COGNITION AND INSTRUCTION, *16*(4), 475–522

A Time for Telling

Daniel L. Schwartz and John D. Bransford
Learning Technology Center
Vanderbilt University

Suggestions for improving text understanding often prescribe activating prior knowledge, a prescription that may be problematic if students do not have the relevant prior knowledge to begin with. In this article, we describe research about a method for developing prior knowledge that prepares students to learn from a text or lecture. We propose that analyzing contrasting cases can help learners generate the differentiated knowledge structures that enable them to understand a text deeply. Noticing the distinctions between contrasting cases creates a "time for telling"; learners are prepared to be told the significance of the distinctions they have discovered. In 3 classroom studies, college students analyzed contrasting cases that consisted of simplified experimental designs and data from classic psychology experiments. They then received a lecture or text on the psychological phenomena highlighted in the experiments. Approximately 1 week later, the students predicted outcomes for a hypothetical experiment that could be interpreted in light of the concepts they had studied. Generating the distinctions between contrasting cases and then reading a text or hearing a lecture led to more accurate predictions than the control treatments of (a) reading about the distinctions between the cases and hearing a lecture, (b) summarizing a relevant text and hearing a lecture, and (c) analyzing the contrasting cases twice without receiving a lecture. We argue that analyzing the contrasting cases increased students' abilities to discern specific features that differentiated classes of psychological phenomena, much as a botanist can distinguish subspecies of a given flower. This differentiated knowledge prepared the students to understand deeply an explanation of the relevant psychological principles when it was presented to them. These results can inform constructivist models of instruction as they apply to classroom activities and learning from verbal materials. In particular, the results indicate that there is a place for lectures and readings in the classroom if students have sufficiently differentiated domain knowledge to use the expository materials in a generative manner.

Requests for reprints should be sent to Daniel L. Schwartz, Learning Technology Center, Box 45–GPC, Vanderbilt University, Nashville, TN 37203. E-mail: dan.schwartz@vanderbilt.edu

The goal of this article is to begin a theoretical and empirical exploration of when to use texts, lectures, and explanations within the total repertoire of instructional methods. These experiments demonstrate that, when students have had an opportunity to generate well-differentiated knowledge about a domain, then teaching through a lecture or text can be an extremely effective form of instruction. This research is important because it focuses explicitly on the degree to which there are points of knowledge development that are indicative of a "time for telling" or a "readiness" for being told something.

Issues relevant to a time for telling are especially important in the context of *constructivist* models that emphasize the active construction of knowledge by learners. Constructivist models are often compared to *transmission* models that assume that students acquire knowledge by having it transmitted to them by a teacher or a text (e.g., Brown, 1992; Cognition and Technology Group at Vanderbilt [CTGV], 1996; Scardamalia & Bereiter, 1991). There are many areas of confusion surrounding constructivist theories and their implications for instruction. A major source of confusion involves assumptions about relations between constructivist theories of knowing and the implications of those theories for instruction. These confusions can be clarified by considering the following continuum:

Total Student Control ↔ Total Teacher Control

Some people assume that constructivist approaches refer primarily to the far left of the continuum and, hence, demand a "discovery learning" approach to instruction in which students explore new domains without teacher guidance. Most constructivists explicitly reject this view and, instead, argue that an emphasis on unconstrained discovery is not a necessary implication of constructivist theorizing. A number of studies show that "guided discovery" and "scaffolded inquiry" are much more effective for learning than unconstrained discovery (e.g., Brown & Campione, 1994; CTGV, 1996; Gagné & Brown, 1961; Littlefield et al., 1988).

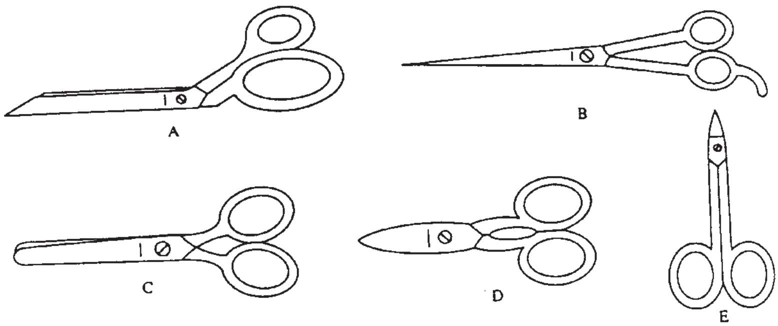
Less extreme (and more prevalent) than the unaided discovery view is the assumption that constructivist models refer to the left half of the continuum and that transmission models refer to the right. In our view, this too is a misconception. Constructivism refers to a theory of knowledge growth that operates whether one is actively exploring or whether one is sitting still and listening to a lecture or reading a book (e.g., Cobb, 1994). The question for constructivists focuses on the kinds of activities needed to help people best construct new knowledge for themselves. Often, the act of listening to a lecture or reading a text is not the best way to help students construct new knowledge. At other times, this may be exactly what students need.

