1

Learning from Observing an Expert's Demonstration, Explanations, and Dialogues[1]

MICHELENE T. H. CHI
Arizona State University, USA

Page Proofs, Not for Distribution

INTRODUCTION

A teacher/mentor is sometimes an expert in the sense that they know the content domain in which they have to teach or mentor students, even though they may not be an expert at pedagogy. Nevertheless, as a content expert, s/he can teach in at least three different ways: (1) by modeling and demonstrating to a large number of students, in the context of a classroom, how a problem is solved or how a chemical reaction can be produced, without any verbal interactions with the students; (2) by explaining verbally some difficult concepts or how a problem is solved; and (3) by guiding and interacting with a specific student through extended dialogue to understand a concept or to solve a problem, with other students observing such interactions. In short, there are many ways that teachers/mentors can deliver instruction even though they may not be an expert at pedagogy. The question is, which method of instruction is best for students' learning?

This chapter is a tribute to my late husband, William G. Chase; I am indebted to him for all the ways (described above) that he had mentored me. Like other students, I had the opportunities of learning from him in formal settings, such as listening to his lectures at colloquia and explanations in classes.

1 This research was supported by a grant from the Spencer Foundation (200800196). I am indebted to Marguerite Roy for the new analyses reported in the last part of this chapter.

However, I had the additional privilege of learning from him in many informal settings. In undertaking his daily research tasks, Bill served as a role model to me by demonstrating how he asked critical questions, designed studies without a confound, analyzed data, and even more importantly, how he revised and repaired his design and analyses. But most importantly, before we would go home each day, I would always sit in a corner in his office waiting for him to finish his day, often watching and listening-in to him explaining and guiding the many students who came in and out of his office daily on how to modify their conceptual questions or analyze their data. I am forever grateful for those years of informal observations since it had a profound influence on my research style and methodology. Little did I realize then that the many years of observing him in these different ways have inspired me to now ask whether some ways of observing are more effective for learning than others.

This chapter discusses the advantages and potential limitations of observing each type of teaching/mentoring. Before doing so, this section will briefly clarify and differentiate what will be discussed here from related work on "vicarious learning." "Vicarious learning" is a term that was introduced in the 1960s to describe learning that results primarily from observing the *behaviors* of others and the consequences of those behaviors for their enactors (Bandura, 1989). The important aspect of Bandura's social learning experiments is that even though the observers did not execute the actions being observed or experience feedback from the consequences of those actions, learning still occurred. For example, if children watch other children behave aggressively, even on television, or if they are exposed, even briefly, to violent video games (Anderson & Bushman, 2001), then they are also more likely to exhibit aggressive behavior (Bandura, Ross & Ross, 1961). In this case, they have learned either the aggressive behaviors themselves or they have learned to lower their inhibitory thresholds to exhibit aggressive behavior. Many other studies have shown that people can learn a behavior by watching a demonstration, such as watching a video of how to tie nautical knots (Schwan & Riempp, 2004). The main difference between the kind of tasks vicarious learning studies use and the tasks in the studies to be described later in this chapter is that in many of the vicarious studies, the tasks tend to be physical behaviors whereas our work examines academic tasks such as solving mathematics or physics problems.

In the preceding paragraphs, learning from observing was described under three instructional contexts: (1) watching an expert *demonstrating* (e.g., solving a problem at the whiteboard); (2) watching and listening to an expert *demonstrating* and *explaining* (e.g., solving and explaining the solution steps beyond describing what was written down at the whiteboard); and (3) observing an expert *guiding* another student as s/he works out a problem or task (e.g., answering the student's questions or scaffolding him/her). (Henceforth, in this third context, this other student with whom the expert is directly interacting will be called a "tutee" and the task of problem-solving will be used as an example throughout the discussion in this chapter.) Although the chapter began with questions about whether there are differential learning advantages to each of

these observing contexts, note that only the first two contexts (e.g., watching an expert solve a problem, and watching an expert solve and explain a solution) are prevalent in current classrooms (e.g., in math and physics classes). Thus, in these two contexts, the students are both observing and receiving instruction directly from the expert/teacher (in that the demonstrations and explanations are intended for the observing-students, and the students can have some minimal interactions with the expert). However, the third context, observing an expert guiding another student/tutee (e.g., while the tutee solves a problem at the whiteboard), is less conventional, but it is the only context (among the three) in which the observing-students and the tutees experience different instructional opportunities, in that the tutees get to interact directly with the expert whereas the observers do not. Thus, the central question raised in this chapter is whether the observing-students in the third context can learn as well as the tutees, and if so, under what circumstances.

But before addressing this last context, the limitations of the first two contexts will be reviewed. That is, in the first context, the following question will be considered: how much can students learn by observing an expert visually demonstrating a skill or perform a complex task such as solving a problem, without directly interacting with the expert? In the second context, the question is: how much can students learn by watching an expert demonstrate a skill or perform a complex task and listening to the expert's verbal explanations while demonstrating it? Finally, in the third context, some preliminary evidence will be provided regarding how much students can learn by observing an expert explaining, guiding, and dialoguing with a tutee. Potential interpretations for our results are offered. In authentic classrooms and other situations, a teacher/mentor may combine several of these methods of instruction. However, for research purposes, each of these contexts will be discussed separately.

EXPERTS MODELING WITH DEMONSTRATIONS ONLY

In many classroom and training contexts, experts typically can teach by visually demonstrating to students. This occurs often for many procedural skills, such as serving in tennis, detecting mines, cooking, or knitting. There are many reasons why a task is sometimes demonstrated without many verbal explanations. One practical reason, as in the case of mine detection (Staszewski, Chapter 2, this volume), is that the expert must demonstrate it from afar for safety reasons. A second reason is that the expert often has no accessible knowledge that s/he can articulate, or the articulated knowledge is essentially useless. For example, a tennis coach might explain that the student player should "pull his shoulders back when serving" but that description is too vague for the student to know exactly how that should be done. So there are many occasions when training occurs from demonstrations alone.

Do students learn from watching experts demonstrate a skill or a procedure? Certainly there has been abundant and consistent evidence suggesting

that one can learn a great deal of physical skills from watching, such as anecdotal evidence from anthropology (in weaving, Lave, 1988), and laboratory evidence such as learning to tie nautical knots from watching a video only (without audio; Schwan & Riempp, 2004). The most compelling evidence comes from work in social psychology, as illustrated above, in the context of behaviors such as aggression (Bandura, 1989).

The fact that people can learn by watching physical behavior without direct interaction should not be surprising given that it is a form of imitation that both chimpanzees (Whiten, 1998) and young children undertake. For example, young children will spontaneously pick up books and imitate adults' reading after exposure to models who were reading aloud. These children were learning the behavior of reading, but not the content of what the adult models were reading (Haskett & Lenfestey, 1974). What about other cognitive behaviors such as learning and collaborating skills? Two studies have shown that observing-students can learn cognitive behaviors such as asking questions by watching an animated agent ask questions (Craig, Gholson, Ventura, Graesser, & the Tutoring Research Group, 2000) and by observing others collaborate (Rummel & Spada, 2005). Note that the observing-students were asked to learn the *skills* of asking questions and collaborating, which are observable overt behaviors, but they were not required to learn the specific content of the topics used in the studies. That is, they were not asked to learn the content of the questions asked, but only to learn to ask questions; and they were not asked to learn the content of the collaborative diagnosis of a clinical case, but only to learn how to collaborate.

For the kind of tasks illustrated above, such as acting aggressively and asking questions, the behavior to be mimicked is directly displayed so that it is observable. However, complex academic task performance often involves many covert inferences and reasoning steps before an overt step is taken, so that there is not a one-to-one correspondence between the actions displayed and the reasoning behind the displayed actions. For instance, suppose a teacher is solving a geometry problem on the whiteboard by writing down a sequence of six steps (some involving equations). Students can obviously "learn" this particular sequence of steps, in that they can reproduce these steps for solving similar problems (VanLehn, Graesser, Jackson, Jordan, Olney, & Rose, 2007). But can they solve other less similar problems? Abundant evidence in the literature shows that students generally cannot transfer what they have learned from studying or watching an example solution to a new, slightly different problem (Reed, Dempster, & Ettinger, 1985; Ross, 1987), because they have not learned the reasoning by which the original steps were derived. In short, the cognitive processes underlying the decision about what step to display (on the whiteboard, for instance) maps many to-one (the actual step displayed). Similarly, if a novice had to learn how to troubleshoot a complicated piece of equipment, watching an expert throw all the switches on a board does not tell the observing novice how the problem was diagnosed. Because there is a lack of one-to-one mapping between the cognitive processes underlying the decisions for which switches to throw and the observable actions of throwing switches, we concluded back in 1991 that one cannot learn by

watching an expert perform a cognitively rich task (Chi & Bjork, 1991), unless the displayed steps of the task map one-to-one with the reasoning behind the steps (Carroll & Bandura, 1990).

