaction between the tutor and the tutee. Although there is no clear definition of what constitutes an interaction, the tutors' actions discussed above, such as giving didactic explanations, seem less likely to lead to an interaction since they require only a generic (‘okay’) type of response. Other actions, such as scaffolding and prompting, seem more conducive to eliciting interactions.

The literature shows conflicting evidence regarding the benefit of collaborative types of interaction. Some data suggest that peer problem solving and collaborative learning in small groups is beneficial (e.g., Skon, Johnson and Johnson, 1981) and others not (Schwartz, 1995). However, it does appear that the more successful dyads are characterized by a higher degree of communication among the partners (Cooper, Ayers-Lopez and Marquis, 1982; Forman and Cazden, 1985), suggesting that communicative types of interaction may result in some kind of co-construction of knowledge. However, these findings do not discriminate between the benefit as arising from self-explaining or from co-constructing, since having a peer provides the opportunity for increased verbalization, and increased verbalization suggests that both self-explaining and co-constructing may be occurring.

Two studies address the role of explanation-construction in collaborative problem solving more directly. Coleman (1992) tested the effectiveness of scaffolded explanation prompts in a group discussion format. She asked three-person groups to arrive at a consensually agreed upon solution to photosynthesis problems after the groups had had 2 weeks of instruction about the topic. In the course of the discussion, she gave each member of the three-person group interchangeable roles, such as the ‘prompter’, the ‘explainer’ and the ‘reader/writer’. The prompter had to choose an appropriate cue from a set of prompting cards provided.

The reader/writer read the cue and wrote down a response, whereas the explainer generated explanations for the prompting cue and provided justifications. The roles rotated among the three-member team.
Examining the individual students' post-test scores, the intervention groups retained and comprehended more about photosynthesis than the control groups that did not receive the intervention. Furthermore, the groups' consensual solution was rated on four conceptual levels from intuitive and incorrect to scientifically correct. The consensual explanations of the intervention groups tended to be more conceptually advanced than those of the control groups. Hence, the first result, namely that individuals in the intervention group retained and comprehended more about photosynthesis, can be attributed to a self-construction of explanation effect, since each 'explainer' had to generate solo explanations. However, the second finding, of more advanced consensual solutions, would support the co-construction hypothesis. Thus, Coleman's findings support both the self- and the co-construction hypotheses.

Teasley (1992) manipulated whether children in collaborative problem-solving pairs were permitted to talk to their partners versus working alone in the context of paired problem solving. In the peer collaboration research, a consistent finding has been that dyads perform better than subjects working alone, using various outcome measures (Azmitia and Perlmutter, 1989; Tudge and Rogoff, 1989). However, in almost all of these studies, talking was confounded with having a partner, so that it is not clear whether the benefits of peer collaboration arose from the social context of having a partner or whether the presence of a partner provided the forum in which greater opportunities and obligations for self- and co-construction were available. Teasley manipulated both the availability of talking and the availability of a partner in a scientific reasoning task with fourth graders. The task was to discover how a new 'mystery' key on a spaceship simulation worked. This required that the children formulate hypotheses and design experiments to test their hypotheses.

The results were straightforward. Using the children's final hypothesis for the function of the mystery key as the dependent measure, the two Talk conditions (either Alone or in Dyads) scored significantly higher than the No Talk conditions. Moreover, the final scores for the Talk Dyads were not significantly better than those for the Talk Alones. Hence, the results suggest that the improved learning and problem solving in a collaborative context can be attributed to the increased opportunity for the students to talk and explain, rather than the social effect of having a partner, consistent also with King's (1994) results.

These two studies, along with other findings from the peer collaboration literature, suggest that the activity of talking, whether naturally available (as in the case of having a conversational partner) or explicitly prompted (either in a group situation as in Coleman's study or alone, as in Chi *et al* 's 1994 self-explanation work), significantly improves learning as measured by a number of different outcome measures. Still, neither study successfully discriminated between the self- and the co-construction of explanation hypotheses.

What is co-construction of scaffolded explanations?