Elsewhere (CTGV, 1997), we conjectured that there are times for telling. One occurs when students enter a learning situation with a wealth of background knowledge and a clear sense of the problems for which they seek solutions. Consider, for example, a clinic for football coaches. Imagine that the most victorious coach of the past decade stands before the group and tells his secrets of victory. Does this make him anticonstructivist? We argued that this method of teaching is appropriate given the level of understanding that the audience brings to the situation. Because of their extensive experiences, the coaches can easily grasp the relevance of what the speaker has to say. The speech can be meaningful because it maps into the knowledge of problem situations that the coaches have already developed.

In many educational settings, however, there is an absence of features that are present at a coaches' clinic. Students often have not had the opportunity to experience the types of problems that are rendered solvable by the knowledge we teach them. Under these conditions, we conjecture that telling is not the optimal way to help students construct new knowledge. When telling occurs without readiness, the primary recourse for students is to treat the new information as ends to be memorized rather than as tools to help them perceive and think.

One common procedure for preparing students for telling is to help them activate relevant prior knowledge before they read a text or listen to a lecture (e.g., Ausubel, 1968; Beck, 1984; Karplus, 1981). There are numerous demonstrations of and theories about the effects of prior knowledge on comprehension and memory (e.g., Bransford & Johnson, 1972; Britton & Graesser, 1996; Dooling & Lachman, 1971; Frase, 1975; Kintsch et al., 1993; McNamara, Kintsch, Songer, & Kintsch, 1996; Voss, Vesonder, & Spilich, 1980). However, the attempt to help students activate prior knowledge presupposes that they have already acquired the relevant prior knowledge in the first place. What if they have not? One approach is to tell them the needed knowledge. This procedure involves "creating a time for telling by doing more telling."

One reason why "more telling" may be ineffective is that many texts and lectures presuppose a level of differentiated knowledge that is not available to novices. Under these conditions, novices can easily think they understand when, in reality, they have missed important distinctions (e.g., Bransford & Nitsch, 1978). As a simple example, consider the following statement: "The dressmaker used the scissors to cut the cloth for the dress." This statement is understandable to most people. If asked to generate an image and elaborate on it, they can easily imagine a person using a pair of scissors to cut some cloth, and they can explain to themselves why cutting might be important for dressmaking. However, what will their concept of the dressmaker's scissors be like? A scissors expert will have a more finely differentiated concept of scissors than most casual comprehenders. Experts understand the many possible features of scissors adapted for different purposes (see Figure 1). Compared to novices, for example, experts would be able to describe the structure of the scissors used by a dressmaker, and they would understand the significance of features such as a flat cutting edge and its advantages. If someone asked an expert to purchase a new pair of scissors, the expert would likely ask questions to clarify the exact type of scissors needed. A novice's knowl-



Structure	Function
a. Dressmaker's shears	
Heavy	Because of heavy use
One hole larger than other	Two or three fingers will fit in larger hole—allows greater steadiness as one cuts cloth on flat surface
Blades off-center and aligned with finger-hole edge	Blade can rest on table surface as cloth is cut—again, greater steadiness
b. Barber's shears	
Very sharp	To cut thin material, for example, hair
Pointed	Permits blades to snip close to scalp and to snip very small strands of hair
Hook on finger hole	A rest for one finger, which allows scissors to be supported when held at various angles—hence, greater maneuverability
c. Pocket or children's scissors	
Blunt ends	Scissors can be carried in pocket without cutting through cloth; children can handle without poking themselves or others
Short blades	Allow greater control by the gross motor movements of the child just learning to cut
d. Nail scissors	
Wide and thick at pivot point	To withstand pressure from cutting thick and rigid materials, that is, nails
Slightly curved blades	To cut slightly curved nails
e. Cuticle scissors	
Very sharp blade	To cut semielastic materials, for example, skin of cuticles
Small, curved blades	To allow maneuverability necessary to cut small curved area
Long extension from finger holes to joint	As compensation for short blades, necessary for holding.