In sum, watching an expert demonstrate a task is adequate for enabling the observing-students to learn how to perform that task to the extent that the task is a physical or behavioral one where each step is overtly displayed. However, if a task requires multiple deep reasoning steps before an overt step is displayed, then learning from watching an expert demonstrate some of the procedural or behavioral steps is more limited. An extreme example might be watching an attending physician inject a drug into a patient. An observing medical student will not learn much from watching this action unless s/he also understands the explanations for the injection. Similarly, students do not understand the solution steps a teacher writes on the whiteboard (because they typically just copy the steps from the whiteboard without knowing the reasoning behind each step), unless the students provide their own reasoning for each demonstrated step, as in self-explaining (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). Thus, learning from watching a demonstration of a complex task is shallow with limited understanding.

EXPERTS MODELING WITH DEMONSTRATIONS PLUS GIVING MONOLOGUE EXPLANATIONS

What about the case of an expert modeling while providing explanations along with it, but without giving feedback to the student (thereby there is no direct interaction)? By *explaining* the steps of a problem, we mean explaining the reasoning that leads from one step to the other, as opposed to *describing* the steps of a problem, which is only uttering verbally what each step is or does. To use knitting as a contrast to a cognitively rich task, telling the student verbally to first put the needle through one loop, then another, is a *description* of the steps, corresponding almost to the student watching the steps. The verbal descriptions did not constitute *explanations* because they do not add any further clarification, justification, or information. One prevalent example of this context is teaching in a large lecture class. Typically, a teacher instructs a class by demonstrating and explaining. For instance, a math teacher can demonstrate how to solve problems by writing down the solution steps on the whiteboard, and explaining each step as s/he writes. Many instructional videos (in the market, or on television) also teach numerous everyday skills through demonstrating and explaining, such as in cooking. Unfortunately, in many of these cases, the explanations are merely descriptions of the demonstrated steps.

The question addressed in this section is: For a complex cognitive task, such as solving geometry and physics problems, in which there are many reasoning steps corresponding to one overt step, are experts' accompanying verbal explanations helpful to students' learning? That is, if instructors are asked to articulate the intervening reasoning for each step (sometimes referred to as making thinking visible, Linn, 1995), will hearing those explanations be helpful, especially for deep learning? Evidence across a variety of research contexts shows

that in general, expert-articulated explanations do not always help students achieve a great deal of learning or transfer, whether it is in a classroom context (Deslauriers, Schelew & Wieman, 2011), or in a tutoring context (Chi, Roy, & Hausmann, 2008). For example, Deslauriers et al. (2011) compared students' learning in quantum mechanics classes that were either instructed by an expert faculty member (one who was motivated, had many years of experience teaching this quantum mechanics course, and had received high student evaluations) who lectured and demonstrated how to solve mechanics problems, with interactive classes that were instructed by a postdoctoral fellow in which students could ask questions and receive feedback, and worked in small dyad groups. A 2.5 standard deviation difference in learning was obtained between the two conditions, consistent with our interpretation that students do not learn very well by listening to an expert explain with demonstrations. Compatible with the minimal learning from listening to an expert's explanations is the finding from expert–novice research showing that learners perform significantly better when instructed by novices than by experts in an electronic wiring task (Hinds, Patterson, & Pfeffer, 2001), suggesting that answers and explanations are helpful to a questioner only if the answers are adjusted to the appropriate level of the question (Webb, 1989). This is also consistent with Feldman, Campbell and Lai's (1999) point that "Students of similar achievement levels may be more effective than teacher–student pairs because peers can discuss strategies at their own novice's zone of proximal development." In a similar vein, in the audience-design literature, it has been shown that experts underestimate the complexity of a subject and their explanations of it (Bromme, Jucks & Runde, 2005; Clark & Murphy, 1982). Thus, this body of evidence suggests that explanations have to be tailored to the level of the recipient's understanding in order for them to be effective, and since experts are not always accurate at gauging a novice's understanding (evidence for this to be presented below), this is one possible reason why experts' explanations, when they accompany demonstrations, may not be that helpful for deep learning.

One reason that experts cannot accurately gauge a novice's understanding (therefore making their explanations ineffective for novices' learning) is that there is usually a mismatch between the normative mental model from which the expert is generating an explanation of a concept, a problem step, or a system, and the naïve mental model to which a student has to assimilate that explanation. We have direct evidence showing that experts, while tutoring, are inaccurate at predicting their tutees' exact level of understanding (or mental model) of the circulatory system (Chi, Siler, & Jeong, 2004). For example, using the technique of categorizing a student's mental model of the circulatory system into six levels, with levels five and six as the most accurate, tutors tend to think that students have a mental model at a fifth or sixth level when, in fact, students have only a second or third level understanding. This finding was based on data collected from interrupting both tutors and their tutees at two points during tutoring to ask the tutors what they thought their respective tutees' mental model of the circulatory system was, as well as asking the tutees

themselves what their current understanding was. The tutors seemed unaware of (or were inaccurate in predicting) the nature of their tutees' flawed mental models. This suggests that an expert's explanations cannot be targeted precisely at what a tutee misunderstands. Instead, a tutor generates feedback on the basis of a mismatch between a tutee's incorrect response as compared with the correct normative response, rather than tailoring the feedback to the tutee's naïve understanding. This finding of a mismatch between an expert's representation and a student's representation or understanding is consistent with the expert blind spot hypothesis, which states that greater content knowledge skews expert teachers' ideas of what students understand. For example, high school teachers assumed that verbal problems were more difficult for students to solve than symbolic problems, when, in fact, the reverse was true (Nathan & Koedinger, 2000). This is because high school math teachers take a domain-centered (or normative) view, reasoning that symbolic problems would be easiest for students to solve because they were written in "pure math."

A second reason for the ineffectiveness of watching and listening to an expert's demonstration along with monologue explanations for cognitive-rich tasks is that experts often cannot articulate all the intervening reasoning steps, because they are often not consciously available to them (Ericsson & Simon, 1984). This may be due to their having automatized or chunked the intervening steps, or alternatively, the intervening reasoning may be implicit (as illustrated in our tennis serving example described above). If the reasoning is not accessible to them, then their explanations often become a description of the procedural steps themselves, which are not a helpful sort of explanation since they do not add any information.

In sum, experts' monologue explanations, when they accompany demonstrations, may not be as helpful as one may think for two possible reasons. The first reason is the mismatch between a novice's and an expert's representations, making it difficult for students to understand experts' explanations when students' representations are naïve. The second reason is that experts' explanations may not be complete or accurate because experts' knowledge may be implicit or automatized, causing it to be inaccessible.

EXPERTS MODELING WITH DEMONSTRATIONS PLUS SCAFFOLDING DIALOGUES

The preceding sections described two common instructional contexts. In the first context, an expert demonstrates a skill or performs a procedure, while students watch the demonstration. Although students can learn in this context, it has its limitations for cognitively rich and complex tasks. The second instructional context involves an expert demonstrating while explaining. Students can learn in this context as well, but again there are limitations for cognitively rich tasks, such as learning to solve complicated problems. In both of these contexts, the students are observing with some minimal opportunities to interact with the

expert/mentor. However, it has been well documented in the tutoring literature that when students have a large number of interactions with an expert, such as interacting with a tutor one-on-one, then learning is maximized. Although such interactions are obviously advantageous, one-to-one interactions with an expert/tutor are costly. However, there is a third possible instructional context, that is for students to observe the dialogue between a tutor and a tutee, without directly interacting with the tutor. In this context, the observing-students do not get to interact with the tutors; only the tutees get to interact directly with the tutors. For this context, the question is, can the observing-students learn from overhearing such dialogues without directly interacting with the experts? In other words, can students learn from watching others learn (Mayes, Dineen, McKendree, & Lee, 2001; McKendree, Stenning, Mayes, Lee, & Cox, 1998)? There are obvious advantages to knowing this answer because it means that one can scale up an expert's dialogue with a tutee by showing or reusing it with many other observing-students.

