APPLIED COGNITIVE PSYCHOLOGY, VOL. 9, 495–522 (1995)

Collaborative Dialogue Patterns in Naturalistic One-to-One Tutoring

ARTHUR C. GRAESSER
The University of Memphis

NATALIE K. PERSON
Rhodes College

JOSEPH. P. MAGLIANO
University of Chicago

SUMMARY

Naturalistic one-to-one tutoring is more effective than traditional classroom teaching methods, but there have been few attempts to examine the features of normal tutoring that might explain its advantage. This project explored dialogue patterns in two samples of naturalistic tutoring with normal unskilled tutors (as opposed to expert tutors): graduate students tutoring undergraduates in research methods and high school students tutoring 7th graders in algebra. We analysed the extent to which those tutoring protocols manifested components that have been emphasized in contemporary pedagogical theories and intelligent tutoring systems: active student learning, sophisticated pedagogical strategies, specific examples and cases, collaborative problem solving and question answering, deep explanatory reasoning, convergance toward shared meanings, feedback, error diagnosis and remediation, and affect. The most prominent components consisted of collaborative problem solving, question answering, and explanation in the context of specific examples. We identify frequent dialogue patterns that characterize these collaborative processes.

It is well documented that one-to-one tutoring is superior to normal learning experiences in traditional classroom settings. The effect size of the advantage of tutoring over classroom has ranged from 0.4 to 2.3 standard deviation units (Bloom, 1984; Cohen, Kulik and Kulik, 1982; Mohan, 1972). These effects are quite robust and have stimulated us to inquire what processes are responsible for the advantages of normal one-to-one tutoring.

Perhaps the robust advantages of tutoring in these studies can be attributed to the selection of skilled tutors. An ideal tutor would have substantial domain

Address correspondence to Arthur C. Graesser, Department of Psychology, The University of Memphis, Memphis, TN 38152, USA (e-mail graesser@memphis.edu).
This research was funded by the Office of Naval Research (N00014-90-J-1492, and N00014-92-J-1826) and the Center for Applied Psychological Research at the University of Memphis. We are indebted to John Cady for providing access to the seventh grade algebra tutoring sessions. We are also indebted to those individuals who served as judges in analyses of the tutoring protocols: Sailaja Ballejepali, Katinka Dijkstra, John Huber, Brenda Johnson, Roger Kreuz, Mark Langston, Joe Magliano, and Rolf Zwaan.

CCC 0888–4080/95/060495–28

Accepted 19 December 1994

knowledge, extensive training on effective tutoring techniques, and several years of tutoring experience. However, the vast majority of tutors in these studies were hardly ideal tutors. Most tutors in a school system are peers of the students, slightly older students, paraprofessionals, and adult volunteers—not highly skilled tutors (Fitz-Gibbon, 1977). Skilled tutors are the exception, rather than the rule. Moreover, according to Cohen *et al.*'s (1982) meta-analysis of 52 tutoring studies, the impact of tutoring on learning is not significantly related to the amount of tutor training or to age differences between tutor and student. Indeed, the peers of students often do an excellent job serving as tutors for students having problems (Fantuzzo, King and Heller, 1992; Mohan, 1972; Rogoff, 1990). These findings are somewhat counterintuitive. Most of us would expect that tutoring age and expertise would improve learning outcomes. Perhaps the training and expertise of tutors are minimal in normal tutoring sessions. Perhaps a tutor needs extensive training on both domain knowledge and tutoring strategies before tutoring expertise shows appreciable gains in learning outcome. Nevertheless, the counterintuitive findings do support one provocative explanation of the advantages of tutoring: the advantage may be attributed to conversational dialogue patterns of unskilled tutors rather than to esoteric pedagogical strategies of skilled tutors.

The notion that conversation mechanisms might have a major impact on learning is hardly a revolutionary idea. The importance of collaborative 'talk' has been emphasized in contemporary theories of education, literacy, and situated cognition (Bakhtin, 1981; Gee, 1989, Goodwin and Heritage, 1990; Greeno, 1991; Hawkins and Pea, 1987; Lave and Wenger, 1987; Roschelle, 1992; Suchman, 1987; Vygotsky, 1978). But what is it about conversational dialogue that might explain its impact on learning? Researchers need to dissect conversational patterns during tutoring and to relate them to learning outcomes.

Other than our own investigations of tutorial dialogue, fewer than a dozen studies have systematically examined the process of naturalistic tutoring at a fine-grained level (Fox, 1991, 1993; Leinhardt, 1987; Lepper, Aspinwall, Mumme and Chabay, 1990; Lepper, Woolverton, Mumme and Gurtner, in press; McArthur, Stasz and Zmuidzinas, 1990; Merrill, Reiser, Ranney and Trafton, 1992; Miyake and Norman, 1979; Putnam, 1987; Van Lehn, 1990). In most of these studies, there were very few tutors and the tutors were relatively skilled. Therefore, it is uncertain whether the findings in these studies would generalize to a larger sample of 'normal' unskilled tutors. It takes a great deal of time and effort to perform in-depth qualitative analyses of tutorial interaction. Because of the limited sample sizes in qualitative process-oriented studies, there has been no attempt to relate the components of the tutorial process to student achievement or to learning outcomes.

Our research has analysed patterns of tutorial dialogue in a comparatively large sample of tutoring sessions with normal tutors (Graesser, 1992, 1993a,b; Graesser and Person, 1994; Person, Graesser, Magliano and Kreuz, 1994; Person, Kreuz, Zwaan and Graesser, in press). These previous studies examined pedagogical strategies, feedback mechanisms, question asking, question answering, and pragmatic assumptions during the normal tutoring process. The present study examines the extent to which these naturalistic tutoring protocols manifest learning components that have been emphasized in contemporary pedagogical theories and intelligent tutoring systems. These components include:

1. *Active student learning.* Instead of the student being a passive recipient of information, the educational experience should encourage active student learning.
2. *Sophisticated pedagogical strategies.* A good teacher/tutor should implement sophisticated pedagogical strategies that are effective in promoting learning.
3. *Anchored learning in specific examples and cases.* A good teacher/tutor should anchor the material in specific examples and cases rather than relying on didactic, declarative information.
4. *Collaborative problem solving and question answering.* A good learning experience involves a balanced collaboration between the teacher/tutor and the student while they solve problems and answer questions.
5. *Deep explanatory reasoning.* The teacher/tutor and student should focus on deep conceptual models and explanations rather than superficial facts.
6. *Convergence toward shared meanings.* The teacher/student should achieve shared knowledge, a 'meeting of the minds'.
7. *Feedback, error diagnosis, and remediation.* A good teacher/tutor should quickly give feedback on the quality of student contributions. When a student makes an error, the teacher/tutor should identify the error, correct the error, diagnose the misconception that explains the error, and rectify the misconception.
8. *Affect and motivation.* A good teacher/tutor bolsters student motivation, confidence, and self-efficacy while mastering the material.

We will ultimately argue that the most salient components of normal tutoring consist of collaborative problem solving, question answering, and explanatory reasoning in the context of specific examples. Thus, components 3, 4, and 5 are most prevalent in normal tutoring. As the student and tutor collaboratively solve problems, answer questions, and construct explanations, each contributor throws fragments of information into a shared workspace in a distributed, incremental fashion. However, the evolution of these dialogue streams is not entirely unsystematic. We will identify some systematic dialogue patterns that impose structure onto the collaborative exchange.

We will argue that the other learning components (1, 2, 6, 7, and 8) are underdeveloped, defective, or virtually non-existent in normal tutoring. There are two important implications of this claim. First, tutors need to be trained to implement these other five components in a successful fashion. Second, the use of these other five components by skilled tutors and intelligent tutoring systems should yield incremental gains in learning outcomes, over and above what is supplied by normal tutors.

The next section describes two samples of normal tutoring that we analysed in detail. We will subsequently report some analyses of these tutoring protocols and examine the status of the eight learning components.

TWO SAMPLES OF NATURALISTIC ONE-TO-ONE TUTORING

Research methods corpus

Graduate students in the psychology department at the University of Memphis tutored undergraduate students on troublesome topics in a research methods course.

All 27 students in the course were tutored as part of a course requirement, so there was a full range of student achievement (i.e. not just underachieving students). The tutoring sessions were a course requirement, counting 6% of the final course grade. There were three tutors, all of whom had received A's in a graduate-level research methods course. Therefore, the corpus involved 'cross-age' tutoring, which is one of the common types of tutoring in school systems. The tutors had never tutored in the area of research methods before this study, but they had occasionally tutored on other topics.

A total of 54 1-hour tutoring sessions were videotaped. The room used for tutoring was equipped with a video camera, a television set, a marker board, coloured markers, and a textbook for the course. The camera was positioned so that the student, the tutor, and the entire marker board were in sight. Therefore, the transcripts of the tutoring sessions included both spoken utterances and messages on the marker board. The transcribers were instructed to transcribe the entire tutoring sessions, including all 'ums', 'ahs', word fragments, broken sentences, and pauses. Messages on the marker board were sketched in as much detail as possible.

The sessions covered six troublesome topics in the undergraduate research methods course. The topics were: operational definitions of variables, graphs, inferential statistics, the evolution of hypothesis to design, factorial designs, and interactions. An index card was prepared for each topic; three to five subtopics were listed under each subtopic. The tutor was asked to cover the topic and subtopics on an index card during the course of the 1-hour tutoring session. The tutors were not given a specific format to follow, but they were told to resist the temptation of simply lecturing to the student. It should be noted that the students were exposed to the material covered on a topic before they participated in the tutoring session. First, the topic had been covered in a classroom lecture by an instructor (not the tutors). Second, the student and the tutor had read specific pages in a research methods text.