FIGURE 1 Differentiated knowledge structures help appreciate specific properties. From "A Sketch of a Cognitive Approach to Comprehension," by J. D. Bransford and N. S. McCarrell, in *Cognition and the Symbolic Processes,* edited by W. Weimer and D. S. Palermo, 1974, Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. Copyright 1974 by Lawrence Erlbaum Associates, Inc. Adapted with permission.

edge, however, would be less differentiated; hence, the novice would be less likely to imagine and elaborate correctly and less likely to generate questions that would clarify which scissors to purchase (for examples of *semantic flexibility* in comprehension as a function of context and expertise, see R. C. Anderson & Ortony, 1975; Barclay, Bransford, Franks, McCarrell, & Nitsch, 1974).

The importance of differentiated knowledge structures can be further illustrated by considering psychology students' attempts to understand the following statement: "The developmental psychologist showed first graders, fifth graders, and college students a set of 30 pictures and found that their memories for the pictures were equivalent." Novices can comprehend this statement at some level, but chances are that their understanding of memory will be relatively imprecise. In contrast, an expert will assume that this experiment involved recognition memory rather than free or cued recall, unless the 30 pictures were chosen to map very explicitly into a domain of organized knowledge in which the children were experts (e.g., dinosaurs, as in Chi, 1976; TV shows, as in Lindberg, 1980). In short, the expert can construct a number of well-differentiated scenarios, whereas the novice understands only superficially.

Our hypothesis is that, rather than doing more telling, a powerful way to create a time for telling comes from theories of perceptual learning that emphasize differentiation (e.g., Arnoult, 1953; Bransford, Franks, Vye, & Sherwood, 1989; Garner, 1974; E. J. Gibson, 1969; J. J. Gibson & Gibson, 1957). These theories propose that opportunities to analyze sets of contrasting cases, such as different pairs of scissors (see Figure 1), can help people become sensitive to information that they might miss otherwise (e.g., Gagné & Gibson, 1947; Garner, 1974; Gick & Paterson, 1992). Contrasting cases help people notice specific features and dimensions that make the cases distinctive. The resulting well-differentiated information provides the bases for guiding other activities such as creating images, elaborating, generating questions, and learning (Bransford & Schwartz, in press).

E. J. Gibson (1969) reported an experiment by Dibble illustrating how opportunities to explore contrasting cases can develop differentiated knowledge that supports further learning. The question addressed was whether actively contrasting letters of a novel alphabet would facilitate the subsequent learning of names for those letters. In preparation to being told the names for each letter, adults completed one of three treatments. In the reproduction treatment, participants had 12 min to copy one letter after another. In the contrasting cases treatment, participants had 10 min to contrast each letter with each other letter. In the control treatment, participants had no prior exposure to the letters. Everybody then learned a name for each letter. People in the control and reproduction treatments took the same amount of time to learn the name–letter pairings, whereas people in the contrasting cases treatment learned the associations significantly faster. By analyzing the letters as contrasting cases, they noticed distinctive features of each letter (e.g., length of stroke, and degree of curve), which in turn, prevented confusion as they paired the names with the letters.

The materials used in these studies involve conceptual rather than perceptual differentiation. In particular, the cases (described later) involve opportunities to actively differentiate among exemplars of psychological phenomena and concepts. The cases were adapted from experiments in cognitive psychology that we have previously taught in our own courses. In our experience, students tended to develop an overly superficial understanding of the experiments and concepts, limiting their ability to transfer to new situations. These experiments use the method of contrasting cases to help students develop knowledge that is more differentiated than we have been able to achieve in the past.

Table 1 summarizes the concepts that were the focus of these experiments. The concepts were taught to students in actual courses; hence, students were presumably motivated to learn effectively. The general experimental design was to compare the students' abilities to learn about the concepts from a text or lecture as a function of whether they prepared by actively contrasting specific cases, or whether they prepared by working with materials that did more telling about the cases and concepts.