There are many situations in which students are expected to learn by observing and listening-in to the dialogue that occurs between a teacher and a student (such as when one of the students is being called to solve a problem at the whiteboard, guided by the teacher while the rest of the students watch and listen); or a master cellist guides and demonstrates to a student while an audience watches and listens), or between an expert and a novice (such as a chef on a cooking show teaching a naïve cook how to prepare a dish; or an adult teaching a young student about science demonstrations in shows such as *The Wizzards* (Field, 1983–1990; 1991–1995). The most serious enterprise that takes advantage of this form of training is in the medical field, where interns go on rounds with an attending physician and a resident. In such clinical training settings, typically the attending physician either does the demonstrating (such as palpating the patient) or gives the resident feedback while s/he palpates the patient. Like a tutor, the attending physician focuses his/her attention and dialogues on the resident/tutee, while the rest of the interns simply listen and watch. They are given opportunities to ask questions, but the main dialogue is between the attending physician and the resident. Becoming more prevalent are related contexts, such as remote telemedicine. Suppose a nurse talks to a remote medical consultant about a child patient, and the child's mother is present and listens in. What can we assume that the mother has learned about the child's condition by overhearing the conversation between the nurse and the medical consultant? Can we assume that she has learned enough so that she can later monitor and treat her child appropriately? Thus, the same question applies in all these cases, that is, do the observers who overhear dialogues, learn anything and if so, what, how and how much?

Two Constructs Relevant to Observational Learning

Before describing the utility and prevalence of an instructional context in which students observe an expert guiding and dialoguing with a novice, we briefly clarify the kind of "observers" we are interested in, and the issue of intentionality.

Types of Observers There are several nuances to the meaning of "observers," especially pertaining to the listening aspect of observing. According to Goffman (1976), there are three types of listeners in the context of a conversation or dialogue. A *side participant* is a participant recognized by both the speaker and addressee as a full member of the conversation. A *bystander* is a participant who both the speaker and addressee are aware can overhear them, but who is not participating in the conversation. Finally, an *eavesdropper* is an over-hearer of which neither the speaker nor the addressee is aware. In this chapter, by "observers," we mean it in an "eavesdropping" sense in that neither the speaker nor the addressee is necessarily aware of the presence of the observer. Therefore, there are no mutual obligations between the speaker and the observer to monitor and repair each other's understanding, in a process called "grounding" (Clark, 1996). With "grounding," individuals who are engaged directly in conversation with each other must make sure that they share common knowledge (such as to whom they are referring, or what event they are talking about) in order to be understood and have a meaningful conversation. However, an observer would not share grounding experiences with the speaker. Therefore, according to Clark's theory, an "eavesdropper-type" of over-hearer would not fare as well as the addressee in understanding what a speaker says, since the over-hearer did not participate in grounding with the speaker.

In summary, in this chapter, we focus on the "eavesdropper-type" of observers, also referred to by Rogoff, Paradise, Arauz, Correa-Chavez and Angelillo (2003) as "third-party observers". We refer to such observers as *observing-students*, and refer to the addressees as *tutees* since they are participating in a one-on-one dialogue with the tutors. The eavesdropper-type of observer is also prevalent in the context of online discussion (Sutton, 2001) and, in fact, they make up the majority of members in online groups, usually referred to as "lurkers" (Nonnecke & Preece, 2001). Because the term *overhearing* connotes passivity and *eavesdropping* connotes a forbidden nature, we agree with Rogoff et al. (2003) that the term *listening-in* is preferable. In short, in this chapter, the term *watching* will be used to refer to observing visual inputs such as gestures, diagrams, or overt behaviors of models; the term *listening* will be used to refer to hearing a monologue explanation; the term *listening-in* will be used to refer to accessing or hearing dialogue that was not meant for the listeners; whereas *observing* will be used as a general term to refer to both *watching* and *listening-in*.

Intentional Participation or Doing

A relevant question for vicarious learning is the sense of *intentionality*. Intentionality can be considered from two perspectives. From the perspective of the teacher/mentor, instruction is typically targeted either at a co-present interacting tutee or no student at all (such as in an online learning environment or even in a large lecture hall); but it is not intended for a listening-in observing-student. However, from the perspective of the observing-students, in many scenarios, they do intend to learn even though instruction is not targeted

at them. For example, in many cultures in which children are integrated into mature community activities, children are expected to learn without direct guidance and monitoring from adults. Instead of directly interacting with adults as conversational partners, children in many cultures learn by *participating* in aspects of adult activities, while *observing* demonstrations and *eavesdropping* on adult talk in context, which includes questions and directives (Rogoff, 1990, p. 122). A key component of third-party observation in such a context is that children are actively and *intently participating in* or *doing* related chores while "keenly observing and listening," (Rogoff et al., 2003, p. 178) because they anticipate undertaking the activity that they are observing. The notion of *intent participation* implies that observational learning is not incidental or accidental, from the perspective of the observing-learners, but requires *actively doing a relevant task* that may lead to learning. Thus, borrowing from Rogoff et al. (2003), we define *intent participation* in a laboratory study as being required to undertake or *do the relevant task* for which the observers are asked to learn.

In this chapter, we focus on the first question, that is, can observing-students learn without directly interacting with the expert/mentor, if they have the goal of *intending-to-learn*? Our hypothesis, consistent with Rogoff et al.'s (2003), is that observational learning can be as beneficial as directly interacting with a tutor, if the observers are *intently* trying to learn by participating or doing the task. We focus on learning complex concepts and procedures, ones that are taught in school, rather than physical skills that can be directly displayed and mimicked.

Three Different Ways of Observing Tutorial Dialogues

There are three different ways of observing tutorial dialogues: (1) alone; (2) collaboratively in pairs; and (3) collaboratively in pairs with *intent participation* or doing. We consider relevant studies to each situation in turn, to see whether and how much students learn.

Watching and Listening-in on Dialogues in Solo There is only scant and inconsistent evidence on how well observers can learn by overhearing instructional or tutorial dialogues, without direct interactions with the instructor or tutor. Three sets of relevant studies are described here. In Schober and Clark (1989), the target task used is the placement of a sequence of random shapes by a Matcher (analogous to a tutee) under the direction of a Director (analogous to a tutor) who tells the Matcher which shape to pick up and place next. Both the Matcher and the Director each have the same set of random shapes (or tangrams), and a screen blocks the views of each other's tangrams. The Director has to describe the shape of each tangram in such a way so that the Matcher will pick the correct one to place down. The Matcher may ask the Director for clarifications and other information. How successfully the Matcher placed the random shapes in the sequence requested by the Director was compared against how successfully a bystander (analogous to an observing-student)

placed his/her own set of tangrams. The bystander could overhear the dialogues between the Director and Matcher, but could not see their tangrams. The results showed that Matchers were significantly more accurate than the solo bystander-observers. This suggested that being directly guided by and interacting with a Director, as in the case of a Matcher, was better than overhearing the dialogue as a solo bystander.

The Schober and Clark study did not measure learning; instead it measured the ability to correctly place the referred-to tangram in the sequence as instructed by the Director, which basically required coming to an agreement about how to refer to a tangram. For example, one ambiguous shape was first described by the Director as "a dancer or something really weird. Um, and has a square head . . . ," and the Matcher followed with a clarification question of, "Which way is the head tilted?" and the Director said "The head is..eh..towards the left, and then th-an arm could be like up towards the right?..and It's-" At this point the Matcher said, "an a big fat leg?" They then settled on referring to this shape as "the dancer with the big fat leg." In short, in the Schober and Clark study, the task mainly consisted of linguistically finding ways to refer to a random shape that both the Director and the Matcher agreed upon, since placing the shape in a sequence is trivial once the correct one is identified.

A similar comparison, very much like the one in the Schober and Clark (1989) study, was carried out by Craig, Driscoll and Gholson (2004), with the exception that the latter study measured learning. Craig et al. (2004) compared how well students learned 12 computer literacy topics through either interacting directly with a tutoring system, or by observing recordings (screen captures of the voices) of the tutorial interactions between the computer tutor and a student. The results of Experiment 1 were consistent with the results of Schober and Clark (1989), in that the tutees who directly interacted with the tutor learned significantly more than students who merely observed the screen captures of tutoring.