The 27 students participated in two tutoring sessions, yielding 54 sessions altogether. Each student was assigned randomly to two of the tutors. However, only 44 of these sessions had sufficient video and auditory quality that they could be transcribed. Therefore, analyses were performed on 44 tutoring sessions (25 out of the 27 students).

Examination scores and final grades were available for the 25 undergraduate students, so we could investigate the relationship between student achievement and tutoring processes. A total examination score was based on three objective examinations completed throughout the semester; there was a total of 150 four-alternative, forced-choice questions. These questions were prepared by authors of an instructor's manual and by graduate students other than the three tutors. The instructor and the tutors did not know what questions were included on the objective examinations. Therefore, the lectures and tutoring session were not tuned to the content of the objective examination. The 25 students had a mean examination score of 100.6 (Standard deviation = 11.4). Regarding the final grade in the course, five students received an A, nine received a B, ten received a C, and four received a C− or D.

Algebra corpus

This corpus consisted of 22 tutoring sessions in which high school students tutored 7th graders on troublesome topics in algebra. There were 13 students who were having trouble with particular topics in their algebra course, according to their teachers. The teachers had recommended that they receive one-to-one tutoring on these troublesome topics. Ten tutors normally provided the tutoring services for the middle school. On average, a tutor had 9 hours of prior tutoring experience before tutoring a student in this sample. The corpus of tutoring sessions included almost all of the tutoring sessions that occurred in the middle school for 7th graders learning algebra for a 1-month period. Therefore, this was a *bona fide* naturalistic corpus of tutoring sessions in a school system. Unlike the research methods corpus, the tutoring sessions in this algebra course were remedial activities for underachieving students. Unfortunately, grades and test scores were not available for these students, so it was not possible to assess the relationship between achievement and tutoring processes.

Almost all of the tutoring sessions covered three tutoring topics that are frequently problematic to 7th graders. These include: (1) calculation of positive and negative numbers in algebraic expressions; (2) constructing equations from algebra word problems; and (3) fractions. An examination and chapter excerpt from a textbook were normally available to the student and tutor when a tutoring session began. The student typically had a poor examination score, with particular problems marked as being incorrect. On the basis of this information, the tutor and student completed the tutoring session in whatever fashion they saw fit. The tutoring session lasted approximately 1 hour, which was comparable to the research methods corpus.

A research assistant from the University of Memphis videotaped the sessions in a similar manner as the sessions were videotaped in the research methods corpus.

Scoring tutoring transcripts on content variables

Previous reports and articles have discussed how transcripts were analysed on content variables (Graesser, 1992, 1993a; Graesser and Person, 1994; Graesser, Person and Huber, 1992, 1993; Person *et al.*, 1994). The details of these scoring procedures are therefore not discussed in this article. Trained research assistants were capable of reliably coding most of the data: segmenting transcripts into speech act units, assigning speech acts to speech act categories, identifying questions, assigning questions to question categories, identifying mechanisms that generate questions, and classifying tutor feedback.

Judges needed to have more expertise in the case of some coding analyses. One such analysis consisted of the quality of a contribution in a tutoring excerpt. There were four levels of contribution quality: (1) error-ridden answer; (2) vague; (3) partially correct; and (4) completely correct. The judges needed to have a large amount of domain knowledge about research methods to make these judgements. Therefore, these judgements were made by professors, postdocs, and 4th year graduate students in experimental psychology. Other analyses that required special expertise involved global levels of tutorial dialogue (e.g. whether an excerpt involved error-remediation or a curriculum script). In these cases, the judges needed to have sophisticated knowledge about the tutoring process in addition to extensive domain knowledge. A pair of judges collaboratively supplied judgements when categories or dimensions required high expertise.

One important analysis was a *collaborative transition matrix*. This analysis traced the evolution of the collaborative exchange between tutor and student when a question was being answered. We observed the status and quality of contribution $n+1$, given that the tutor and student had together achieved a particular level of quality via contributions through n. A separate transition matrix was prepared for the tutor and the student. This analysis permitted us to quantify the quality of the information that was supplied by each participant. We could also identify patterns of tutorial dialogue at a fine-grained level. Dialogue facilities in intelligent tutoring systems have been developed at this fine-grained level (Blandford, 1993; Clancey, 1987; Moore and Paris, 1993; Swigger, 1991; Woolf, 1991; Woolf and McDonald, 1984), but there have been no comparable analyses of normal tutorial dialogue.

ACTIVE STUDENT LEARNING

There is the optimistic vision that learners are active, self-motivated, inquisitive individuals who are sensitive to deficits in their knowledge and who initiate self-regulatory strategies that correct the deficits. Active learners select their own problems to work on, ask questions, and construct their own learning environment. This vision is generally regarded as a fantasy for most learners in most situations. However, researchers have frequently advocated educational settings that engage students in active learning or that train students how to acquire self-regulatory learning strategies (Brown, 1988, 1992; Bruner, 1961; Carroll, 1990; Papert, 1980; Piaget, 1952; Scardamalia and Bereiter, 1991; Schank and Farrell, 1988; Wittrock, 1990; Zimmerman, Bandura and Martinez-Pons, 1992). Perhaps the advantage of tutoring over classroom settings can be attributed to more active learning during tutoring. Whereas it is impractical to have 30 students actively learning in a classroom simultaneously, one-to-one tutoring provides this opportunity.

Nevertheless, we found little support for active student learning when we analysed our naturalistic tutoring protocols. It was the tutor, rather than the student, who dictated the course of tutoring. Students rarely initiated exchanges that exerted control over the tutorial dialogue. Only 5% of the subtopics were initiated by the students in the research methods corpus, whereas 10% were student-initiated in the algebra corpus. When students did initiate a new subtopic, they normally brought up an example problem or concept that they were having difficulty with (e.g. 'I had trouble with problem 4', 'I don't understand what an antagonistic interaction is'). The students normally brought up these subtopics after being prompted by the tutor (e.g. 'So what are you having problems with?'). The students never set the agenda for the tutoring session. Thus, the tutors carried the burden of setting the agenda, introducing subtopics, and proposing problems to solve.

At the micro-level, there was one finding that indicated that students are somewhat more active in tutoring contexts than in classroom settings. Student questions are more frequent in tutoring settings than in classroom settings (Graesser and Person, 1994). The mean number of student questions per hour was 21.1 in the research methods corpus and 32.2 in the algebra corpus. In contrast, a particular student in a classroom setting asks only 0.11 question per hour; an entire classroom of students asks only 3.0 questions per hour (Dillon, 1988; Graesser and Person, 1994). In spite of the comparatively high incidence of student questions during tutoring, it is the tutor who asked the lion's share of the questions. We found that

80% of the questions in a tutoring session were asked by the tutor (82% in the research methods corpus and 78% in the algebra corpus). This percentage is somewhat lower than the percentage of teacher questions in a classroom (96%).

Most of the questions that students asked during tutoring did not address their own knowledge deficits. Knowledge deficit questions occur under the following conditions: (1) when the student encounters an obstacle in a plan or problem; (2) when the student detects a contradiction; (3) when the student detects an anomalous fact or event; (4) when there is an obvious gap in the student's knowledge base; or (5) when the student needs to make a decision among a set of alternatives that are equally attractive (Graesser and McMahen, 1993; Ram, 1991; Schank, 1986). Only 29% of the student questions during tutoring were knowledge-deficit questions (Graesser and Person, 1994), which amounts to 7.7 questions per hour. Most of the student questions (54%) were attempts to confirm the validity of their own beliefs (e.g. 'Doesn't a factorial design have two independent variables?') or to confirm common ground (e.g. 'Are you talking about the second condition?').

The good students in the research methods corpus did not ask more questions, and did not ask more knowledge-deficit questions (Person *et al.*, 1994). The correlations were low between the examination scores and: (1) the total number of student questions ($r = -0.22$); and (2) the proportion of student questions that addressed knowledge deficits ($r = 0.15$). These correlations were low when the final course grade was the measure of achievement instead of examination scores. Other researchers have also failed to show a positive relationship between the incidence of question asking and achievement (Fishbein, Eckart, Lauver, van Leeuwen and Langmeyer, 1990). Students may need to be trained how to ask good questions that reflect their knowledge-deficits (King, 1992; Pressley, 1990).

Given these results, intelligent tutoring systems need to impose special strategies of transferring control to the student if there is a commitment to promote active learning. Such strategies are not in the repertoire of the normal tutor. For example, there are hypermedia systems that require students to actively explore the material and to ask questions (Graesser, Langston and Baggett, 1993; Schank, Ferguson, Birnbaum, Barger and Greising, 1991; Sebrechts and Swartz, 1991; Spiro, Feltovich, Jacobson and Coulson, 1992). Computer systems with mixed initiative dialogue facilities prompt students to ask questions, answer questions, generate examples, and exert control over their learning environment (Clancey, 1987; Woolf, 1991; Woolf and McDonald, 1984).