No precise operational definition has been provided in the literature concerning what exactly co-construction is. Its definition has various meanings depending on the specific literature in which it is embedded. In the Vygotskian developmental context, co-construction can be seen as adult-child interaction such that the adult scaffolds the child within the child's zone of proximal development (Newman, Griffin and Cole, 1989). Co-construction is having the adult tailor the task in a way that allows the child to perform successfully. This is scaffolded co-construction, in the sense that the adult obviously knows more, and can craft the problem into subproblems that a child can handle. In the situated action perspective (Suchman, 1987) co-construction is generally viewed as having two (or more) people collaboratively construct a solution, an understanding, a shared meaning of knowledge, which neither partner possesses (Barwise and Perry, 1983; Perret-Clermont, Perret and Bell, 1991; Roschelle, 1992). Schliemann and Nunes (1990), for example, suggested that primitive fishermen can understand proportions and devise calculation procedures through everyday practices. Presumably, they invent this knowledge between themselves, without the help of an authoritative or more knowledgeable group member.

Because tutoring practices often entail a more knowledgeable person tutoring a less knowledgeable tutee, the context may be more analogous to the Vygotskian adult-child interaction. Thus, if co-construction occurs in tutoring, understanding is not so much jointly created by the tutor and tutee as it is passed on in the form of the tutor's expertise. The tutor's goal can thus be seen as trying to get the tutee to share the tutor's knowledge and understanding (through scaffolding), rather than to create or negotiate a shared meaning which neither one originally possessed.

One can conjecture aspects of dialogues that can be construed as scaffolded co-construction, even if no clear operational definition has been provided. There are some general guidelines from existing work. From the work on reciprocal teaching (Brown and Palinscar, 1989; Palinscar and Brown, 1984), we can infer that many activities, such as helping students to improve their clarifying, summarizing, question-asking and predicting skills, would constitute a form of co-construction. From the literature on cognitive apprenticeship and guided participation (Collins, Brown and Newman. 1989; Rogoff, 1990), we see these same activities in the processes of scaffolding and fading. Scaffolding, guided participation, and fading refer to the support and guidance and their gradual removal that a master provides an apprentice in performing some skill (Collins *et al.,* 1989). The support and guidance consist of hints (Hume, Michael, Rovick and Evens, 1996), maintaining goal orientation, highlighting critical task features (Stone, in press), developing the task, actual execution of parts of the skill (Rogoff, 1990), or providing physical props such as cue cards (Scardamalia, Bereiter and Steinbach, 1984). Scaffolding thus involves cooperative execution by the expert and the novice in a way that allows the novice to take an increasingly larger burden in performing the skill.

With these guidelines, we can define a list of actions that a tutor can take that would elicit a series of responses from a tutee such that the chain of conversational exchanges would constitute co-construction. That is, a response such as ‘okay’ would not qualify. These tutor actions can be subsumed under two broad categories, prompting and scaffolding. The main difference between prompting and scaffolding is that prompting can be generated without knowing the content domain, whereas scaffolding cannot. Prompting would be comments from tutors such as ‘What do you think should be done next?’ or ‘What does this mean?’ Prompting is basically an invitation by the tutors to elicit self-explanations from the students. Scaffolding, on the other hand, includes some aspects of the content. From the literature, we can discern a preliminary set of scaffolding actions to be: (1) describing the problem so as to orient the student to the important feature; (b) comparing the current problem/concept to a previous problem/concept; (c) providing the student with a goal; (d) querying the student in the context of the tutor's goal; (e) completing the student's reasoning step (or ‘splicing in’ the correct answer, Graesser *et al.,* 1995); (f) initiating the beginning of a reasoning step and asking the student to complete the step; and (6) redirecting the student. (A comprehensive list of scaffolding actions can be derived for individual sets of protocol data.) In many ways, all of these actions can be considered a form of hint. Thus, scaffolding is a joint activity in which the tutor plays some role and expects the student to respond.

The nature of prompting and scaffolding differs dramatically from a discrete one shot question-answering mode of interaction that is typically held in a classroom. Question-answering means that a specific and complete content question is asked, usually on the basis of the curriculum script, and a coherent answer to that question is expected. For instance, the initiating question that a tutor asks (as in step 1) is usually a complete question. For example, the tutoring dialogue presented by Graesser *et al.* (1995) shows an initiating question posed by a tutor as ‘Now what is a factorial design?’ However, on the basis of the student's response (‘The design has two variables’), the subsequent scaffolding (step 4) question was ‘So there are two or more independent variables and one ... [pause]’. Note that a scaffolding question differs from the initiating question in that it is more specific and fragmented, with an expectation (in this case) of a fragmented fill-in type of answer (‘dependent variable’) rather than a complete answer. In the classroom, once an initiating question is posed, the teacher expects a complete answer, so that if the answer of ‘The design has two variables’ was given, then the teacher is more likely to move on to other students and query ‘What else?’ in order to get a complete answer, rather than scaffold the initial student's response.