One way to understand and interpret the results of these two studies is that in contrast to the tutees, the solo observers were not very generative or constructive. Being generative, in the traditional associative recall type of memory studies (Slamecka & Graf, 1978; Stein & Bransford, 1979), meant that the participants were able to generate their own elaborations or associations. The classic finding from the memory literature was that when participants generated their own associations between two words, they tended to have better recall of the words. This has been referred to as the generation effect. Similarly, being constructive (as in self-explaining) means that learners are constructing a relationship between two solution steps or generating an inference that ties two sentences together, and doing so enhances learning (Chi, et al., 1989; Chi, de Leeuw, Chiu, & LaVancher, 1994). Based on the snippet of negotiation on how to refer to a shape described in the above example, a simple interpretation for the advantage of the Matcher's performance over the observer's performance is that the Matcher was the one who was generating part of the referral label, especially the part of "the big fat leg." An observer in this task, however, was not at liberty to generate any references.

In sum, an interpretation for the advantage of the Matcher's performance is simply that s/he had opportunities to be generative, which has been shown to be advantageous in many memory studies and self-explanation studies, whereas the solo observers had no opportunities to be generative. Thus, an alternative interpretation for the poorer performance of the solo observers in the Schober and Clark (1989) study is their lack of opportunities to be generative, as opposed to the interpretation that the Matchers had opportunities to be interactive with the Director. Of course, it is possible for observers to be covertly generative, but it is reasonable to assume that one is more likely to be generative when participating in a conversation, as in the case of the Matchers, than when merely observing alone.

In the tangram-placing task, even though the solo observers did not perform as well as the Matchers, overhearing a dialogue between the Matcher and the Director was nevertheless better than overhearing just the Director explaining the placement and sequencing of the tangrams in a monologue (Tree, 1999). This finding, that listening to dialogues is more effective for performance than monologues, has now been supported by several other studies that did measure learning (Craig et al., 2000; Craig, Chi, & VanLehn, 2009; Muller, Sharma, & Reimann, 2008). The advantage of dialogues over monologues confirms our assumption in the prior section that students do not learn very much from listening to an expert's monologue explanations with demonstrations.

In sum, the overall findings from both a matching tangram task and a learning task show that listening to an expert dialoguing with a novice is not an effective instructional method for solo observers. ~~inviting the conclusion favored by Schober and Clark (1989) that observing dialogues is not as effective a means to learn as directly interacting with an expert.~~

However, listening to a dialogue is superior to listening to a monologue, but neither is as good as directly interacting with an expert.

Watching and Listening-in on Dialogues in Pairs As stated above, our interpretation for the ineffectiveness of an observing dialogue paradigm is that the solo observers were neither generative nor constructive. To increase active participation, Craig et al. (2004, Experiment 2) asked students to observe tutoring dialogues in pairs. However, the paired observers still did not learn as much as the tutees who had the opportunities to directly interact with the tutor, and the paired observers also did not learn any more than the solo observers. Given that collaboration typically improves learning over solo activity, this latter finding is surprising. Our interpretation for the equivalent learning of the paired and solo observers in the Craig et al. (2004) study is that the dyads did not interact much with each other; in fact, on average they each took only three conversational turns in a 35-minute session. One reason that they were not very interactive could be that they were not motivated to learn, in the sense that they were not *intently participating* (i.e., they were not required to do a task while watching the video). Thus, even though the observers were paired up, because they were not very interactive with each other (Craig et al., 2004,

Experiment 2), they essentially were observing alone. So, in essence, we can only conclude that observing without *intent participation* (even if in pairs) is not a very effective way to learn, and is comparable to observing alone.

To summarize, for the studies cited above, three hypotheses can explain the inferior learning of the observing-students. The first hypothesis, favored by Shober and Clark (1989), is simply that a student needs to interact directly with an instructor in order to learn. A second hypothesis, presented here and in Chi et al. (2008), is that the observing-students did not have opportunities to be generative (either with a tutor or with a peer-partner). A third related hypothesis, inspired by Rogoff's work, is that observing-students, even when paired up, may not be motivated to intently learn, and so are not very interactive. In contrast, tutees typically interact much more frequently with their tutors. For example, in our natural tutoring data with inexperienced tutors (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001), the tutors took on average 256 turns whereas the tutees took 224 turns, over a span of around one-and-a-half to two hours. This suggests that in a 30-minute interval, the tutees took at minimum, over 60 turns, as compared with the three turns in Craig et al.'s (2004) data. Thus, there is no question that tutees have many opportunities to interact with a tutor and can thereby be constructive in their interactions. The observers in Craig et al.'s study, even though paired, were obviously not as interactive as tutees could be, as shown in the data of Chi et al. (2001). Thus, it seems that a scenario in which observing-students really have opportunities to be interactive (thereby constructive) and intently participate, is needed in order to test whether observers can, in fact, learn as effectively as tutees from listening-in on dialogues. The next section describes a study that created a situation in which observing-students had to *intently participate,* in the sense of doing the task at hand, and that doing may cause them to interact more frequently.

conversational

2004

Watching and Listening-in on Dialogues in Pairs with Doing In Chi et al. (2008), we asked pairs of students to solve physics problems collaboratively while simultaneously watching a video of a tutee being guided by a tutor in solving the same problem. Making the observing-students collaboratively solve the same problems while watching and listening to a video of tutorial dialogue likely forced them to participate more intently (as opposed to the Craig et. al. 2004 study, where the paired observers only had to listen/watch the video and were not required to do any other activity). Our study consisted of a tutoring condition (Tutoring), which served as the benchmark condition, since tutoring typically offers the largest learning gains in terms of effect size for students. The second condition (Pairs + Video) was the target condition consisting of dyads collaboratively observing a video of tutoring while doing a task, but without opportunities to interact directly with the tutor. These two conditions (Tutoring, and Pairs + Video) were compared in terms of the amount learned by the tutees and the paired observing-students. The learning gains of the tutees and the observing-students were also compared with three other conditions: pairs of students collaborating to solve problems with the aid of a

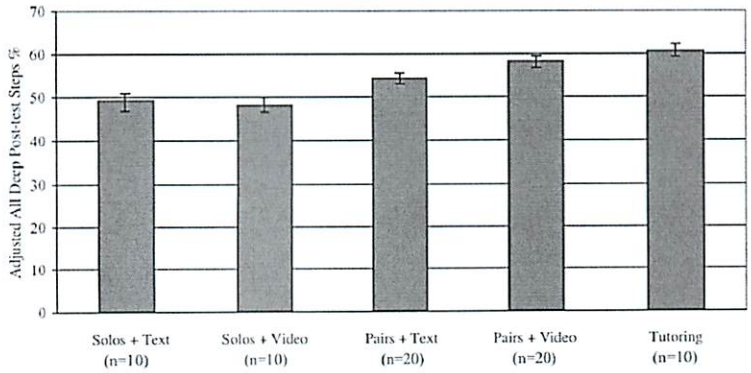

Figure 1.1 Adjusted mean proportion correct of all deep post-test steps, controlling for all deep pre-test steps for each condition. The error bars are standard errors.

textbook (instead of a tutoring video; Pairs + Text); solos solving problems while studying a tutoring video alone (of the same videos used in the Pairs + Video condition; Solos + Video), and, finally, solos solving problems with the aid of a textbook (Solos + Text).

Ten students served as tutees, and so 10 videos were created from an expert instructor who tutored each individual student. The instructor was a 30-year veteran college physics professor and tutored each tutee in solving three mechanics problems. The 10 videos were then shown to pairs of collaborative observers (the target Pairs + Video condition), and to individual observers (in the Solos + Video condition). In the two text conditions, pairs or solos had access to the relevant portions of a physics textbook (Chapters 1–5) that contained worked examples similar to the problems that they had to learn to solve. All conditions received the same pre-test and post-test. In the four non-Tutoring conditions (whether studying the video or text, either alone or in pairs), all the students had to solve the same problems that the tutees in the Tutoring condition were learning to solve. Therefore, this manipulation should have required all students to participate intently.

Figure 1.1 shows the results expressed as the mean (and standard error) of post-test scores adjusted for pre-test scores in an ANCOVA. Scores were based on correctness of deep problem-solving steps. (Details of the scoring procedure are presented in Chi et al., 2008.) Foremost, consistent with the tutoring literature, the Tutoring benchmark condition (right-most bar) had the largest learning effect. This is not surprising especially given that our tutor was not only an expert physicist who had taught physics for over 30 years, but moreover, he was employed full-time by another project as a physics tutor, where his main job was to tutor students and then analyze transcripts of his tutoring to find ways to improve it.