SOPHISTICATED PEDAGOGICAL STRATEGIES

Researchers have advocated a number of sophisticated tutoring strategies that promote deep comprehension of the material. Examples of these strategies are the Socratic method (Collins, 1985; Stevens, Collins and Goldin, 1982), inquiry teaching (Collins, 1988), the reciprocal training method (Brown and Palincsar, 1989; Palincsar and Brown, 1984), and modelling–scaffolding–fading (Collins, Brown and Newman, 1989; Rogoff, 1990). In the Socratic method, for example, the tutor asks questions that lead students to discover their own misconceptions during the course of answering the questions. This method is illustrated below (see Collins, 1985).

Tutor: Do you know why it rains a lot in Oregon and Washington?
Student: There is a warm current passing over cool land.
Tutor: Do the Cascade Mountains there affect the amount of rainfall?
Student: No, no, no.
Tutor: How can the Andes affect the amount of rain in the Amazon and the Cascades not affect the rain in Oregon?

When students are confronted with this line of questioning, they discover their misconceptions and revise their explanatory hypotheses. The tutor never tells the students they are wrong, never directly supplies the correct information, and never articulates the students' misconceptions.

We examined the research methods corpus and the algebra corpus for the occurrence of the above four sophisticated tutoring strategies: (1) the Socratic method; (2) inquiry teaching; (3) the reciprocal training method; and (4) modelling–scaffolding–fading. These strategies were virtually non-existent in the two samples of normal tutoring. It apparently takes a large amount of training and experience for tutors to use these sophisticated pedagogical strategies. It is therefore not surprising that the strategies were non-existent in our sample of 13 tutors, and presumably are non-existent in real school settings. There will, hopefully, be high payoffs in learning outcomes when these strategies are implemented in skilled tutors and intelligent tutoring systems.

In the absence of sophisticated tutoring strategies, what strategies did the tutors in our sample typically adopt? The most prevalent strategies at the macro-level involved the use of *curriculum scripts* (McArthur *et al.*, 1990; Putnam, 1987). Curriculum scripts consist of a set of subtopics, example problems, and questions that the tutor selects for the tutoring session. We found that the vast majority of the tutor's subtopics, example problems, and questions were motivated by their curriculum scripts (67% in the research methods corpus and 79% in the algebra corpus). In the case of the research methods corpus, the tutor selected the subtopics in a top-down fashion. The selected subtopics had a close correspondence to the information in the chapter excerpts and the index cards supplied by the tutors (with the major topic and three to five subtopics). The example problems selected by the tutor usually came directly from the book. A tutor frequently introduced the same examples, subtopics, and questions to several students that were tutored on a particular topic. In the algebra corpus, the tutor normally selected a problem from the student's examination or the textbook. After selecting the problem, the tutor typically coached the student to a solution, or the tutor and student collaboratively solved the problem. It should be noted that the curriculum script is not necessarily a rigid structure in terms of the selection and ordering of the material. According to McArthur *et al.* (1990), the tutor revises and replans the agenda throughout the course of the tutoring session. The revising and replanning are presumably influenced by the student's performance.

ANCHORED LEARNING, CASES, AND EXAMPLE PROBLEMS

The importance of grounding learning in particular examples or cases is widely acknowledged (Anderson, Conrad and Corbett, 1989; Hammond, Seifert and Gray, 1991; Kolodner, 1993; LeFevre and Dixon, 1986; Schank and Jona, 1991; Sleeman and Brown, 1982; Sweller, 1988; Williams, 1992). At one extreme, there are *authentic anchored cases* (Bransford, Goldman and Vye, 1991; Goldman, Pellegrino and Bransford, 1993; Schank and Jona, 1991); these cases are challenging (requiring several minutes or hours to solve), are anchored in real-world situations, and genuinely motivate the student. At the other extreme, *symbolic examples* are decontextualized and lack concrete referents (e.g. 'if you have one variable A and another variable B, what would a curvilinear relationship look like?'). *Concrete examples* are decontextualized, but have concrete referents (e.g. 'Suppose there are two types of therapy (drugs versus psychotherapy) and the dependent measure is hours of sleep per night . . . ?'). Symbolic and concrete examples are the mainstay of textbooks, workbooks, and classroom activities.

Examples were very prevalent in the tutoring protocols that we analysed. Most of the tutor questions were asked in the context of a particular example in the research methods corpus (67%) and in the algebra corpus (92%). These example problems were normally selected by the tutor rather than by the student. Most example problems were selected from the textbook or examination rather than being invented by the tutor to fit the idiosyncratic needs of the student. Approximately half of the examples were concrete and half symbolic, whereas virtually none of them were authentic anchored cases. Therefore, learning environments and intelligent systems that have authentic anchored cases go a step beyond normal tutoring activities.

COLLABORATIVE PROBLEM SOLVING AND QUESTION ANSWERING

One of the salient trends in contemporary education is the increasing interest in collaborative learning and problem solving (Bransford *et al.*, 1991; Brown, 1992; Dansereau, 1988; Goldman *et al.*, 1993; Kourilsky and Wittrock, 1992; Roschelle, 1992). There are several pedagogical and social benefits when two or more individuals collaboratively solve problems or work on projects of mutual interest. For example, when one learner encounters a knowledge gap or obstacle, the other learner can fill in missing links and alternative solutions. The size of the learning groups have ranged from dyads to research teams and 'learning communities' with five to ten students (Brown and Campione, 1990; Rogoff, 1990; Slavin, 1983). The group is normally more successful when there is an expert in the group, although there sometimes are learning gains when the group consists entirely of peers. One-to-one tutoring is merely one of many possible configurations of collaborative learning and problem solving.

Researchers in discourse processing, sociolinguistics, and communication have emphasized the fact that conversations are joint collaborative efforts between speech participants (Clark and Schaefer, 1989; Kreuz and Roberts, 1993; Levin and Moore, 1977; Schegloff, 1991; Suchman, 1987). Communication is not simply a matter of two speakers taking turns transmitting messages, such that speaker A transmits a complete message to speaker B, B interprets the message, and then B formulates a complete message to speaker A. Instead, both speakers jointly help each other during the process of constructing messages. Meanings accumulate collaboratively and incrementally, with ongoing repair. Speaker A fills in fragments of what speaker B is trying to say, and vice versa. The listener assists the speaker by filling in words and by providing *backchannel feedback* that acknowledges that the listener is following what the speaker is saying (e.g. 'uh huh', head nod). The listener does this while the speaker is speaking.

Perhaps the most salient characteristic of normal tutoring is the prevalence of collaborative communication while the tutor and student work on problems, discuss subtopics, and answer questions. This was quite obvious when we analysed the two most frequent activities in our tutoring protocols: working on example problems and answering questions asked by tutors. When the tutor and student worked on an example problem, there were dozens/hundreds of turns throughout the collaborative evolution of the solution. When the tutor asked a question, the median number of turns in the answer was 5 in the research methods corpus and 10 in the algebra corpus. Only one turn would be needed if it were simply the case that the tutor asked the question and the student answered it.

A pervasive dialogue pattern consisted of a 5-step dialogue frame that was initiated by a tutor question (Graesser, 1993a,b; Graesser and Person, 1994; Person *et al.*, in press).

Step 1: Tutor asks question.
Step 2: Student answers question.
Step 3: Tutor gives short feedback on the quality of the answer.
Step 4: Tutor and student collaboratively improve the quality of the answer.
Step 5: Tutor assesses student's understanding of the answer.

Figure 1 specifies further the components of this 5-step dialogue frame. A short example of the frame is provided below.

1. *Tutor:* Now what is a factorial design?
2. *Student:* The design has two variables.
3. *Tutor:* Uh-huh.
4. *Tutor:* So there are two or more independent variables and one [pause]
   Student: dependent variable.
5. *Tutor:* Do you see that?
   Student: Uh-huh.

It should be noted that teachers in classrooms normally enact a 3-step dialogue frame instead of the 5-step dialogue frame. Mehan (1979) identified a persistent dialogue pattern in classrooms, which includes initiation, response, and evaluation. The teacher elicits information from the student (frequently in the form of a question), the student responds, and then the teacher evaluates the response. This classroom frame corresponds to the first three steps of our 5-step dialogue frame in tutoring. It is conceivable that the extra two steps account for the advantage of tutoring over classroom settings.

Step 1: Tutor asks questions

In step 1, the tutor normally asks a single question. However, sometimes the question is not posed clearly or as intended, so the tutor revises the question. Successive tutor questions drifted systematically in a manner that made it easier for the student to answer the question (Graesser, 1992). For example, in the excerpt

STEP 1: TUTOR ASKS QUESTION

IF the tutor cannot understand the question or the question is not posed as intended, THEN the tutor asks a revised question.
IF the student does not understand the question, THEN the student asks a counter-clarification question.

↓

STEP 2: STUDENT ANSWERS QUESTION

↓

STEP 3: TUTOR GIVES SHORT FEEDBACK

The tutor's feedback is positive, negative, or neutral.
The feedback is linguistic or paralinguistic (e.g., head nod).
Intonation is important.

↓

STEP 4: TUTOR IMPROVES QUALITY OF ANSWER

The tutor splices in a complete or partial answer.
The tutor summarizes answer.
The tutor gives hint.
The tutor pumps the student for more information.
The tutor traces explanation or justification.
The tutor elaborates on answer.
The tutor asks question to elaborate on answer.
The tutor presents an example.
The tutor corrects a misconception.
The tutor issues a command or indirect request for student to complete an activity.

↓

STEP 5: TUTOR ASSESSES STUDENT'S UNDERSTANDING

The tutor asks whether the student understands.
The tutor asks a simple question.
The tutor asks a complex question.
The tutor requests the student to solve a similiar problem.