Both prompting and scaffolding occur in rapid succession in a natural conversational turn-taking manner because the expectation is to build an answer over several turns. Thus, instead of expecting a discrete answer to a question, scaffolding and prompting expect a continuous series of interchanges. Moreover, the questions asked in a scaffolding context are driven by the student's response. For instance, suppose a student is working on a division problem and forgot to put the remainder down, then a scaffolding hint might be ‘What about the remainder?’ Thus, a hint is a local question tailored to the response the tutee had provided: it is not a generic question driven by a curriculum script.

Evidence that tutors do prompt and scaffold

Using this guideline of what constitutes scaffolding, we can collapse the 44 tutoring strategies identified by McArthur *et al.* (1990) in their problem-solving protocols into three broad categories: prompting for self-explanation, scaffolding, and providing feedback. The frequency of use of each kind of strategy is: 84 instances of prompting to self-explain, 98 cases of scaffolding, and 161 cases of feedback. Out of the 161 cases of feedback 134 (or 84%) of them were direct corrective feedback, either responses such as ‘Right’, or ‘If you want to use division, that's fine’. As discussed earlier, direct corrective feedback means that the student is told explicitly whether what s/he did was correct, and if incorrect, what the correct answer is. In contrast, indirect or suggestive feedback can be a query, such as ‘So you think it's 126?’ Suggestive feedback can be considered a form of scaffolding, so that means 16% (21 of the 161) of the feedback should be re-classified as scaffolding. The large proportion of direct corrective feedback also implies that feedback seldom involves extensive explanations of deep conceptual issues, consistent with Merrill *et al.'s* (1992) results. The data sampled (84 cases of prompting and 98+21 cases of scaffolding) support both the self-explaining and co-constructing as viable hypotheses for tutoring effectiveness, along with receiving corrective feedback.

Similarly, a study of expert tutors also shows that they tend to prompt tutees for self-explanations. The very best tutors virtually never provide answers to problems (Lepper, Woolverton, Mumme and Gurtner, 199 1). Instead, they prompt students to generate their own corrections with statements such as ‘So, you think it's 126?’ request explanations (e.g., ‘Now tell me how you got that 6’), or ask leading questions (e.g., ‘Now which column is the one's column?’). This is consistent with Evens, Spitkovsky, Boyle, Michael and Rovick's (1993) data as well, in which 44% of the negative responses given by the tutors in face-to-face sessions can be considered general prompts. Even the general diagnoses the tutors used in McArthur's protocols cited above (‘Do you understand?’ ‘Have you heard of something called the additive inverse?’) are comparable to general prompts for self-explanations. Clearly, the upshot of these kinds of prompts is that they motivate the construction of self-explanations.

In sum, two broad categories of actions on the part of the tutors can result in a chain of responses from the students in a collaborative co-constructing way. The tutors can either prompt the students for self-construction or they can scaffold the students for co-construction.

ANALYSING TUTORING PROTOCOLS FOR EVIDENCE OF HOW LEARNING OCCURS: A CASE STUDY

In this section, a detailed analysis of learning in a physics tutoring context is presented as an exercise to see when and how learning occurs. First, we want to capture what is learned besides procedures for solving problems. Second, we want to discover the conditions that triggered learning; that is, which and whose actions are responsible for learning. The protocols for this analysis were taken from VanLehn, Chi, Baggett and Murray (1995). The tutoring protocols consist of a tutor tutoring a college student (tutee A) on three physics problems on mechanics, totalling 2592 lines of protocol. This analysis discusses primarily the first problem of tutee A, although some preliminary analyses of problems 2 and 3 have been done, as well as analysis of tutee Q. The tutoring takes place in a non-face-to-face interaction situation: they sat in different rooms. The interaction is achieved via telephoning. That is, when the tutee wants help, she can type ‘help’ and the tutor comes on. The tutor can see the student's ‘scratch pad’ on the computer screen at all times. The ‘scratch pad’ contains all of the student's steps and actions, such as her drawings and the equations she is using. This student, tutee A, had taken high school physics a semester ago. When confronted with the to-be-solved problem, the student clearly had trouble solving it.