What is amazing is that in our target condition (Pairs + Video – second column from the right in Figure 1.1), the paired observers learned about as well as the tutees in the Tutoring condition. (Figure 1.1 shows adjusted post-test

scores; but to get a sense of the magnitude of gains, we report pure gain scores, which are 17% for the paired observers and 21% for the tutees.) Both of these conditions were significantly better than the other three conditions, as analyzed via pair-wise comparisons, which supports our assertion that students from the Pairs + Video condition learned about the same amount as those from the Tutoring condition, our benchmark. This equivalence of the Tutoring and Pairs + Video conditions suggests that pairs of observing-students can learn without directly interacting with a tutor.

The results of this study thus support the second and third hypotheses described above, that because students in the Pairs + Video condition had to participate intently (in solving the same problems), they had opportunities to be generative and interactive, resulting in learning gains equivalent to those of students in the Tutoring condition. Thus both factors (generative and interactive), together, account for the superior learning of the paired observers (superior in the sense that they did learn as well as the tutees even though they had no opportunities to interact directly with the tutor). Although intent participation (Rogoff, et al., 2003) or hands-on practice while observing (Shebilske, Jordan, Goettl, & Paulus, 1998) is an effective way to learn, this factor alone is not sufficient to enhance paired observers' learning. This is confirmed by the results from the Solos + Video condition, wherein the observing-students also had opportunities to intently solve problems, but they did not learn as well as the students in the Pairs + Video condition, suggesting that intently participating alone is not a sufficient factor for learning from observing dialogues, but being interactive with a partner also seems to be critical.

If interacting with a peer-partner might have contributed to the paired observers' learning, is there evidence that they were, in fact, interacting? In our data, the paired observers (in the Pairs + Video condition) interacted, on average, 121 conversational turns per 35-minute interval, as compared with three conversational turns per 35-minute interval in Craig et al.'s (2004) data. Thus, in our study, *intent participation* may have led to more interactions between the dyads, and greater collaborative interactions must have approximated the amount of interactions that tutees experience with tutors.

How Did the Paired Observing-students Learn?

In the preceding section, we showed that in a study described in Chi et al. (2008), observing-students can learn as well as tutees if they observed in pairs, because that allows them to be interactive with a peer-partner. In this section, we explain *how* students learned from observing a video of tutorial dialogues collaboratively.

In general, tutees learn because they usually have the benefit of: (1) being constructive while interacting with a tutor (such as opportunities to ask questions, make suggestions, and attempt steps; Chi et al., 2001); (2) receiving corrective feedback in terms of whether a solution step or explanation is correct or incorrect; and (3) receiving a variety of scaffolding guidance from tutors that

might elicit more constructive learning. Likewise, the paired observers in our study, as collaborators, can also learn by: (1) being constructive while interacting with each other; (2) receiving feedback from each other; and (3) receiving a variety of scaffolding guidance. However, paired observers, in contrast to tutees, do not receive corrective feedback or scaffolding from an expert/tutor. So, because tutees benefit from receiving direct feedback and guidance from the tutor, logically, observing-student pairs must have learned in a different way compared to the tutees. Is there any evidence to suggest that observing-student pairs learn via a different mechanism from the way tutees learn? In this section, we present new data and new analyses beyond the ones reported in Chi et al. (2008) in order to shed light on this question of how the paired observers might have learned.

One would assume that tutees typically pay attention to their tutors, especially when the tutors request responses from them, as dictated by Grice's conversational pragmatics. The natural tutoring data presented in Chi et al. (2001) showed this assumption to be true. In the tutoring protocols of that 2001 study, we identified three categories of tutor requests (such as asking questions, scaffolding tutees with a hint, and asking whether tutees understand). There were a total of 94 such requests and there were a corresponding 104 tutee responses. These approximately equivalent counts of the number of tutor and tutee statements point out that tutees typically respond to tutors' requests, confirming our assumption that tutees do pay attention to what their tutors ask and may have learned from the tutors' scaffoldings. If paired observers cannot interact with a tutor, will they still pay attention to the tutor?

To answer the last question, we analyzed the dialogues of the paired observers from the Chi et al. (2008) study in the following way. We segmented the paired observers' conversation into episodes, with each episode addressing only one concept or one problem solving step. We then narrowed down and focused only on episodes in which the paired observers were trying to resolve and make sense of what a segment in the video was explaining or doing. Within each of these resolving episodes, we coded whether the paired observers' discussion referred to what the tutee did and said, or what the tutor did and said (as well as referrals to the whiteboard or referrals to both the tutor and the tutee). The analysis of interest is to contrast referrals to the tutor versus the tutee. We found, surprisingly, that the paired observers referred significantly more frequently ($p < .02$) to what the tutees said (6.8 episodes per problem) than what the tutor said (1.8 episodes per problem). This result is counter-intuitive, as we expected more references to what the tutor said, consistent with the common assumption stated above that tutees do pay attention to (and presumably learn from) what tutors say. The fact that the paired observers referred more often to what the tutees said over what the tutor said, could be interpreted as undesirable, in that it was done out of necessity to understand what tutees said. For example, the paired observers could have attended to the tutor's comments, but possibly then had to spend more time assessing the tutees' understanding of what the tutor said, which may have been confusing and error-prone.

To disambiguate whether the paired observers simply *preferred* to attend more to what the tutees said, or *needed* to take more time understanding what the tutees did and said, we assessed whether attending to what tutees say is beneficial to learning. To assess learning, we divided the tutees into Good-learner versus Poor-learner tutees, based on how well the tutees learned on their own (as assessed by a pre-test that was administered *prior* to tutoring but *after* tutees learned some relevant background knowledge on their own). The results reported in Figure 3 of Chi et al. (2008, p. 316) showed that the Pairs + Video observing-students learned significantly more from Good than the Poor tutees (mean gain of 21.9% vs. 7.8%, $p < .01$); and they also referred to what Good tutees said in 9.75 episodes, but referred to what Poor-tutees said in only 3.8 episodes. Thus, this pattern of results suggests that when observers attend even more to what tutees say (in the case of Good tutees), they then learn significantly more. This confirms the interpretation that observers *prefer* to attend to what tutees say, and rejects the interpretation that observers *have to* attend to what tutees say in order to understand them better. If observers were attending simply because they *had to* understand tutees better, then one might expect them to have paid more attention to the Poor tutees, since poorer tutees made more errors and their tutorial dialogues were more confusing (for example, the Poor tutees expressed confusion twice as often as the Good tutees (Chi et al., 2008, p. 315). Instead the opposite was found – observers paid far more attention to Good-learner tutees.

than to what tutors say

Why did the paired observers overall prefer to pay more attention to the tutees over the tutor? Aside from the social reason that observers may prefer peer models (Davidson & Smith, 1982), we speculate on four possible cognitive reasons for why paired observers might find what the tutees say and do more learnable than what a tutor says and does.

Zone of Proximal Representational Match One reason that paired observers may have preferred to pay attention to what the tutees said is because the observers' understanding (or representation) matches more closely with the tutees' understanding (since they are both novices), than with the tutor's normative understanding. We might call this the "zone of proximal representational match." Chi, Feltovich and Glaser (1981) and Chi et al. (2004) have shown evidence for a *mismatch* between novices' naïve representation and experts' normative representation. Thus, we suggest that the paired observers in our 2008 study paid more attention to the tutees perhaps for the same reason, i.e., that their naïve representation matched more closely with the naïve representation of the tutees, and therefore they could understand better what the tutees were asking and saying.

Explaining a Tutee's Errors A second reason for why attending to tutees' contributions might help observers learn has to do with the errors that tutees make. Although there is a large body of evidence in several areas of literature showing that one can learn better from making "errors" (Ben-Zeev, 1995),

"impasses" (VanLehn, 1988), "conflicts" (Dreyfus, Jungwirth, & Eliovitch, 1990; Piaget, 1952), and "productive failures" (Kapur, 2008), there is very little evidence showing that one can learn better by observing someone else make errors.