Figure 1. The 5-step dialogue frame.

below, an answer to the first question would invite an elaborate construction of information, whereas a simple 'yes' or 'no' would be an adequate answer to the second question.

Tutor: So how could we do that [operationally define intelligence]? I mean, do you think that everyone agrees on what intelligence is?

In the following example, the tutor restates the question in different words that provide a more succinct focus on the intended question. It illustrates that the process of constructing a question is iteratively distributed over time.

Tutor: Did you see how they did that? How did they manage to do that? What did they do there?

Question revisions such as these are presumably more frequent when the student appears puzzled or unengaged.

Sometimes the student did not understand the question, particularly when the question was not adequately specified. In these situations, the student sometimes asked a counter-clarification question and then the student answered the original question, as illustrated below.

Tutor: Why would a researcher ever want to use more than two levels of an independent variable in an experiment?
Student: More than two levels?
Tutor: Uh-huh.
Student: They would, um, it'd be real accurate 'cause it would show if there's a curvilinear.

The tutors frequently did not pose questions with a sufficiently high level of specification. They elliptically deleted words, phrases, and clauses from their questions under the assumption that the context is sufficiently rich for the student to reconstruct the intended question. However, the tutors were frequently incorrect in making this assumption. As a consequence, the student ended up misinterpreting the question or the student answered the wrong question. Tutor questions were scaled on degree of specification, with values of high, medium, and low (Graesser and Person, 1994). Only 2% of the questions had high specification and 50% had low specification. The students sometimes did not have enough context to interpret the question so they asked counter-clarification questions. The likelihood of a student asking a counter-clarification question significantly decreased as a function of higher specification of the tutor question, 0.17, 0.08, and 0.00 for tutor questions that were low, medium, and high in specification. Tutors should probably try to formulate their questions with a higher degree of specification.

Step 2: Student answers question

The student's answer within this single turn was typically piecemeal, incremental, and semi-coherent, as the above example illustrates. Answers are not normally articulated in a clear, succinct, coherent form. The student frequently produced single words, short phases, or incoherent fragments of information. The tutor ended up working with these fragments (in step 4) in a fashion that allowed a reasonable answer to evolve. When a student's initial answer was incomplete, the tutor frequently pumped the student for additional information by expressing neutral feedback in step 3 (e.g. 'uh huh'). There was an iteration of steps 2 and 3 when the tutor pumped the students for more answer information. Tutors and intelligent tutoring systems should use this pump mechanism more extensively if they want the student to contribute more information in the dialogue and thereby be more active.

Step 3: Tutor gives short feedback on the quality of the student's answer

In our analyses, we classified the short feedback as positive, negative, or neutral. Most of the time the feedback was expressed verbally. Occasionally, the tutor nodded or shook his head to express feedback. When the feedback was neutral on the written transcript, it was necessary to view the videotape in order to code the intonation and timing of the feedback. Intonation and timing is an important consideration when classifying feedback as positive, negative, or neutral (Fox, 1991, 1993). We found that 34% of the neutral feedback ended up being either positive or negative when the videotape was viewed.

Fox (1991, 1993) reported that skilled tutors sometimes use lengthy pauses or hesitations to signify negative feedback. Instead of verbally giving negative feedback (e.g. 'no', 'that's not correct'), the tutors in her sample used hesitations, pauses, and intonation contours to express negative feedback. However, the unskilled tutors in our sample did not use these cues to signify negative feedback. The likelihood of the tutor pausing or hesitating in step 3 did not vary (as Fox would predict) as a function of the quality of the student's answer in step 2; the mean likelihoods were 0.08, 0.13, 0.15, and 0.13 when the students' answers were error-ridden, vague, partially correct, and completely correct, respectively. The use of non-verbal negative feedback apparently is a consequence of the training, experience, and conversational style of the tutor.

Step 4: Tutor and student collaboratively improve the quality of the answer

Figure 1 lists some of the strategies that the tutors in our sample used to improve the quality of the answer. We believe that this is a critical step that differentiates good tutors from mediocre tutors, and tutors from teachers in classrooms. It is beyond the scope of this article to enumerate all of the strategies and tactics that tutors can use to enhance collaborative activities. We will point out some of the frequent ones that our unskilled tutors used and will suggest ways that they may be implemented in computerized tutorial dialogue.

Pumping

We have already discussed this method. The tutor pumps the student to elaborate an answer in step 2 by giving positive or neutral feedback in step 3 (e.g. 'uh huh', 'okay', head nod) and then pausing for the student to supply more information. This pumping mechanism was frequently used when the student's answer was vague in step 2. Tutors were reluctant to give a completely correct answer in step 4 when the quality of the student's contribution in step 2 had a vague status (Graesser, 1993b).

There was a 0.08 likelihood of a *tutor* giving a completely correct answer at the beginning of step 4, given that the student's answer was vague in step 2. In contrast, there was a 0.26 likelihood of a *student* giving a completely correct answer after being pumped, given that the student's answer was vague in step 2. Therefore, there was some pedagogical payoff in using the pump mechanism: the students had more knowledge than was manifested in their initial step 2 answer and this knowledge was elicited by the pump mechanism. The tutors were reluctant to rush in with a complete answer at the beginning of the answer evolution. The following production rule captures this pumping method.

IF [quality of student answer in step 2 = vague (or nothing)]
THEN [tutor pumps student for more information]

An intelligent tutoring system could implement this pump mechanism, according to this production rule, by producing a neutral feedback (e.g. 'okay') or by a directive response (e.g. 'tell me more').

Prompting students to fill in words
The tutors frequently prompted the student to fill in a word or phrase in the middle of a tutor's contribution in step 4. They did this by pausing or by intonation.

Tutor: There are two independent variables and one———[pause]
Student: dependent variable.

The tutor already knew the information that needed to be filled in. The tutor encouraged the student to contribute more actively by prompting the student to supply the missing information. This can be captured in the following production rule:

IF [tutor's contribution has missing information x and the tutor knows x and the tutor believes the student knows x]
THEN [tutor prompts the student to fill in x]

An intelligent tutoring system could be designed to implement this production rule. The system would periodically stop, print out an underline prompt, and request the student to fill in the information. After the student supplies this information, the computer would print out the correct information and thereby provide feedback. The computer system could dynamically adjust the type and amount of information the student is supposed to fill in, depending on the student's ability. At the beginning of the session, the computer would prompt the student to fill in easy information or single words. As the session progresses and the student's knowledge increases, the student would be expected to fill in more difficult information and lengthier responses. This method would gradually encourage the students to take a more active role in the learning experience.

Splicing
When a student's contribution was error-ridden, the tutors tended to jump in quickly and splice in a good answer. Thus, immediate corrective feedback was given by the tutor, in the spirit of Anderson's LISP tutor (Anderson, Boyle and Reiser, 1985;

Anderson *et al.*, 1989). We use the metaphor of 'splicing' because normally the tutor and student are jointly constructing a connected structure of ideas when the errors occur. The splicing method was apparent when we performed detailed analyses of the evolution of answers to 'deep' questions asked by the tutor (e.g. why, how, what-if, as will be defined later). The likelihood of a tutor giving a partially correct or completely correct answer on contribution $n+1$ varied as a function of the quality of the student's contribution n; the likelihoods were 0.59, 0.62, 0.58, and 0.81 for the quality states of completely correct, partially correct, vague, and error-ridden at contribution n (Graesser, 1993b). The 0.81 value for error-ridden answers was significantly higher than the other values. The tutors frequently spliced in correct information without informing the student that the student's contribution was error-ridden. The production rule below captures this strategy.

IF [the student's contribution is error-ridden]
THEN [tutor splices in an answer that is partially or completely correct]

Summarizing

The tutors carried the burden of summarizing or recapping the answer to a deep question asked by the tutor. This was apparent when we analysed the evolution of answers to these questions. Tutors had a higher likelihood than students of giving a completely correct contribution (at $n+1$) when the cumulative exchange had already reached the quality state of completely correct, 0.16 versus 0.04 (Graesser, 1993b). The following production rule captures this strategy.

IF [quality of the cumulative collaborative exchange = completely correct]
THEN [tutor supplies a summary or recap of the answer]

It would be preferable for the student to take on the burden of providing these summaries because such activities improve organization and retention. Skilled tutors and intelligent tutoring systems should request that the students supply answer summaries.

Pumping, prompting, splicing, and summarizing hardly exhaust the strategies that the tutors used in our two corpora. But they were the most frequent forms of scaffolding in step 4. Other strategies were occasionally applied in step 4 of the 5-step dialogue frame (see Figure 1). For example, various types of hints were provided when students encountered obstacles and produced errors. However, we found that our tutors' hints were less than impressive from the standpoint of existing intelligent tutoring systems and ideal theories of tutoring (Fox, 1991, 1993; Lepper *et al.*, 1990; Lepper *et al.*, in press; Merrill *et al.*, 1992; Stevens *et al.*, 1982).

Step 5: Assessing the student's understanding of the answer

In most cases (92% of the time) the tutor simply asked the student whether the student understood the answer (e.g 'Do you understand'?, 'Do you follow?', 'Okay?'). Unfortunately, student answers to these comprehension-gauging questions are notoriously inaccurate, as will be discussed later. The tutors occasionally asked a simple follow-up question that tested the student's understanding of the answer (7% of the time). Very rarely did the tutor thoroughly test the student's understanding by asking a complex question or by requiring the student to solve a problem, as illustrated below.