Analysing misconceived and missing knowledge pieces

One could analyse the protocols in terms of what solution steps the tutee took, what equations were used, and what strategy of problem solving the solution steps implicate. This would be the standard approach to analysing problem-solving protocols. Here, in the present analysis, the focus was on capturing the knowledge that the tutee needed in order to proceed with the solution steps. Moreover, instead of doing a task analysis first to come up with all the knowledge pieces that a student needs to solve a problem (the template approach—again, another standard approach in cognitive task analysis), the protocols were analysed instead to extract the misconceived knowledge that the tutee has.

The protocols were first coded into segments that consisted of an *action* taken either by the tutor or the tutee. After several iterations, the set of interesting/ important actions taken by the tutor were: *reviews, summarizes, reminds, analogizes, prompts, gives didactic explanations, telling tutee what plan to do or what steps to take, gives corrective feedback, hints, motivates, asks questions, diagnoses misconceived knowledge, assesses missing knowledge, and assesses deviations from ideal;* and the actions taken by the tutee were: *calculates, applies equations, exhibits confusion, and asks questions.* General responses such as ‘Oh’ or ‘Okay’ were not coded. Thus, the result of the first coding of the protocols is an annotated version with 41 coded actions.

The second pass at the protocols consisted of identifying the pieces of knowledge that the tutee either lacked or was confused about. For problem 1, the tutee exhibited seven pieces of misconceived knowledge [see Table 1, Knowledge Pieces (KP) 1-7]. These are knowledge pieces that the student was confused about. For example, KP1 says that the student confused mass and weight. This is supported by comments made in line 132, in which she said, ‘I guess since it's falling downward, then it's the mass would equal the weight’. This confusion persisted in line 154 when she substituted a weight quantity for mass in the equation F = ma. All the misconceived KPs occur on more than one occasion in the protocols, so we are quite confident about their existence. KPs 8 and 9 (in Table 1) are not misconceived knowledge pieces. Rather they are simply missing knowledge pieces that the tutor assumed that the student lacked, and he taught these knowledge pieces to the student. They were not exhibited as misconceptions.

Table 1. Knowledge pieces that are misconceived (1-7) or missing (8, 9)

Yes	KP1)	Confuses mass with weight
Yes	KP2)	Does not realize that weight is the force due to gravity
Maybe	KP3)	Confuses direction of motion with direction of acceleration
No	KP4)	Does not realize that the **F** in **F** = m**a** *is* the sum of forces and not a single force
No	KP5)	Does not realize that an unknown force f can be one of the forces embedded in **F** for the sum of forces
No	KP6)	Does not realize that all forces acting on a body can simply be summed, without necessarily needing to know the exact equation corresponding to each force
Maybe	KP7)	Confusion about signs
No	KP8)	Draw forces on a diagram (a procedure)
Yes	KP9)	If an object in motion is slowing down or speeding up, then the object has an acceleration

Yes or No means the tutee either learned it or not. Maybe means there is no direct evidence one way or the other.

There are several things to note about these knowledge pieces. First, the majority of these misconceived knowledge pieces are general, in the sense that they cut across different physics problems and different students (such as confusing mass with weight, KP 1 in Table 1), but also in the sense that some of them have been discussed in the physics education literature (Reiner, Slotta, Chi and Resnick, in press). However, a small subset of them is specific to the problem at hand. One could say that KP7 (confusion about the signs) is specific to this problem, although this could be interpreted as a more general confusion as well. Thus, for two students anyway, there seems to be a common set of knowledge pieces that are misunderstood by them. Both tutee A and tutee Q were confused about KPs 1 to 7, for problem #1.

Besides misconceived knowledge, tutees may lack of course many knowledge pieces (what we referred to earlier as missing knowledge). The tutor seemed to have no trouble assessing what knowledge pieces the tutee was missing and taught these directly, even though the tutee had not exhibited any confusion about them. Missing KPs across the two students can easily be characterized as condition-action rules. A good example is KP9, where the tutor told the tutee the rule that 'IF something is speeding up, THEN there is acceleration'. This is clearly a *missing* knowledge piece as opposed to a *misconceived* knowledge piece because the tutee was not aware that she lacked such knowledge, and never exhibited any confusion about it. For tutee A in problem #1, the tutor assessed two pieces of missing knowledge.