As noted previously, one hypothesis guiding our work is that the active comparison of relevant contrasting cases (as opposed to simply being told about the cases and concepts) helps foster well-differentiated knowledge. These activities by themselves, however, are usually insufficient for students to induce the principles necessary to understand a domain at a satisfactorily deep level. This is because novices often lack an overriding framework that helps them develop a theory or model to explain the significance of the distinctions they have discovered. This is one place in which telling can have powerful effects on people's abilities to learn; it can help them make sense of the distinctions that they have noticed. Thus, our

TABLE 1
The Content of Instruction: The Eight Target Concepts and Their Operationalizations

Concept Type	*Operationalization*
Schema concepts	
Stereotypical recall	People remember events of high stereotypy
Script intrusions	People falsely remember events from own script
Ordered recall	People remember events in chronological sequence
Obstacle recall	People remember obstacles to goal completion
Encoding concepts	
Total recall	People remember more if material is meaningful
Primacy and recency	People remember first and last items of a stimulus list
Gist and verbatim	Meaning leads to gist recall; nonsense leads to verbatim recall
Inference intrusions	People remember inferences made while reading a passage

complete hypothesis is that opportunities to develop well-differentiated knowledge structures set the stage for learning through telling but usually do not replace it. We explore this hypothesis across three experiments by providing different groups of students with particular experiences (e.g., summarizing a text vs. analyzing contrasting cases) and then assessing whether and how well they learn from a subsequent lecture or text.

Central to our experiments is the notion of what it means to assess understanding of the materials being studied. We do not expect our experimental and comparison groups to differ in their ability to repeat factual statements about the concepts being learned (e.g., see Bransford et al., 1989; Michael, 1989). To return to the dressmaker example, we would expect both novices and experts to remember that the dressmaker used scissors to cut the cloth. More subtle assessments are needed to discriminate qualitative differences in the understanding of experts and novices. In the case of the dressmaker example, one might ask students to draw the pair of scissors being used. In the case of these experiments on psychological concepts, we use an assessment that asks students to make detailed predictions about a new experiment. The nature of this assessment will become clearer in the description of Experiment 1.

EXPERIMENT 1

The first experiment addresses whether analyzing contrasting cases, as compared to reading about summaries of those same cases, prepares students for a lecture on cognitive psychology. In a within-subject design, undergraduate students analyzed a set of contrasting cases for one group of target concepts and read about another set of cases for a second group of target concepts. Afterward, they heard a lecture that covered both groups of concepts. If analyzing contrasting cases creates a time for telling, then we should expect the students to learn more from the portion of the lecture that complements the cases they analyzed than from the portion of the lecture that complements the cases that they read about. Next, we describe our instructional manipulations and learning materials more thoroughly. We then explain how we assessed student learning.

The Instructional Manipulations

The instruction revolved around the eight target concepts shown in Table 1. For experimental purposes, the concepts were separated into two groups that we label *schema concepts* and *encoding concepts.* The schema concepts primarily concern the effects of schemas on the organization of and retrieval from long-term memory. The encoding concepts primarily concern encoding from short-term to long-term memory.

Concepts and cases. For the analysis activity, we developed two sets of contrasting cases, one for the schema concepts and one for the encoding concepts. For the schema concepts, the contrasting cases were drawn from studies on mental scripts (e.g., Abelson, 1981). For the encoding concepts, the contrasting cases were drawn from studies on the effects of a meaningful context on memory. To help distinguish the two sets of contrasting cases from the two clusters of target concepts that the cases exemplify, we give more specific labels to the contrasting cases. We label the cases for the schema concepts as the *doctor visit* because the experiments and data sets that create the contrasts were adapted from a set of studies that examined people's memories for a doctor visit (Bower, Black, & Turner, 1979). We label the cases for the encoding concepts as the *balloon passage* because the materials were adapted from a set of studies that examined people's memory for a text that was presented with and without a picture that included balloons (Bransford & Johnson, 1972).

Analyses of contrasting cases. One of the studies and data sets included in the doctor visit cases is illustrated in Table 2 (the Appendix includes the complete doctor visit and balloon passage contrasting cases). These cases involve examples of person-by-person data rather than overall summaries of these data. When students analyzed the cases, their task was to find and graph what they thought were the revealing patterns for each study. We asked students to graph the patterns because we have found that graphing encourages students to look for contrasts at the level of overall patterns rather than at the level of the "little contrasts" between one data point and another.

By analyzing the cases, students should differentiate empirical phenomena relevant to the target concepts. For example, students analyzing the study shown in Table 2 should notice that some doctor visit events are mentioned by more participants than others (e.g., "read magazine" vs. "following nurse"). This phenomenon is relevant to the target concept of stereotypical memory (see Table 1). The additional doctor visit materials (see the Appendix) provide further contrasts that bear on the target concepts. The third study, for example, shows that participants remember nonstereotypical events about a patient's doctor visit at a very low rate, except one event: The patient forgot his wallet. Students should notice the contrast between memory for this event and the other nonstereotypical events. Of course, they may not have an explanation for this brute fact; namely, people remember obstacles to goal completion. This is one reason why we believe that some form of telling following the case analyses is important; it provides explanations that students are unlikely to develop on their own.