> *Tutor:* Do you have any problem with these kinds of word problems (referring to a section in a book). Where they say . . .
> *Student:* (interrupts) Uh, not really.
> *Tutor:* You don't. You don't? You don't have any problem with that?
> *Student:* No.
> *Tutor:* Let's just do one of them. Um, Dan earned 56 dollars, which is twice more than what Jim earns. Now you're supposed to write an equation.
> *Student:* Uh, I can't write the equations.

Tutors apparently assume that students understand anything that gets discussed during tutoring. If something gets said, tutors assume that it is understood and they quickly verify this understanding. But this is a bad assumption. A good tutor should assess the student's understanding more rigorously. The tutor could ask follow-up questions that are diagnostically discriminating and that troubleshoot potential misunderstandings. The tutor could present a similar problem and ask the student to solve it without any coaching.

DEEP EXPLANATORY REASONING

A deep understanding of domain knowledge is achieved when students have mastered appropriate conceptual models for making predictions, solving problems, and generating explanations (Cobb, Wood, Yackel and McNeal, 1992; diSessa, 1993; Gentner and Stevens, 1983; Kozma, 1992; Mayer, 1992; Ohlsson and Rees, 1991; Stevens *et al.*, 1982; White, 1993). Perhaps tutoring permits students to penetrate deeper levels of reasoning and understanding.

Our analyses of normal tutoring did not directly evaluate the depth of the acquired conceptual models and the depth of the student's reasoning. However, we did indirectly analyse the depth of the questions that students and tutors asked during the tutoring sessions. Deep questions were defined as those questions that manifested logical reasoning, causal reasoning, or goal-oriented reasoning. In logical reasoning, the statements expressed in the answer consist of the premises and conclusions of a logical syllogism. In causal reasoning, the answer conveys the antecedents and consequences of events. In goal-oriented reasoning, the answer traces the goals and plans of agents. It is well documented that comprehension and memory for technical material improves to the extent that the learner actively constructs explanations and justifications that reflect these types of reasoning (Chi, Bassok, Lewis, Reimann and Glaser, 1989; Chi, Leeuw, Chiu and LaVancher, in press; Cobb *et al.*, 1992; Pressley, Symons, McDaniel, Snyder and Turner, 1988; Van Lehn, Jones and Chi, 1992).

We developed a taxonomy that identifies those question categories that expose deep reasoning (Graesser and Person, 1994; Graesser *et al.*, 1992). Deep reasoning questions include the following six categories.

1. *Antecendent questions* (why? how?). What caused a state or event? What logically explains or justifies a proposition?
2. *Consequence questions* (what if? what next?). What are the causal consequences of a state or event? What are the logical consequences of a proposition?
3. *Goal orientation* (why?). What are the goals or motives behind an agent's actions?
4. *Enablement* (why? how?). What object, state, or resource allows an agent to perform an action? What state or event allows another state or event to occur?
5. *Instrumental-procedural* (how?). What instrument or plan allows an agent to accomplish a goal?
6. *Expectational* (why not?). Why did an expected state or event not occur. Why didn't an agent do something?

These questions are manifested in a tutoring session to the extent that the tutor and student explore deeper levels of comprehension. It should be noted that these deep reasoning questions were highly correlated ($r = 0.64$) with the deeper levels of cognition in Bloom's taxonomy of educational objectives in the cognitive domain (Bloom, 1956). Low level questions in Bloom's taxonomy inquire about specific facts, terminology, and explicit information in the text; deeper level questions involve reasoning, analysis, application, synthesis, and evaluation (see also Scardamalia and Bereiter, 1991).

Our analyses of normal tutoring uncovered an impressive number of deep reasoning questions (Graesser and Person, 1994). In a typical tutoring session, a student asked approximately eight deep reasoning questions per hour, whereas a tutor asked 19 deep reasoning questions. The incidence of deep reasoning questions was much higher in tutoring sessions than in normal classroom settings, according to published studies on classroom questioning (Dillon, 1988; Graesser and Person, 1994). The incidence of student questions in a classroom is extremely low in published studies (0.11 question per student per hour), so deep reasoning questions would also be very low. Only 4% of teacher questions in a classroom are deep questions in Bloom's taxonomy; the vast majority of teacher questions are short-answer questions that grill students on explicit material (Dillon, 1988; Kerry, 1987).

We found that the good students in the research methods corpus asked a higher proportion of deep reasoning questions (Person *et al.*, 1994). There were significant positive correlations between the proportion of student questions that were deep reasoning questions and: (1) examination scores ($r = 0.44$) and (2) final grades ($r = 0.58$). Therefore, good students penetrated the deeper levels of comprehension.

In summary, our analyses of normal tutoring indicate that deep reasoning is more prevalent in tutoring than classroom settings. The reasoning is a collaborative activity between student and tutor, and typically occurs while solving example problems. In traditional classroom settings, there is greater emphasis on the delivery of didactic content by the teacher and the posing of shallow questions about explicit material. We believe these differences may explain the well-documented learning advantages of tutoring over classroom settings.

CONVERGENCE TOWARD SHARED MEANINGS

Perhaps another advantage of tutoring is an enhanced 'meeting of the minds' between tutor and student. Roschelle (1992) has argued that there is a convergence toward shared meanings in collaborative learning. It would be impractical to achieve such common ground in a classroom of 30 students. However, model tracing facilities in intelligent tutoring systems attempt to track the knowledge of particular students and to shape the interaction around the student model (Anderson *et al.*, 1989; Clancey, 1987; Larkin and Chabay, 1992; Ohlsson, 1986; Sleeman and Brown, 1982). The establishment of common ground is very critical in these systems. Similarly, researchers in discourse processing have argued that the establishment of common ground is a prerequisite for successful communication (Clark and Schaefer, 1989).

It is conceivable, nevertheless, that normal tutoring does not achieve a close and rapid convergence to shared meanings. Just as the goals are substantially different for teachers and students in a classroom (Lave, Smith and Butler, 1988; Schoenfeld, 1988), there may be a serious mismatch in the goals of tutors and students. Moreover, analyses of our tutoring protocols present a somewhat pessimistic picture with respect to establishing common ground. There appeared to be a very slow refinement of common ground. At times we have been tempted to go out on a limb and exaggerate the picture: the tutor and student seem to be operating in frighteningly different mental spaces; the collaborative activities mutually constrain each other's mental spaces in a slow and sporadic fashion.

Consider first the question of how the tutor infers what the student knows. One method at the tutor's disposal is to simply ask the student. This occurs naturally in step 5 of the 5-step dialogue frame, when the tutor asks the student a comprehension-gauging question (e.g. 'Do you understand?', 'Okay'). When these questions are asked, the student either answers 'yes' ('I understand'), answers 'no' (I don't understand'), or gives an indecisive response (no answer, 'I don't know'). We found that students' answers were not a valid reflection of the student's true understanding. There was nearly a zero correlation between student achievement and the likelihood of answering 'yes'. In fact, this relationship was found to be significantly curvilinear, 0.46, 0.62, 0.61, and 0.52 for students receiving final grades of A, B, C, and C – /D, respectively (Person *et al.*, 1994). Regarding the 'no' answers, there was a significant positive correlation between exam scores and the likelihood of students' answering 'no' (I don't understand). This is a counterintuitive outcome. It was the good students who said that they did not understand. Chi *et al.* (1989) also reported a positive correlation, in the domain of physics, between student understanding and the likelihood of their answering 'no'. Therefore, available evidence indicates that a tutor cannot simply ask students whether they understand and expect students to supply accurate feedback. The feedback is misleading. Part of the problem is that students are notoriously poor at calibrating their own comprehension of material (Glenberg, Wilkinson and Epstein, 1982; Weaver, 1990).

The tutors in our sample also could not accurately infer student knowledge from the questions that students asked. As reported earlier, there was no correlation between student achievement and: (1) the frequency of questions students asked; and (2) the proportion of student questions that were knowledge-deficit questions. There was a positive correlation between student achievement and the proportion of student questions that were deep reasoning questions. However, it would be difficult for the tutor to gauge student understanding by this index. An average student asked only eight deep reasoning questions per hour, so the tutor would be basing the computation on a low frequency event.

The students' answers to questions provided the most reliable basis for inferring student knowledge (Person *et al.*, 1994). There was a modest positive correlation between achievement and the proportion of student answers that were completely correct ($r = 0.32$, and 0.43 for exam scores and final grades, respectively). There was a negative correlation between student achievement and the proportion of student answers that were categorized as error-ridden, vague, or no-answer ($r = -0.52$, and -0.49). The tutors asked a large number of questions (104 questions per hour), so there was ample opportunity for the students to give answers and for the tutor to evaluate the quality of the answers. Therefore, it is the tutor's challenge to select those questions that diagnose the student's knowledge deficits, bugs, and deep misconceptions. Perhaps this is why tutor questions are so frequent in tutoring sessions.

Although it is possible for tutors to infer student knowledge from student answers, it is important to acknowledge that student contributions are normally very fragmentary, semi-coherent, and distributed over many turns during the collaborative exchange. In fact, the tutors ended up supplying more information in route to an answer than did the students, even though the tutors asked the original questions (Graesser, 1993b). This is hardly a solid foundation for the tutor to infer student knowledge.