KP8 is the other missing KP, although it is unique in that it is not only procedural in nature, but it encompasses all the other KPs plus more. It is a procedure of ‘drawing forces on the diagram’ that was introduced by the tutor, and which the tutee lacked. But ‘drawing the forces’ procedure incorporates knowledge about finding them, and realizing such things as that weight is a force but not mass (KP1), or that gravity contributes toward the weight (KP7). Thus, KP8 includes other KPs for which the tutee had misconceptions, but the tutor obviously was not aware of this since he did not break this KP8 into its components. This evidence is consistent with the interpretation that tutors generally cannot diagnose students' misconceived knowledge. In sum, KPs 1-7 are confusions exhibited by the tutee, whereas KP8 and KP9 are missing KPs that the tutor wanted the tutee to learn.

Perhaps the most important thing to note about KPs 1-7 is that they are subtle misunderstandings that are not explicated by the text or by the tutor in any explicit or direct way. In fact, it took several iterations of analyses to figure them out. The two that are most salient in this respect are KPs 5 and 6. They are subtle but important variations of KP4. That is, KP4 simply elaborates the big **F** in the syntactic form **F** = ma by specifying that **F** is the sum of forces; let us call them f_1, f_2, f_3 and so on. However, there are two corollaries to this KP4 represented by KP5 and KP6. KP5 is the realization that an unknown force (f_1, let us say) to be sought can be one of the forces that sums with other forces to the big **F**, and yet it may not have a name (such as the normal force) nor be given a label in the problem statement (i.e., no variable has been assigned to the force). KP6 is the misunderstanding that an unknown force (f_1) must have a name, and can only be sought by an equation that explicitly contains f, as a variable. The idea that students need to acquire is that any unknown force can simply be one force that has to be added with the other forces. Thus, if frictional force (f_1) is a component of the big **F**, one can actually find the value of the frictional force by using $F = f_1, + f_2 + ... = ma$, since the frictional force is a component of **F**, without needing to use an equation that contains f, explicitly as its unknown. This kind of confusion is exhibited by remarks about not knowing how to find the friction force (since they do not know an equation that has the frictional force as the unknown), instead of treating friction force as just one component force that can be summed in the **F**s.

What was learned

In the third pass through the protocols, we looked for evidence to see which of the nine knowledge pieces were learned. There are three categories of learning: yes, no, and maybe (shown on the left column in Table 1). Basically, a KP is coded as learned (a ‘yes’) if the tutee exhibited the correct application or use of that KP in relation to how it was used or conceived of before. For example, for KP1, the tutee was initially confused about weight and mass in line 132 of the protocol where she said explicitly that ‘I guess since it's falling downward then it's the mass would equal the weight’, and then again in line 153 when she explicitly used F = ma by substituting the weight of 670 N for the mass. Thus, there were at least two occasions when this confusion was initially exhibited.

This confusion was apparently removed later not only because she applied it in the equation correctly in line 383 ‘For the mass ... um, you wanted me to take 670 and divide it by 9.8?’ and again in line 391 ‘Um, well, yeah, that's the weight and I need to find mass’, but also because she realized that she had committed an error previously, in her comment in line 371 ‘I know he was saying about the mass ... I actually took the weight . . .’. Thus, she learned KP1 because there were three instances of getting the correct conception and applying it correctly in an equation form. Out of a possible nine KPs, she had learned three by the end of problem #1.

A KP is categorized as not learned (the ‘no’ category) if either of the following is true. First, the tutee exhibited a lack of understanding in some later portion of the protocols even though she could use the KP correctly earlier in the protocols. These early occasions of correct use usually arose from following the tutor's instruction closely, so that one is hard pressed to say that she had understood and learned it. Hence, the evidence constituting a failure to learn arises from manifestations of misunderstanding again at a later portion of the same problem. Similarly, if a KP is subsequently confused again in the protocols of problem #2, then it has not been learned. Four of the nine KPs were not learned by tutee A, according to these criteria.

The ‘maybe’ categories are those two KPs for which we have no definitive evidence that the tutee had either learned them or not. In both of these cases, the tutee followed the tutor's instruction in the application and use of these KPs. Because there were no further occasions to use or exhibit knowledge of these KPs, we could not ascertain whether, in fact, they were learned or not, hence, the ‘maybe’ category.