There is, however, suggestive evidence from three diverse sets of literature. From developmental literature, Siegler (2002, Figure 1.5), for example, has shown that asking children to explain an adult's correct and incorrect reasoning led to greater learning than explaining an adult's correct reasoning only. Given that tutors rarely, if ever, display errors, but that tutees, in general, do tend to make errors while being tutored, the observing pairs of students in our study might have preferred to attend to the tutees more than the tutor in order to explain (and learn from) the tutees' errors.

Related findings also occur in social psychology. For example, Schunk, Hanson, and Cox (1987) investigated how children learn by observing a model solve fraction problems. Children learn significantly better when they observed a model who had to cope (or struggle to solve the problem) versus a model who demonstrated a smooth master performance. A master model is like a tutor who performed all the operations in the problem correctly without any hesitations or uncertainties, stating confidently that, "I can do this one" or "I'm good at this." A coping model is analogous to a tutee who made errors, was hesitant, and made comments such as, "I'm not sure I can do that" or "I'm not very good at these." The observing-students not only learned better by watching the coping model, but they had significantly enhanced self-efficacy.

Finally, there is also suggestive evidence in the vicarious learning literature. In particular, Monaghan and Stenning (1998, Study 1) found that students who watched a video of a model struggling to solve syllogism problems learned slightly more than students who watched a model sailing through the problem solving, although the difference was not significant.

If observers learn from watching tutees struggle, one might wonder then, why did the paired observers learn less from the Poor-learner tutees, since the Poor tutees made many more errors? We can provide an explanation based only on informal analyses of both the Good and the Poor tutees' video tapes. The informal analyses of the videos suggest that the dialogues with the Poor tutees were very disorganized and confusing. Although we had reported that the Poor tutees gained 15.6% from the pre-test to the post-test (whereas the Good tutees gained 24.6%), we failed to emphasize in the 2008 study (Chi et al., 2008, p. 315) that the Poor tutees' pre-test scores were around 30%, whereas the Good-learner tutees' pre-test scores were around 50%. After tutoring, the Poor tutees' post-test scores were still below 50%. Therefore, one could argue that the Poor tutees' performance was below a critical threshold, and therefore their dialogues could not serve as good learning materials.

In sum, there are various findings from disparate literature showing that there is some benefit to observing and explaining a tutee's errors, accounting for why observing-students might learn via a mechanism that requires them to focus on the tutees. The benefit of doing so may overcome the limitation of not being able to interact directly with a tutor.

Abstract versus Grounded Descriptions In reading and studying the tutorial protocols, we noticed an intriguing and prominent difference between the way the tutor talked about a concept or a principle and the way a tutee talked about it. We characterize this difference as being *abstract* versus *grounded.* That is, when a concept/principle is discussed in an abstract way, the description involves context-general information about characteristics or definitions that are not specific to the current problem situation, whereas when it is discussed in a grounded way, the description relates to specific features of the problem situation at hand. Here are two examples of comments coded as abstract; the first one made by the tutor and the second one made by a tutee, taken from our data (Chi et al., 2008):

Tutor: Normal force is always normal and is perpendicular to the surface in contact.
Tutee: The acceleration is due to gravity?

In contrast, grounded comments are ones that are context-dependent, such as these following two examples:

Tutor: This normal force . . . first look at the weight . . . uh, weight force. Is there any other force acting downward?
Tutee: There, weight is pushing down. [Gesturing pushing down]

Context-sensitive comments can be identified easily by deitic references such as, "this," "there," "look at," and so forth.

To reduce the amount of data we had to code, we examined only the "critical nodes." A problem solving node, as defined in the original coding in Chi et al. (2008, see Figure 1 on p. 310), corresponded to a state in a problem space. Each node defined a subgoal, such as "draw a free-body diagram of forces on Mass A in the X direction." Critical nodes were a subset of all the nodes, and they were the more important ones, defined as involving sub-procedures for which critical parts of the problem solution were generated, or for which important physics principles were applied to attain each of the main goals stated in the target problem.

For each critical node, we further determined whether the substantive contributions to solving that node (segmented into a "statement" size units) were made primarily by the tutor, the tutee, or jointly by both (the term "substantive" refers to content-relevant contributions). Jointly covered meant that both the tutor and a tutee made substantive contributions to the node. Table 1.1 shows the number of statements made when a critical node was covered or contributed either by the tutor alone, the tutee alone, or jointly by both. When a node is covered only by the tutor, the tutor made, on average, 33.8 statements, and when it was covered primarily by a tutee, the tutee made on average, 13.6 statements. This difference makes sense since the tutor obviously could say more about how to solve a critical node.

Taylor & Francis Not for Distribution

TABLE 1.1 Number of Statements Contributed in Discussing a Critical Node and Correlations with Paired Observers' Deep Learning

	Frequency	Observers' Learning
Tutor Alone	33.8	N.S.
Abstract	12.8	N.S.
Grounded	21.0	N.S.
Tutor and Tutee	77.9	N.S.
Abstract	14.7	N.S.
Grounded	63.2	$r = .416, p = .068$
Tutees Alone	13.6	N.S.
Abstract	2.2	N.S.
Grounded	11.3	$r = .544, p = .013$

However, we further coded each of the tutor and tutee critical node statements to see whether the contributions were more abstract or grounded. The results confirm only half of our intuitive impression, in that the tutees did tended to discuss the situation of each node in a more grounded way versus an abstract way (11.3 grounded statements vs. 2.2 abstract statements), a five-to-one difference. However, the tutor also described the critical nodes in a more grounded way than in an abstract way, although the difference was more modest (21.0 grounded vs. 12.8 abstract statements). But what is interesting is that there were no significant correlations between the tutor's descriptions (related to these nodes) and the paired observers' learning, regardless of whether the tutor described them in an abstract or grounded way. This is consistent with the overall lack of correlation found earlier (see Chi et al., 2008, Table 3) between the tutor's explanations and both the tutees and observers' learning, using either the entire set of protocols coded at a statement level or at a node level (see Chi et al., 2008, Table 6). However, when a critical node is covered either by the tutees alone or between the tutor and tutee jointly, then the frequency of such coverage correlates with observers' learning, but only when the node is discussed in a grounded way, either by the tutees alone (shown by the significant correlation between tutees' grounded discussion and observers' learning, $r = .544$, $p = .013$), or approaches significance when a node is discussed jointly by the tutor and tutees in a grounded way ($r = .416$, $p = .068$). This pattern of correlations confirms our prior analyses showing that observers paid more attention to what the tutees said over what the tutor said, and supports our conjecture that observers may have preferred to pay attention to the tutees' statements because they were more grounded, thereby making them more understandable (see Table 1.1).

Learning the Skills of How-to-Interact-with-an-Expert to Learn The fourth possible reason for why paired observers might learn by paying more attention to tutees than a tutor (in a tutorial dialogue tape) is the conjecture

that an observer might be acquiring *how-to-interact-with-an-expert* learning skills from the tutees. Although a great deal of research (Chi, 2009; Webb & Palincsar, 1996), along with practices in a classroom setting, have focused on student learning when interacting with another peer, little work has addressed understanding what *interacting-with-an-expert* learning skills are, and how we can optimize students' acquisition of such skills. There are increasingly more settings that require students or trainees to interact with an expert, such as other team members who have different expertise, a knowledgeable animated agent, or an older and more knowledgeable peer (Roscoe & Chi, 2007).

We conceive of *interacting-with-an-expert* learning skills as a form of 21st-Century skills that are different from *interacting* skills *per se* in a collaborative learning sense. *Interacting-with-an-expert* learning skills are also different from the burgeoning literature on team interactions and coordination. That literature, by and large, focuses more on understanding team communication, such as effective communication strategies for calling attention to problems and getting action and attention from the other team members (Fischer & Orasanu, 2000), or effective communication related to sharing critical information (Fisher, McDonnell, & Orasanu, 2007). The ideas of *interacting-with-an-expert* learning skills proposed here are also different from the traditional learning-to-learn skills introduced by Ann Brown (Brown, Campione, & Day, 1981) over two decades ago. Much of the early learning-to-learn skills involved monitoring one's own understanding, such as acknowledging that one does not understand, gauging that one needs to spend more time on a specific concept, and so forth. In short, they involve self-reflection and self-regulation skills (Butler & Winne, 1995; Schraw, Crippen & Hartley, 2006; Zimmerman, 2000).