There are other reasons for being sceptical about the notion that there is a rapid convergence toward shared meanings during tutoring. First, the tutors' contributions tend to be fragmentary, semi-coherent, and distributed over many turns. So it may be difficult for students to get into the mind set of the tutor. Second, the students already had been exposed to the material in lectures and textbooks prior to the tutoring, yet they still had problems comprehending the material. Given they frequently failed to comprehend the material delivered in coherent discourse, it is unlikely that they would pick it up during a scruffy tutoring exchange. Third, the students received inaccurate feedback from the tutor about whether the student was understanding the material. This claim will be elaborated in the next section. Fourth, as discussed earlier, the questions posed by the tutors tended to be underspecified and were frequently unclear to the students (as manifested by counter-clarification questions). In summary, available evidence is incompatible with the intuitively compelling idea that there is a robust 'meeting of the minds' in normal tutoring sessions.

FEEDBACK, ERROR DIAGNOSIS, AND ERROR REMEDIATION

An effective tutor presumably should give the student feedback on the quality of the student's contributions. When the student does well, the student should be given positive feedback and praise. When the student does poorly, the tutor should give negative feedback or withhold positive feedback. When the student commits errors, the tutor should acknowledge the error and correct it. This simple rational view of feedback would appear to be a sensible component for skilled tutors and intelligent tutoring systems. However, the matter of feedback is far more complex.

Consider the treatment of error-ridden contributions of the student. The tutor may handle these errors by: (1) acknowledging the error occurred; (2) identifying where the error occurred; (3) instructing the student how to repair the error; (4) diagnosing the bugs and misconceptions that generated the error; and (5) setting new goals that remediate the error, bugs, and misconceptions. This approach to handling errors has been incorporated in a number of intelligent tutoring systems (Anderson *et al.*, 1985, 1989; Burton and Brown, 1982; Lesgold, Lajoie, Bunzo and Eggan, 1992; Reiser, Connelly, Ranney and Ritter, 1992; Van Lehn, 1990). This straightforward feedback on student errors may prevent the student from floundering and becoming discouraged.

There are two major drawbacks to the above treatment of errors. The first drawback is that students don't learn how to discover their own errors, to repair them, and to develop the metacognitive skill of self-regulating their knowledge (Bangert-Drowns, Kulik, Kulik and Morgan, 1991; Collins and Brown, 1988; Merrill *et al.*, 1992; Scardamalia and Bereiter, 1991; Scardamalia, Bereiter, McLean, Swallow and Woodruff, 1989; Schoenfeld, 1988). The second drawback is that students may become discouraged when they are confronted with a large amount of negative feedback and error remediation (Lepper *et al.*, 1990, in press; Merrill *et al.*, 1992). Thus, there is a tradeoff between the cognitive goal of imparting correct knowledge and the affective goal of building student confidence. An alternative approach would be to give gentle and indirect guidance when students manifest errors, bugs, and misconceptions, rather than direct feedback and remediation. Indeed, there is some evidence that skilled tutors adopt the gentle indirect approach rather than the harsh direct approach (Fox, 1991, 1993; Lepper *et al.*, in press).

So how do normal tutors handle the problem of feedback and student errors? We found that our tutors did not spend much time diagnosing, dissecting, and troubleshooting the student errors that occurred in the dialogue. Only 3% of the tutors' contributions were devoted to rectifying the bugs and deep misconceptions that explain errors (Graesser, 1993b). Only 5% of tutor questions were driven by student errors. It is very difficult for tutors to identify these underlying bugs and misconceptions, let alone to repair them. In the last section we presented evidence that there is no fine-tuned convergence toward shared meanings, so it is not surprising that tutors failed to pursue this approach. Intelligent tutoring systems that are designed to recognize and repair bugs and misconceptions may therefore reap benefits that are not supplied by normal tutors. Even skilled tutors rarely provide this sophisticated feedback (Lepper *et al.*, 1990, in press; McArthur *et al.*, 1990; Putnam, 1987).

The tutors in our study applied a variety of strategies when students gave error-ridden contributions. Graesser (1993b) analysed the tutors' responses to error-ridden student contributions. The tutors directly acknowledged student errors only 24% of the time when we examined their short feedback (step 3 of the dialogue frame) and their long feedback (step 4). Consequently, some students might not get the message that they are committing an error. However, the tutors did not gloss over the errors. The tutors spliced in correct answers 38% of the time, as discussed earlier. The tutors initiated lengthy explanations and problem solutions 25% of the time. The tutors provided hints 28% of the time; 19% of the time they asked the student a question in order to direct the exchange toward a correct answer. Very rarely (2% of the time) did the students catch their own errors. These results support the position that tutors adopt a gentle, indirect stance when students commit errors. Normal tutors refrain from telling the student they are committing an error, but the tutors do invoke an indirect strategy to correct the error. This is illustrated in the example below.

Tutor: . . . and that's our frequency distribution . . . What is that one called again? (pointing to a bar graph)
Student: A histogram.
Tutor: Allright, or a bar graph.

Although the tutor spliced in the correct answer, it is not clear whether the student understood his answer was in error. Technically speaking, a histogram is not equivalent to a frequency distribution.

Graesser (1993b) analysed the feedback that tutors gave to student contributions at all levels of quality: error-ridden, vague, partially correct, and completely correct. Short feedback consisted of brief positive, negative, or neutral response in step 3 of the dialogue frame (e.g. 'yeah', 'right', 'good', 'okay', 'uh huh', 'not so', head nod, pause, sceptical intonation pattern). Long feedback consisted of lengthier comments on answer quality during step 4 (e.g. 'that is correct because . . .', 'there is a problem with prediction . . .'). Corrective feedback is a more complex form of negative feedback; the tutor contributes information in step 4 that corrects erroneous or misleading information in the student's contribution (but does not tell the student an error was made). The results of this analysis are presented in Table 1.

From one perspective, it would appear that the tutors gave sensible and discriminating feedback to the students regarding the quality of their contributions. Positive feedback robustly increased as a function of answer quality, whereas negative feedback robustly decreased. From another perspective, however, the feedback was inaccurate for error-ridden and vague student contributions. Tutors tended to give positive feedback more often than negative feedback when student contributions were error-ridden or vague. Tutors did give corrective feedback to error-ridden contributions, but they did not acknowledge the contribution was error-ridden. Once again, tutors gave gentle, indirect feedback to error-ridden answers rather than harsh, direct negative feedback. These results are compatible with available research on skilled tutors (Lepper *et al.*, 1990, in press; McArthur *et al.*, 1990).

AFFECT AND MOTIVATION

Lepper has argued that a tutor should consider affective and motivational goals of the learning experience in addition to cognitive goals (Lepper *et al.*, 1990; in press). The affective and motivational goals include: (1) enhancing the student's self-efficacy and confidence; (2) challenging the student; (3) giving the student a sense of control; and (4) eliciting curiosity. Lepper identified approximately 20 tactics that a tutor could implement to achieve these goals. One tactic, for example, is to give the student an easy problem that the student should be able to solve, but to mention that it is a challenging problem that might be beyond the student's abilities. If the student solves it, the student gains a sense of self-confidence and self-efficacy. If the student cannot solve it, the failure can be attributed to the problem rather than the student. Another tactic is to select real-world problems that arouse curiosity and interest.

The tutors in this study rarely implemented most of the tactics identified by Lepper. Apparently, training and experience are needed before these skills emerge in the tutor's repertoire. The only tactics proposed by Lepper that occurred among the tutors in this study have already been discussed: tutors avoid giving negative feedback, they give indirect feedback to poor student contributions, and they give positive feedback to good student contributions.

One of the salient features of our tutorial dialogues was that the tutors were polite conversation partners (Person *et al.*, in press). They refrained from confronting, embarrassing, and humiliating the student. Researchers in discourse processing and sociolinguistics have proposed that a 'politeness principle' is an overarching conversational norm of conversation (Brown and Levinson, 1987; Leech, 1983). A polite conversationalist has tact, has modesty, avoids disagreement, minimizes impositions on the listener, and so on. These politeness goals are sometimes incompatible with cognitive pedagogical goals (Person, *et al.*, in press). When a cognitive goal calls for the tutor to make requests, demands, and criticisms of the student, then there is a potential clash with politeness goals. The tutor compromises by substantially softening the requests, demands, and criticisms. So the tutor issues indirect requests (e.g. Why don't we work on problem 4?) instead of direct requests (e.g. 'Solve problem 4'). The tutor uses indirect feedback rather than direct feedback and criticisms when student contributions are error-ridden.

It is interesting to speculate what tutors would be like if they violated politeness maxims and the Gricean maxims of conversation (e.g. quality, quantity, manner, relation; Grice, 1975). Violating some of these maxims would produce a tutor who is abrasive, rude, boring, pedantic, or incoherent. But there may be specific conditions in which it is advantageous, pedagogically or stylistically, to violate specific maxims. Under what conditions might it be worthwhile to crack the polite conversation barrier in service of pedagogical goals? One direction for future research is to identify potential context-sensitive rules that violate specific conversational maxims, but enhance pedagogical or stylistic goals. A tutor who violates some of these maxims, under specific conditions, may end up having a more interesting personality and may thereby increase student motivation.