What actions triggered the learning

The most relevant analysis for this paper addresses how the tutee learned the three KPs. To find out, we examined the occasions when a KP was learned and noted which type of tutorial interactions resulted in the learning. KP1 was learned by the tutor's hint. The tutor merely warned ‘But the mass might give you a little trouble’. KP2 (that weight is a force due to gravity) was learned by the interchange occurring in lines 387-402. (Although we do not include the protocols in the appendix for brevity's sake, it still makes sense to refer to the line numbers to give the reader a sense of where in the protocols the event is occurring.) In this case, the tutee asked the tutor directly about her confusion in line 392 ‘But why don't I need to take the, um, the gravity into account, though?’ and the tutee learned from the response given by the tutor in line 400 and again in line 405. KP9, that speeding up means there is acceleration, is a piece of knowledge that was missing from the tutee's knowledge. The tutor, in the context of an error exhibited by the tutee, corrected the direction of the sign by telling the tutee in line 451 and line 456 that ‘IF it's accelerating downward (taking it as positive), THEN that would mean she's speeding up’. The evidence for her learning it is that she was then able to use and rephrase it as ‘IF it is slowing down, THEN the acceleration is decreasing’. In sum, the two misconceived KPs are learned either by the tutor's hint or answer to tutee's explicit question, whereas the missing knowledge piece was learned by an explicit explanation to the tutee's query.

A complementary analysis is to see what instructional actions did the tutor not take, or what actions did the tutor take that did *not* result in learning? For example, did the tutor diagnose misunderstanding? Diagnosing student's misunderstanding, as mentioned earlier. is when a student clearly exhibited confusion, and the tutor would pursue it and try to figure out what the misunderstanding was and correct it. Out of the 41 tutorial actions coded for problem #1, there were at least 6 occasions on which the tutee manifested confusion. In five of these cases, the tutor simply ignored them. For example, in lines 133-154, the tutee was confused about mass and weight. The tutor simply ignored this confusion rather than clarifying it. Instead, the tutor pursued his own plan of making the tutee find all the forces.

In contrast to diagnosing misconceived knowledge or misunderstanding, the tutor seemed quite competent at assessing missing knowledge. Assessing missing knowledge occurred when the tutor, from prior experiences in tutoring or teaching, realized that either a certain concept is difficult or some aspect of the problem may present difficulties. In such a case, the tutor assessed whether the tutee knew this piece of knowledge even though the tutee may not have exhibited any lack of knowledge or misunderstanding. A good example of this assessment of missing knowledge occurred in lines 247-251 where the tutor plunged ahead and told the tutee to beware of the signs even though there was no evidence that the tutee would be confused about the signs. This is because tutee A was the second subject, and the tutor experienced the previous tutee Q's difficulties with the signs.

What did the tutor do mostly?

If the tutor was not diagnosing and addressing the tutee's misunderstandings, what was it that the tutor did? The tutor mostly pursued his own plan. There were two related subplans within his plan. First, the tutor wanted the tutee to draw all the forces on the diagram, and then to sum the forces. The tutor executed his plan at a fairly high level, resorting to a more detailed level only when the tutee requested information and failed to execute any actions. The tutor also gave feedback, of both the corrective and reinforcing kinds. As consistent with Merrill *et al.*'s data (1992), the corrective feedback tended to occur mostly in syntactic actions. For instance, when the tutee added some wrong forces together, or when mass was substituted for weight in the equation, the tutor interjected and gave corrective feedback. In contrast, the tutor seldom gave corrective feedback when there was a conceptual misunderstanding (which was coded above as a failure to diagnose). The tutor also gave *reinforcing feedback* when some correct action was taken, usually in equation form or in drawing of forces on the diagram. Here, even though the tutee was not yet sure that she was doing it correctly, the tutor interjected quickly and said that it was correct. This is another form of ignoring tutee's understanding, but this is not counted as such in the above analyses on ignoring tutee's misunderstanding.

There were few occasions of prompting and scaffolding, partly because of the nature of this tutoring task (problem solving), which tended to be very scripted in getting the right equation down, and partly because the non-face-to-face nature of this task precluded spontaneous scaffolding that can occur in a face-to-face dialogue situation.

In sum, what this exercise tells us is the following. Overall, the student did ultimately learn how to solve the problem, primarily because the tutor insisted on her following his prescripted plan. This is completely consistent with results from an Anderson, Conrad and Corbett (1989) type of computer-based intelligent tutoring system. That is not to say that the tutee has learned that much. But by following a prescripted plan a couple of times in a tutorial session, a tutee will pick up the right procedures for solving problems, consistent with findings in Sleeman, Ward, Kelly, Mantinak and Moore, (1991) and Chi and VanLehn (1991). However, they learned the correct conception for only two of the seven exhibited misconceptions. Two of the three learned knowledge pieces were learned not by following a prescripted *plan,* but rather, by the opportunities for the tutee and tutor to interact. To summarize, KP1 was learned from the tutor's hint, KP2 was learned by a question-asking exchange initiated by the tutee, and KP9 was learned from being directly told of the condition-action rule. No learning ever resulted from long-winded didactic explanations (of which there were several), possibly because these explanations did not address the tutee's misunderstanding directly; nor did teaming arise from correctly diagnosing misconceived knowledge. Thus, the tutee learned not from the tutor's instructional skills such as diagnosing misconceived knowledge or giving didactic explanations, but rather, from interactions with the tutor in the form of question-asking by the tutee and scaffolding in the form of hinting by the tutor, as well as from being directly told.