What might be considered *interacting-with-an-expert* learning skills? Several examples can be considered productive *interacting-with-an-expert* learning skills, such as asking questions, initiating new topics for discussion, and so forth. Such tutee behaviors, if exhibited, would demonstrate to observers that the tutees are actively taking control of their own learning, and that they are not merely responding to a tutor's leads and scaffoldings. Essentially such behaviors would illustrate that the tutees are equal partners in the tutorial dialogues. We hypothesize that Good-learner tutees (previously defined as tutees who were more successful at learning on their own prior to tutoring) would engage in these types of *interacting-with-an-expert* learning skills more often than Poor-learner tutees, thus enabling observing-students to learn more from the Good-learner tutees' dialogue.

To test this hypothesis, we re-examined our tutorial dialogues (Chi et al., 2008). We again focused strictly on the critical nodes, but this time we focused on those nodes that were jointly covered, since we were interested in why the paired observers paid more attention to the tutees, it seemed only fair to focus on nodes for which there was joint coverage, so that we could understand why there was a bias towards attending to the tutees. We coded the frequency of tutees' overall initiated comments, questions and meta-comments when

tutees jointly covered a critical node with the tutor. On average, among the 27.8 critical nodes that were jointly covered, 20.8 of them had comments initiated by the tutees, which were comments not elicited by the tutor. This is in stark contrast to the number of questions that students typically initiate in class, which is 0.11 per hour (Graesser, Person, & Magliano, 1995). Although we know in general that tutorial dialogues often involve tutees *responding* to a tutor's elicitations, such as answering a tutor's questions and scaffolding hints, as shown by the equivalent number of tutor requests (94) and tutee responses (104) described above (from data reported in Chi et al., 2001), what we did not know is that the tutees also do *initiate* many comments, questions, and meta-comments. Furthermore, across all the critical nodes, the Good-learner tutees initiated comments in 27.5% of the nodes on average, compared to 6.43% for the Poor-learner tutees, and the difference was significant ($F(1,7) = 29.851$, $p = .011$). This contrast confirms our interpretation that perhaps the paired observers learn more from the Good-learner tutees' dialogues because those dialogues illustrate *interacting-with-an-expert* learning skill (such as initiating) more often than the dialogues with Poor-learner tutees.

Another *interacting-with-an-expert* learning skill might relate to a tutee making joint contributions with the Tutor, as explained above. A tutee contributing jointly with the Tutor implies that the tutee has the confidence to assume that what s/he has to say is important and meaningful, whereas students in a classroom may often be timid and more reluctant to make contributions. Although in a tutoring context, the majority of the critical nodes were covered jointly by both the tutor and the tutees (consistent with the statement data shown in Table 1.1 above, indicating that 77.9 statements were made jointly in the context of covering critical nodes, vs. 33.8 statements for the tutor's primary coverage and 13.55 for the tutees' primary coverage), nevertheless, for dialogues with the Good tutees, 90% of the critical nodes were covered jointly, whereas for dialogues with the Poor tutees, only 50% of the critical nodes were covered jointly. This difference provides some support that Good-learner tutees were more skillful at making joint contributions.

Behaviors such as initiating new comments and jointly contributing to solving a critical node imply that a tutee is treating himself/herself as an equal to the tutor, and that the tutee can take control of the learning task. If our interpretation that observing-students acquire *interacting-with-an-expert* learning skills is correct, then there ought to be a correlation between the display of such behaviors from the tutees and the observers' behaviors. We did, in fact, find a very strong correlation ($r = .662$, $p = .003$) between the number of joint explanations that tutees made with the tutor, and the number of joint explanations paired observers made with each other. Such correlations further support our hypothesis that paired observers may have acquired *interacting-with-an-expert* learning skills. In sum, this fourth conjecture postulates that by paying more attention to the tutees rather than the tutor, the observing-students might have benefitted from observing displays of these *how-to-interact-with-an-expert* learning skills.

Summary of How Paired Observers Might Have Learned

We have provided preliminary evidence supporting the hypothesis that observers of tutees learned to solve physics problems in ways different from the tutees themselves. Observers might learn by paying attention to what the tutees said because what they said was more grounded and understandable as compared to the tutor. Doing so may mean that observers might be explaining the tutees' errors, and may be picking up other *interacting-with-an-expert* learning skills, such as taking initiative and making joint explanations. In addition, observers might understand a tutee's initiations because they have a closer zone of proximal representational match with a tutee than a tutor. Thus, interacting with a partner and observing tutoring, for various reasons, can compensate for the lack of opportunities to interact directly with a tutor.

CONCLUSION

In this chapter, we address the broad question of whether students can learn from experts or teachers/mentors without directly interacting with them, since experts are a costly resource with regard to one-on-one instruction. We analyzed three ways that experts could impart their knowledge to novices without directly interacting with them: (1) by demonstrating their expertise in terms of overtly performing the task(s) for which they are skilled; (2) by accompanying their demonstrations with monologue explanations; or (3) by guiding a novice's problem-solving through dialogues and letting other novices observe. The first method (demonstrating expertise from overt performance only—such as a math professor standing at the board and writing out the solution steps) is not an effective educational practice, and moreover, difficult academic tasks cannot be easily demonstrated overtly. So we will ignore this method as a viable approach to academic instruction. Educational practices often involve the second method, which is for experts to demonstrate while explaining. Although this is a common practice, its effectiveness is limited. The third way is to have an expert (e.g. a tutor) guide and interact with a novice (e.g. a tutee) while students observe. Our finding that the paired observing-students were able to learn as well as the tutees, even though they were not directly interacting with a tutor but they could interact with a peer partner, suggests a way of multiplying the benefits of investments in tutoring (which we know is extremely effective for the students being tutored; however, one-on-one tutoring is expensive and labor-intensive).

How did paired observers learn from watching and listening-in to tutoring dialogues? We postulated that the observing-students must have learned in a way that was different from the tutees, since the observing-students could not interact with the tutor. Although there was also likely some benefit of interacting with each other, we confirmed this difference by showing that the observers paid more attention to what the tutees said and did (especially the Good-learner tutees), rather than what the tutor said and did. We hypothesized several reasons for why observers had a biased preference to attend to the tutees'

comments over the tutor's comments. One reason is that the observers, being novices, might have had representations that more closely matched those of the tutees (who are also novices) than the tutor's. Such a closer zone of proximal representational match would have made the tutees' comments and questions more understandable to the observers, as compared to the tutor's normative comments and explanations. Related to this mismatch idea, tutees' comments are also most often more grounded than a tutor's comments, and therefore, observers might be able to understand tutees' comments better. Third, there is evidence in the literature that by explaining tutees' errors, the observers might learn. Since tutees often make errors, the observers might be intrigued or challenged to explain and resolve the tutees' errors, and such effort might have caused the observers to learn. Fourth, observers might be able to acquire *interacting-with-an-expert* learning skills from tutees, since tutees tend to display these skills when interacting with a tutor. One interacting-learning skill might be in initiating new comments and ideas, and this contrasts the typical tutorial dialogue pattern in which the tutor initiates and the tutee responds. In sum, these reasons might account for our most counterintuitive finding that observers are biased to attend to the tutees more than to the tutor.

In summary, this chapter suggests that a teacher/mentor's instructional expertise can be amplified for other student observers by having the expert interact with a novice (tutee). Aside from the advantages of learning, watching a video of someone else learn also has scale-up economic advantages, such as amortizing the expense of one-to-one tutoring over a multitude of learners, as with the concept of re-usable instructional dialogues (McKendree et al., 1998; Stenning et al., 2000). This suggests that creating instructional videos of an expert tutoring a novice student might have exciting applications in distributed and networked environments, and in online discussion boards, in which direct discussion with an expert is limited.

REFERENCES

Anderson, C. A., & Bushman, B. J. (2001). Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review of the scientific literature. *Psychological Science, 12*(5), 353–359.

Bandura, A. (1989). Social cognitive theory. In R. Vasta (Ed.), *Annals of child development. Vol.6. Six theories of child development* (pp. 1–60). Greenwich, CT: JAI Press.

Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *The Journal of Abnormal and Social Psychology, 63*(3), 575–582.

Ben-Zeev, T. (1995). The nature and origin of rational errors in arithmetic thinking: Induction from examples and prior knowledge. *Cognitive Science, 19*(3), 341–376.