Table 1. Tutor feedback as a function of the quality of the student contributions

	Quality of student contribution			
	Error-ridden	Vague	Partially correct	Completely correct
Positive feedback				
Short	0.31	0.32	0.56	0.68
Long or short	0.31	0.38	0.62	0.78
Negative feedback				
Short	0.23	0.08	0.06	0.00
Long or short	0.24	0.10	0.07	0.00
Corrective, long, or short	0.62	0.18	0.12	0.04
Neutral feedback	0.21	0.52	0.28	0.21

CLOSING COMMENTS

Our exploration of normal tutors has identified those components of learning that are prevalent in tutorial dialogue. The tutor and student engage in distributed collaborate dialogue while solving example problems. They jointly construct deeper explanations than in traditional classroom settings. There is more opportunity for students to ask questions. The tutor controls the dialogue in the form of a curriculum script and a 5-step dialogue frame. Tutors are polite conversational partners; they issue directives in the form of indirect requests and they indirectly correct student errors. A number of learning components were conspicuously absent, deficient, or underdeveloped: active student learning, sophisticated pedagogical strategies, cases anchored in real-world scenarios, rapid convergence to shared meanings, error diagnosis, discriminating feedback on poor student contributions, and tactics to bolster affect and motivation. Thus, our analyses have narrowed down the space of potential learning mechanisms that exist during normal tutoring.

This project has sharpened some of the boundaries that discriminate normal unskilled tutoring, skilled tutoring, and intelligent tutoring systems. For a number of reasons, it is important to identify these boundaries. One reason is that the vast majority of tutors in actual school systems are unskilled tutors like the 13 tutors in our samples. Given that unskilled tutors are the norm, it is important to understand how they tutor and how the resulting learning experiences influence learning outcomes. It is already well documented that unskilled tutors produce better learning outcomes than traditional classroom activities. What has been missing is an understanding of the process of unskilled tutoring. A second reason for sharpening the boundaries addresses the contrast between skilled and unskilled tutors. Skilled tutors presumably are better than unskilled tutors. We can now identify what skilled tutors do, but unskilled tutors fail to do. This is needed in order to train effective tutors; we know what learning components and tutoring strategies are natural for tutors to implement versus those that require training. A third reason for sharpening the boundaries addresses the contrast between unskilled tutors and intelligent tutoring systems. We have a better understanding of how intelligent tutoring systems provide more sophisticated pedagogical strategies than unskilled one-to-one tutoring. We also understand what it would take for the dialogue facility on an intelligent tutoring system to mimic naturalistic tutoring between humans.

REFERENCES

Anderson, J. R., Boyle, C. F. and Reiser, B. (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**, 467–505.

Bakhtin, M. (1981). *The dialogic imagination*. Austin: University of Texas Press.

Bangert-Drowns, R. L., Kulik, C-L., Kulik, J. A. and Morgan, M. T. (1991). The instructional feedback in test-like events. *Review of Education Research*, **61**, 213–238.

Bloom, B. S. (Ed.). (1956). Taxonomy of educational objectives. *Handbook I: Cognitive domain*. New York: McKay.

Bloom, B. S. (1984). The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, **13**, 4–16.

Bransford, J. D., Goldman, S. R. and Vye, N. J. (1991). Making a difference in people's ability to think: reflections on a decade of work and some hopes for the future. In R. J. Sternberg and L. Okagaki (Eds), *Influences on children*. Hillsdale NJ: Lawrence Erlbaum Associates, Inc.

Brown, A. L. (1988). Motivation to learn and understand: on taking charge of one's own learning. *Cognition and Instruction*, **5**, 311–321.

Brown, A. L. (1992). Designing experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, **2**, 141–178.

Brown A. L. and Campione, J. (1990). Communities of learning and thinking, or A context by any other name. *Human Development*, **21**, 108–126.

Brown A. L. and Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: essays in honor of Robert Glaser* (pp. 393–451). Hillsdale NJ: Lawrence Erlbaum Associates, Inc.

Brown, P. and Levinson, S. C. (1987). *Politeness: some universals in language use*. Cambridge: Cambridge University Press.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, **31**, 21–32.

Burton, R. R. and Brown, J. S. (1982). An investigation of computer coaching for informal learning activities. In D. Sleeman and J. S. Brown (Eds), *Intelligent tutoring systems* (pp. 79–98). New York: Academic Press.

Carroll, J. S. (1990). *The Nurnberg funnel: designing minimalist instruction for practical computer skill*. Cambridge, MA: MIT Press.

Chi, M. T., Bassok, M., Lewis, M. W., 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., Leeuw, N., Chiu, M. H. and LaVancher, C. (in press). Eliciting self-explanations improves understanding. *Cognitive Science*.

Clancey, W. (1987). *Knowledge-based tutoring*. Cambridge, MA: MIT Press.

Clark, H. H. and Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, **13**, 259–294.

Cobb, P., Wood, T., Yackel, E. and McNeal, B. (1992). Characteristics of classroom mathematics traditions: an interactional analysis. *American Educational Research Journal*, **29**, 573–604.

Cohen, P. A., Kulik, J. A and Kulik C. C. (1982). Educational outcomes of tutoring: a meta-analysis of findings. *American Educational Research Journal*, **19**, 237–248.

Collins, A. (1985). Teaching reasoning skills. In S. F. Chipman, J. W. Segal and R. Glaser (Eds), *Thinking and learning skills*, vol 2 (pp. 579–586). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Collins, A. (1988). Different goals of inquiry teaching. *Questioning Exchange*, **2**, 39–45.

Collins, A. and Brown, J. S. (1988). The computer as a tool for learning through reflection. In H. Mandl and A. Lesgold (Eds), *Learning for intelligent tutoring systems* (pp. 1–18). New York: Springer-Verlag.

Collins, A. Brown, J. S. and Newman, S. E. (1989). Cognitive apprenticeship: teaching the craft of reading, writing and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: essays in honor of Robert Glaser* (pp. 453–494). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Dansereau, D. F. (1988). Cooperative learning strategies. In C. L. Weinstein, E. T. Goetz and P. A. Alexander (Eds), *Learning and study strategies: issues in assessment, instruction, and evaluation*. San Diego, CA: Academic Press.

Dillon, J. T. (1988). *Questioning and teaching: a manual of practice*. New York: Teachers College Press.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, **10**, 105–225.

Fantuzzo, J. W., King, J. A. and Heller, L. R. (1992). Effects of reciprocal peer tutoring on mathematics and school adjustment: a component analysis. *Journal of Educational Psychology*, **84**, 331–339.

Fishbein, H. D., Eckart, T., Lauver, E., Van Leeuwen, R. and Langmeyer, D. (1990). Learners' questions and comprehension in a tutoring setting. *Journal of Educational Psychology*, **82**, 163–170.

Fitz-Gibbon, C. T. (1977). *An analysis of the literature of cross-age tutoring*. Washington, DC: National Institute of Education (ERIC Document Reproduction Service No. ED 148 807).

Fox, B. (1991). Cognitive and interactional aspects of correction in tutoring. In P. Goodyear (Ed.), *Teaching knowledge and intelligent tutoring* (pp. 149–172). Norwood, NJ: Ablex.

Fox, B. (1993). *The human tutorial dialogue project*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gee, J. P. (1989). *What is literary*? (Tech. Rep. No. 2). Newton, MA: Literacy Institute.

Gentner, D. and Stevens, A. (Eds). (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Goldman, S. R., Pellegrino, J. W. and Bransford, J. D. (1993). Assessing programs that invite thinking. In H. O'Neill and E. Baker (Eds), *Technology assessment: estimating the future*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Goodwin, C. and Heritage, J. (1990). Conversational analysis. *Annual Review of Anthropology*, **19**, 283–307.

Graesser, A. C. (1992). *Questioning mechanisms during complex learning*. Memphis State University, Memphis, TGN (ERIC Document Reproduction Service ED No. 350 306).

Graesser, A. C. (1993a). Dialogue patterns and feedback mechanisms during naturalistic tutoring. *Proceedings of the 15th Annual Cognitive Science Society* (pp. 126–130). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Graesser A. C. (1993b). *Questioning mechanisms during tutoring, conversation, and human–computer interaction*. Memphis State University, Memphis, TN (ERIC Document Reproduction Service No. TM 020 505).

Graesser, A. C., Langston, M. C. and Baggett, W. B. (1993). Exploring information about concepts by asking questions. In G. V. Nakamura, R. M. Taraban and D. Medin (Eds), *Categorization by humans and machines* (pp. 411–436). Orlando, FL: Academic Press.

Graesser, A. C. and McMahen, C. L. (1993). Anomalous information triggers questions when adults solve quantitative problems and comprehend stories. *Journal of Educational Psychology*, **85**, 136–151.

Graesser, A. C. and Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, **31**, 104–137.

Graesser, A. C., Person, N. K. and Huber, J. (1992). Mechanisms that generate questions. In T. Lauer, E. Peacock and A. Graesser (Eds), *Questions and information systems* (pp. 167–187). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Graesser, A. C., Person, N. K. and Huber, J. (1993). Question asking during tutoring and in the design of educational software. In M. Rabinowitz (Ed.), *Cognitive science foundation of instruction* (pp. 149–172). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Greeno, J. G. (1991). Number sense as situated knowing in a conceptual domain. *Journal for Research in Mathematics Teaching*, **22**, 170–218.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. Morgan (Eds), *Syntax and semantics, vol. 3: Speech acts* (pp. 41–58). New York: Academic Press.

Hammond, K. J., Seifert, C. M. and Gray, K. C. (1991). Functionality in analogical transfer: a hard match is good to find. *Journal of the Learning Sciences*, **1**, 111–152.

Hawkins, J. and Pea, R. D. (1987). Tools for bridging the culture of everyday and scientific thinking. *Journal for Research in Science Teaching*, **24**, 291–307.