CONCLUSION

The survey of the literature suggests that tutoring benefits may arise from two of the three hypotheses considered in this paper. Specifically, there is indirect evidence to suggest that tutoring effectiveness in the form of promoting deep understanding (such that misconceptions are removed) results from the self-construction of knowledge on the part of the tutee, as well as co-construction from scaffoldings. On the other hand, tutoring effectiveness in the form of deep understanding does not seem to arise from tutoring skills per se, such as diagnosing misunderstandings or giving didactic explanations.

The exercise of analysing this case study points out another important dimension of assessing learning. That is, in measuring tutoring effectiveness and what accounts for learning, one needs to discriminate between assessing the tutee's ability to solve problems (which can result from tutors' actions that are more directive and less interactive because these directive actions require that the tutees follow the tutors' prescripted plan) versus an analysis that examines tutor-tutee interaction in the context of understanding concepts and correcting misunderstandings. A tutee seems to be able to achieve the latter goal primarily from the co-construction of knowledge with the tutor, by having the tutor scaffold the tutee, or by having the tutor answer the tutee's question.

What is to be taken away from this analysis of a single subject is not just the generalization that self-explanations and scaffolded explanations are definitely the sources of tutoring benefits, but also a method of analysis that is not focused on learning the procedures of problem solving. Instead, it focuses on what new knowledge is learned, and whether old misconceived knowledge is removed.

REFERENCES

Anderson, J. R., Boyle, C. F. and Reiser, B. J. (1985). Intelligent tutoring systems. *Science,* 228, 456-468.

Anderson, J. R., Conrad, F. G. and Corbett, A. T. (1989). Skill acquisition and the LISP tutor. *Cognitive Science,* 13(4), 467-506.

Azmitia, M. and Perlmutter, M. (1989). Social influences on children's cognition: state of the art and future directions. In H. W. Reese (Ed.), *Advances in child development and behavior* (pp 89-144). New York: Academic Press.

Barwise, J. and Perry, J. (1983). *Situation and attitudes.* Cambridge, MA: MIT Press.

Brown, A. L. and Palinscar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Cognition and instruction: issues and agendas* (p. 395). Hillsdale, NJ: Lawrence Erlbaum.

Bruer. J. T. (1993). *Schools for thought: a science of learning in the classroom.* Cambridge, MA: MIT Press.

Chi, M. T. H. and Bjork, R. (1991). Modelling expertise. In D. Druckman and R. Bjork (Eds.), *In the mind's eye: understanding human performance* (pp. 57-79). Washington, DC: National Academy Press.

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

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P. and 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. and La Vancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science,* 18, 439-477.

Coleman, E. (1992). Reflection through explanation. Paper presented at the American Educational Research Association meeting, San Francisco, April.

Collins, A., Brown, J. S. and Newman, S. E. (1989). Cognitive apprenticeship: teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: essays in honor of Robert Glaser (pp.* 435-494). Hillsdale, NJ: Lawrence Erlbaum.

Cooper, C. R., Ayers-Lopez, S. and Marquis, A. (1982). Children's discourse during peer learning in experimental and naturalistic situations. *Discourse Processes,* 5, 177-191.

Evens, M., Spitkovsky, J., Boyle, P., Michael, J. and Rovick, A. (1993). Synthesizing tutorial dialogues. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 137-142). University of Colorado: Boulder.

Forman, E. A. and Cazden, C. B. (1985). Exploring Vygotskian perspectives in education: the cognitive value of peer interaction. In J. V. Wertsch (Ed.), *Culture, communication and cognition: Vygotskian perspectives* (pp. 323-347). New York: Cambridge University Press.

Glenberg, A. M., Wilkinson, A. C. and Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory and Cognition*, 10, 597-602.

Graesser, A. C., Person, N. K. and Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, 9, 495-522.