Bromme, R., Jucks, R., & Runde, A. (2005). Barriers and biases in computer-mediated expert-layperson-communication. In R. Bromme, F.W. Hesse, & H. Spada (Eds.), *Barriers, biases and opportunities of communication and cooperation with computers and how they may be overcome* (pp. 89–118). New York: Springer.

Brown, A. L., Campione, J. C., & Day, J. D. (1981). Learning to learn: On training students to learn from texts. *Educational Researcher, 10*(2), 14–21.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245–281.

Carroll, W. R., & Bandura, A. (1990). Representational guidance of action production in observational learning: a causal analysis. *Journal of Motor Behavior, 22*(1), 85–97.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*, 73–105.

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145–182.

Chi, M. T. H., & Bjork, R. A. (1991). Modeling expertise. In R. Druckman, & A. Bjork (Eds.), *In the mind's eye: Enhancing human performance* (pp. 57–79). Washington, DC: National Academy Press.

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

Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121–152.

Chi, M. T. H., Roy, M. & Hausmann, R. G. M. (2008). Observing tutoring collaboratively: Insights about tutoring effectiveness from vicarious learning. *Cognitive Science, 32*(2), 301–341.

Chi, M. T. H., Siler, S. A., & Jeong, H. (2004). Can tutors monitor students' understanding accurately? *Cognition and Instruction*, 22(3), 363–387.

Chi, M. T. H., Siler, S., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from tutoring. *Cognitive Science*, 25, 471–533.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. In J. F. Le Ny, & W. Kintsch (Eds.), *Language and comprehension*. Amsterdam: North-Holland.

Craig, S., Chi, M. T. H., & VanLehn, K. A. (2009). Improving classroom learning by collaboratively observing human tutoring videos while problem solving. *Journal of Educational Psychology, 101*(4), 779–789.

Craig, S. D., Driscoll, D. M., & Gholson, B. (2004). Constructing knowledge from dialog in an intelligent tutoring system: Interactive learning, vicarious learning, and pedagogical agents. *Journal of Educational Multimedia and Hypermedia, 13*, 163–183.

Craig, S. D., Gholson, B., Ventura, M., Graesser, A. C., & Tutoring Research Group. (2000). Overhearing dialogues and monologues in virtual tutoring sessions: Effects on questioning and vicarious learning. *International Journal of Artificial Intelligence in Education, 11*, 242–253.

Davidson, E. S., & Smith, W. P. (1982). Imitation, social comparison, and self-reward. *Child Development, 53*, 928–932.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improving learning in a large-enrollment physics class. *Science*, 332, 861–864.

Dreyfus, A., Jungwirth, E., & Eliovitch, R. (1990). Applying the "cognitive conflict" strategy for conceptual change: Some implications, difficulties, and problems. *Science Education, 74*, 555–569.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Feldman, A. I., Campbell, R. L., & Lai, M. K. (1999). Improving elementary school science teaching by cross-level mentoring. *Journal of Science Teacher Education, 10*(1), 55–67.

Field, F. (1983–1990/1991–1995). *Mr. Wizard's World*. Calgary: Nickelodeon

Fischer, U., McDonnell, L., & Orasanu, J. (2007). Linguistic correlates of team performance: Toward a tool for monitoring team functioning during space missions. *Aviation, Space, and Environmental Medicine, 78*, B86–B95.

Fischer, U., & Orasanu, J. (2000). Error-challenging strategies: Their role in preventing and correcting errors. In Proceedings of the International Ergonomics Association 14th Triennial Congress and Human Factors and Ergonomics Society, 44th Annual Meeting San Diego, CA, August.

Goffman, E. (1976). Replies and responses. *Language in Society, 5*, 257–313.

Graesser, A. C., Person, N., & Magliano, J. (1995). Collaborative dialog patterns in naturalistic one-on-one tutoring. *Applied Cognitive Psychology, 9*, 359–387.

Haskett, G. J., & Lenfestey, W. (1974). Reading-related behavior in an open classroom: Effects of novelty and modeling on preschoolers. *Journal of Applied Behavior Analysis, 7*(2), 233–241.

Hinds, P. J., Patterson, M., & Pfeffer, J. (2001). Bothered by abstraction: The effect of expertise on knowledge transfer and subsequent novice performance. *Journal of Applied Psychology, 86*, 1232–1243.

Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*, 379–424.

Lave, J. (1988). *Cognition in practice*. New York: Cambridge University Press.

Linn, M. C. (1995). Designing computer learning environments for engineering and computer science: The scaffolded knowledge integration framework. *Journal of Science Education and Technology, 4*(2), 103.

Mayes, J. T., Dineen, F., McKendree, U., & Lee, J. (2001). Learning from watching others learn. In C. Steeples & C. Jones (Eds.), *Networked learning: Perspectives and issues* (pp.1–16). London: Springer.

McKendree, J., Stenning, K., Mayes, T., Lee, J., & Cox, R. (1998). Why observing a dialogue may benefit learning: The vicarious learner. *Journal of Computer Assisted Learning, 14*, 110–119.

Monaghan, P., & Stenning, K. (1998). Learning to solve syllogisms by watching others' learning. Research Paper HCRC/RP-98. Human Communication Research Centre, University of Edinburgh.

Muller, D. A., Sharma, M. D., & Reimann, P. (2008). Raising cognitive load with linear multimedia to promote conceptual change. *Science Education, 92*(2), 278–296.

Nathan, M. J., & Koedinger, K. R. (2000). An investigation of teachers' beliefs of students' algebra development. *Cognition and Instruction, 18*, 207–235.

Nonnecke, B., & Preece, J. (2001). Why lurkers lurk. American Conference on *Information Systems*, 1–10.

Piaget, J. (1952). *The child's concept of number*. New York: W.W. Norton.

Reed, S.K., Dempster, A., & Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. *Journal of Experimental Psychology: Learning, Memory and Cognition, 11*, 106–125.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in sociocultural activity*. New York: Oxford University Press.

Rogoff, B., Paradise, R., Arauz, R. M., Correa-Chavez, M., & Angelillo, C. (2003). Firsthand learning though intent participation. *Annual Review of Psychology, 54*, 175–203.

Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, *77*, 534–574.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *13*, 629–639.

Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *The Journal of the Learning Sciences*, *14*, 201–241.

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

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective in learning. *Research in Science Education*, *36*(1), 111–139.

Schunk, D. H., Hanson, A. R., & Cox, P. D. (1987). Peer-model attributes and children's achievement behaviors. *Journal of Educational Psychology*, *79*, 54–61.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction*, *14*, 293–305.

Shebilske, W. L., Jordan, J. A., Goettl, B. P., & Paulus, L. E. (1998). Observation versus hands-on practice of complex skills in dyadic, triadic, and tetradic training-teams. *Human Factors*, *40*, 525–540.

Siegler, R. S. (2002). Microgenetic studies of self-explanation. In N. Granott, & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 31–58). Cambridge: Cambridge University Press.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 592–604.

Stein, B. S., & Bransford, J. D. (1979). Constraints on effective elaboration: Effects of precision and subject generation. *Journal of Verbal Learning and Verbal Behavior*, *18*, 769–777.

Stenning, K., McKendree, J., Lee, J., Cox, R., Dineen, F. & Mayes, T. (2000). Various learning from educational dialogue. *Proceedings of Computer-Supported Cooperative Learning (CSCL, '99)*, 341–347.

Sutton, L. A. (2001). The principles of vicarious interaction in computer-mediated communications. *International Journal of Educational Telecommunication*, *7*, 223–242.

Tree, J. E. F. (1999). Listening in on monologues and dialogues. *Discourse Processes*, *27*(1), 35.

VanLehn, K. (1988). Toward a theory of impasse-driven learning. In H. Mandl, & A. Lesgold (Eds.), *Learning issues for intelligent tutoring systems* (pp. 19–41). New York, NY: Springer.

VanLehn, K., Graesser, A., Jackson, G., Jordan, P., Olney, A., & Rose, C. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, *31*, 3–62.

Webb, N. M. (1989). Peer instruction, problem-solving, and cognition: Multidisciplinary perspectives. *International Journal of Educational Research*, *13*, 21–39.

Webb, N. M., & Palincsar, A. S. (1996).Group processes in the classroom. In D. C. Berliner, & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 841–873). New York, NY: Prentice Hall.

Whiten, A. (1998). Imitation of the sequential structure of actions by chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *112*, 270–281.

Zimmerman, B. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25*, 82–91.

T&F Proofs, Not for Distribution