Kerry, T. (1987). Classroom questions in England. *Questioning Exchange*, **1**(1), 32–33.

King, A. (1992). Comparison of self-questioning, summarizing, and notetaking-review as strategies for learning from lectures. *American Educational Journal*, **29**(2), 303–323.

Kolodner, J. L. (1993). *Case-based reasoning*. San Mateo, CA: Morgan Kaufmann Publishers, Inc.

Kozma, R. (1992, April). *Mapping external representations onto internal meaning: toward a science of symbolic design*. Paper presented at the meeting of the American Educational Research Association, San Francisco.

Kourilsky, M. and Wittrock, M. C. (1992). Generative teaching: an enhancement strategy for the learning of economics in cooperative groups. *American Educational Research Journal*, **29**, 861–876.

Kreuz, R. J. and Roberts, R. M. (1993). When collaboration fails: consequences of pragmatic errors in conversation. *Journal of Pragmatics*, **19**, 239–252.

Larkin, J. H. and Chabay R. W. (Eds). (1992). *Computer-insisted instruction and intelligent tutoring systems*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lave, J., Smith, S. and Butler, M. (1988). Problem-solving as an everyday practice: In E. A. Silver and R. I. Charles (Eds), *Research agenda in mathematics education: the teaching and assessing of mathematical problem solving* (pp. 61–81). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lave, J. and Wenger, E. (1987). *Situated learning: legitimate peripheral participation*. Cambridge: Cambridge University Press.

Leech, G. (1983). *Principles of pragmatics*. London: Longman.

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

Leinhardt, G. (1987). Development of an expert explanation: an analysis of a sequence of subtraction lessons. *Cognition and Instruction*, **4**, 225–282.

Lepper, M. R., Aspinwall, L. G., Mumme, D. L. and Chabay, R. W, (1990). Self-perception and social-perception processes in tutoring: subtle social control strategies of expert tutors. In J. M. Olson and M. P. Zanna (Eds), *Self-inference processes: the Ontario Symposium* (pp. 217–237). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lepper, M. R., Woolverton, M., Mumme, D. L. and Gurtner, J. L. (in press). 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 Associates, Inc.

Lesgold, A., Lajoie, S., Bunzo, M. and Eggan, G. (1992). SHERLOCK: a coached practice environment for an electronics troubleshooting job. In J. H. Larkin and R. W. Chabay (Eds), *Computer-assisted instruction and intelligent tutoring systems* (pp. 201–238). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Levin, J. A. and Moore, J. A. (1977). Dialogue-games: metacommunication structures for natural language interaction. *Cognitive Science*, **1**, 395–420.

Mayer, R. (1992). Cognition and instruction: their historic meeting within educational psychology. *Journal of Educational Psychology*, **84**, 405–412.

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

Mehan, H. (1979). *Learning lessons: social organization in the classroom*. Cambridge, MA: Harvard University Press.

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.

Miyake, N. and Norman, D. A. (1979). To ask a question, one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior*, **18**, 357–364.

Mohan, M. (1972). *Peer tutoring as a technique for teaching the unmotivated*. Fredonia, NY: State University of New York, Teacher Education Research Center (ERIC Document Reproduction Service No. ED 061 154).

Moore, J. D. and Paris, C. L. (1993). Planning text for advisory dialogues: capturing intentional rhetorical information. *Computational Linguistics*, **19**, 651–694.

Ohlsson, S. (1986). Some principles of intelligent tutoring. *Instructional Science*, **14**, 293–326.

Ohlsson, S. and Rees, E. (1991). The function of conceptual understanding in the learning of arithmetic procedures. *Cognition and Instruction*, **8**, 103–179.

Palincsar, A. S. and Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, **1**(2), 117–175.

Papert, S. (1980). *Mindstorms: children, computers and powerful ideas.* New York: Basic Books.
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.
Person, N. K., Kreuz, R. J., Zwaan, R. and Graesser, A. C. (in press). Pragmatics and pedagogy: conversational rules and politeness strategies may inhibit effective tutoring. *Cognition and Instruction.*
Piaget, J. (1952). *The origins of intelligence*. New York: International University Press.
Pressley, M. (1990). *Cognitive strategy instruction that really improves children's academic performance*. Cambridge, MA: Brookline Books.
Pressley, M., Symons, S., McDaniel, M. A., Snyder, B. L. and Turner, J. E. (1988). Elaborative interrogation facilities in the acquisition of confusing facts. *Journal of Educational Psychology*, **80**, 301–342.
Putnam, R. T. (1987). Structuring and adjusting content for students: a live and simulated tutoring of addition. *American Educational Research Journal*, **24**, 13–48.
Ram, A. (1991). A theory of questions and question-asking. *The Journal of the Learning Sciences*, **3&4**, 273–318.
Reiser, B. J., Connelly, J. W., Ranney, M. and Ritter, C. (1992). *The role of explanatory feedback in skill acquistion*. Unpublished manuscript, Princeton University, Princeton, NJ.
Roschelle, J. (1992). Learning by collaboration: convergent conceptual change. *Journal of the Learning Sciences*, **2**, 235–276.
Rogoff, B. (1990). *Apprenticeship in thinking*. New York: Oxford University Press.
Scardamalia, M. and Bereiter, C. (1991). Higher levels of agency for children in knowledge building: a challenge for the design of new knowledge media. *The Journal of the Learning Sciences*, **1**, 37–68.
Scardamalia, M., Bereiter, C., McLean, R. S., Swallow, J. and Woodruff, E. (1989). Computer-supported intentional learning environments. *Journal of Educational Computing Research*, **5**, 51–68.
Schank, R. C. (1986). *Explanation patterns: understanding mechanically and creatively*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Schank, R. C. and Farrell, R. (1988). Creativity in education: a standard for computer-based teaching. *Machine-Mediated Learning*, **2**, 175–194.
Schank, R., Ferguson, W., Birnbaum, L., Barger, J. and Greising, M. (1991). ASK TOM: an experimental interface for video case libraries. *The Proceedings of the 13th Annual Conference for the Cognitive Sciences Society* (pp. 570–575). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Schank, R. C. and Jona, M. Y. (1991). Empowering the student: new perspectives on the design of teaching systems. *The Journal of Learning Sciences*, **1**, 7–35.
Schegloff, E. A. (1991). Conversation analysis and socially shared cognition. In L. B. Resnick, J. Levine and S. D. Behrend (Eds), *Socially shared cognition* (pp. 150–171). Washington, DC: APA.
Schoenfeld, A. (1988). Mathematics, technology, and higher order thinking. In R. S. Nickerson and P. P. Zodhiates (Eds), *Technology in education: Looking toward 2020* (pp. 67–96). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Sebrechts, M. M. and Swartz, M. L. (1991). Question-asking as a tool for novice computer skill acquisition. *Proceedings of the International Conference on Computer–Human Interaction*. 293–297.
Slavin, R. E. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin*, **94**, 429–445.
Sleeman, D. H. and Brown, J. S. (Eds) (1982). *Intelligent tutoring systems*. New York: Academic Press.
Spiro, R. J., Feltovich, P. J., Jacobson, M. J. and Coulson, R. L. (1992). Cognitive flexibility, constructivism, and hypertext: random access instruction for advanced knowledge acquisition in ill-structured domains. In R. M. Duffy and David H. Jonassen (Eds), *Constructivism and the technology of instruction* (pp. 57–75). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Stevens, A., Collins, A. and Goldin, S. E. (1982). Misconceptions in students' understanding. In D. Sleeman and J. S. Brown (Eds), *Intelligent tutoring systems* (pp. 13–24). New York: Academic Press.
Suchman, L. A. (1987). *Plans and situated actions: the problem of human–machine communication.* Cambridge, MA: Cambridge University Press.
Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science*, **12**, 257–285.
Swigger, K. (1991). Managing communication knowledge. In H. B. Burns, J. W. Parlett and C. L. Redfield (Eds), *Intelligent tutoring systems: evolutions in design* (pp. 13–34). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Van Lehn, K. (1990). *Mind bugs: the origins of procedural misconceptions.* Cambridge, MA: MIT Press.
Van Lehn, K., Jones, R. M. and Chi, M. T. (1992). A model of self explanation effect. *The Journal of the Learning Sciences*, **2**, 1–59.
Vygotsky, L. S. (1978). *Mind in society.* Cambridge, MA: Harvard University Press.
Weaver, C. A. III (1990). *Calibration and assessment of comprehension.* Unpublished doctoral dissertation, University of Colorado, Boulder, CO.
White, B. (1993). ThinkerTools: causal models, conceptual change, and science education. *Cognition and Instruction*, **10**, 1–100.
Williams, S. M (1992). Putting case-based instruction into context: examples from legal and medical education. *The Journal of the Learning Sciences*, **2**, 367–427.
Wittrock, M. C. (1990). Generative processes of comprehension. *Educational Psychologist*, **24**, 345–376.
Woolf, B. (1991). Representing, acquiring, and reasoning about tutoring knowledge. In H. B. Burns, J. W. Parlett and C. L. Redfield (Eds), *Intelligent tutoring systems: evolutions in design* (pp. 127–149). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Woolf, B. and McDonald, D. D. (1984). Context-dependent transitions in tutoring discourse. In *Proceedings of the National Conference on Artificial Intelligence* (pp. 355–361).
Zimmerman. B. J., Bandura, A. and Martinez-Pons, M. (1992). Self-motivation for academic attainment: the role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, **29**, 663–676.