Hume, G. D., Michael, J. A., Rovick, A. A. and Evens, M. W. (1996.). Hinting as a tactic in one-on-one tutoring. *Journal of the Learning Sciences,* 5(1), 23-47.

King, A. (1994). Guiding knowledge construction in the classroom: effects of teaching children how to question and how to explain. *American Educational Research Journal,* 31, 338-368.

Lepper, M. R., Woolverton, M., Mumme, D. L. and Gurtner, J. L. (1991). Motivational techniques of expert human tutors: lessons for the design of computer-based tutors. In S. P. Lajoie and S. Derry (Eds.), Computers as cognitive tools. Hillsdale, NJ: Lawrence Erlbaum.

McArthur, D., Stasz, C. and Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction,* 7, 197-244.

Merrill, D. C., Reiser, B. J. and Landes, S. (1992). Human tutoring: pedagogal strategies and learning outcomes. Paper presented at the annual meeting of the American Research Association, San Francisco, April.

Merrill, D. C., Reiser, B. J., Ranney, M. and Trafton, J. G. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal of the Learning Sciences*, 2, 277-305.

Newman, D., Griffin, P. and Cole, M. (1989). *The construction zone: working for cognitive change in school.* Cambridge: Cambridge University Press.

Palinscar, A. S. and Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. *Cognition and Instruction,* 1, 117-175.

Perret-Clermont, A. N., Perret, J. F. and Bell, N. (1991). The social construction of meaning and cognitive activity in elementary school children. In L. B. Resnick, J. M. Levine and S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 41-62). Washington, DC: American Psychological Association.

Person, N. K., Graesser, A. C., Magliano, J. P. and Kreuz, R. J. (1994). Inferring what the student knows in one-to-one tutoring: the role of student questions and answers. *Learning and Individual differences,* 6, 205-229.

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

Reiner, M., Slotta,. J. D. Chi, M. T. H. and Resnick, L. B. (In press). An underlying materialistic commitment in naive thought. *Cognition and Instruction.*

Renkl, A. (in press). Learning from worked-out examples: a study on individual differences. *Cognitive Science.*

Rogoff, B. (1990), *Apprenticeship in thinking: cognitive development in social context*. New York: Oxford University Press.

Roschelle, J. (1992). Learning by collaborating: convergent conceptual change. *The Journal of the Learning Sciences, 2,* 235-276.

Scardamalia, M., Bereiter, C. and Sternbach, R. (1984). Teachability of reflective processes in written composition. *Cognitive Science,* 8, 173-190.

Schliemann, A. D. and Nunes, T. A. (1990). A situated schema of proportionality. *British Journal of Developmental Psychology,* 8, 259-268

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

Skon, L., Johnson, D. and Johnson, R. (1981). Cooperative peer interaction effort versus individual effort: effect on the acquisition of cognitive reasoning strategies. *Journal of Educational Psychology,* 73, 83-92.

Sleeman, D., Ward, R. D., Kelly, E., Martinak, R. and Moore, J, (1991). Overview of recent studies with PIXIE. In P, Goodyear (Ed.), *Teaching knowledge and intelligent* tutoring (pp. 173-185). Norwood, NJ: Ablex.

Stone, C. A. (In press). What's missing in the metaphor of scaffolding? In E. A. Forman, N. Minick and C. A. Stone (Eds.), *Context for learning.* New York: Oxford University Press.

Suchman, L. A. (1987). *Plans and situated actions.* New York: Cambridge University Press.

Teasley, S. D. (1992). Communication and collaboration: the role of talk in children's peer collaborations. Unpublished doctoral dissertation. University of Pittsburgh.

Tudge, J. R. H. and Rogoff, B. (1989). Peer influences on cognitive development: Piagetian and Vygotskian perspectives. In M. Bornstein and J. Bruner (Eds.), *Interaction in human development* (pp. 17-40). Hillsdale, NJ: Lawrence Erlbaum.

VanLehn, K. A., Chi, M. T. H., Baggett, W. and Murray, R. C. (1995). Progress Report: towards a theory of learning during tutoring. Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

VanLehn, K. A., Jones, R. M. and Chi, M. T. H. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences*, 2(1), 1-60.

Webb, N. M. (1989). Peer interaction and learning in small groups. In N. Webb (Ed.), Peer interaction, problem solving, and cognition: multidisciplinary perspectives [Special Issue] *International Journal of Education Research*, 13, 21-39.