Reading about cases. To compare the effects of analyzing contrasting cases to the effects of reading about those same cases, we created *read-only* ver-

TABLE 2
The First of the Doctor-Visit Contrasting Cases That the Students Analyzed

In the first experiment, researchers asked 6 participants to write down the events that occur when they visit the doctor. The results are shown here. Notice that the participants' prose has been reduced to each main idea each participant came up with. For example, if a participant wrote, "You enter the doctor's office," and another participant wrote, "You walk through the door of the doctor's suite," both would be simplified to "Enter office." Although this simplification may miss some subtle details and differences in the participants' protocols, it makes it much easier to compare ideas across different participants.

Participant 1:	Enter office. Check in with receptionist. Sit down. Wait. Name called. Enter exam room. Sit on table. Doctor examines. Doctor asks questions. Make another appointment. Leave office.
Participant 2:	Check in with receptionist. Read magazine. Look at other people. Name called. Sit on table. Nurse tests. Doctor examines. Leave office.
Participant 3:	Check in with receptionist. Sit down. Read magazine. Talk to nurse. Nurse tests. Talk to doctor about problem. Leave office.
Participant 4:	Enter office. Sit down. Read magazine. Enter exam room. Undress. Sit on table. Nurse tests. Doctor examines. Get dressed. Leave office.
Participant 5:	Enter office. Check in with receptionist. Sit down. Read magazine. Name called. Follow nurse. Enter exam room. Nurse tests. Doctor enters. Doctor examines.
Participant 6:	Check in with receptionist. Sit down. Read magazine. Nurse tests. Wait. Doctor greets. Doctor examines. Get medicine. Leave office.

Note. Students analyzed the data and made a visualization of what they thought were the important patterns (the complete doctor visit and balloon passage cases may be found in the Appendix).

sions of the doctor visit and balloon passage cases. This way, the students could analyze one set of cases (e.g., doctor visit) and read about the other set of cases (e.g., balloon passage). The read-only materials were similar to the contrasting cases, except that the students did not receive the raw data. Instead, the read-only materials included graphs and descriptions of the important patterns and distinctions. Table 3 shows the read-only equivalent of Table 2. With these materials, the students read about the important features instead of actively discovering them.

Different treatments. As a homework assignment, one half of the students analyzed the doctor visit cases and read about the balloon passage cases. The other half of the students analyzed the balloon passage cases and read about the doctor visit cases. After the students completed the assignments, an instructor delivered an in-class lecture to both groups of students simultaneously. The instructor described the experiments included in the cases as well as a few others, showed graphs that summarized the relevant data patterns, and explained the meaning of the experiments in terms of their implications for human behavior in various domains (e.g.,

TABLE 3
The Read-Only Version of the Doctor Visit Case Shown in Table 2

In the first study, the researchers asked 6 participants to write down the events that occur when they visit the doctor. The researchers reduced a participant's prose to each main idea the participant came up with. For example, if a participant wrote, "You enter the doctor's office," and another participant wrote, "You walk through the door of the doctor's suite," both would be simplified to "Enter office." Although this simplification may miss some subtle details and differences in the participants' protocols, it makes it much easier to compare ideas across different participants. The results are shown here. The capitalized phrases were mentioned by 5 participants, the italicized phrases were mentioned by 3 participants, and the plain text phrases were mentioned by 1 participant.
Enter office
CHECK IN WITH RECEPTIONIST
SIT DOWN
Wait
Look at people
READ MAGAZINE
Name called
Follow nurse
Enter exam room
Undress
Sit on table
Talk to nurse
NURSE TESTS
Wait
Doctor enters
Doctor greets
Talk to doctor about problem
Doctor asks questions
DOCTOR EXAMINES
Get dressed
Get medicine
Make another appointment
LEAVE OFFICE

Note. Adapted from Bower, Black, & Turner (1979).

prejudice, studying for a test, etc.). The instructor also described theories (e.g., schema theory, short-term memory, etc.) and metaphors (e.g., filing cabinets, a register with seven slots, etc.) that can account for the patterns found in the experiments.

Assessing Levels of Understanding

To assess deep understanding, 1 week after the lecture, the students wrote their predictions about the outcomes of a hypothetical study adapted from Bransford and

Johnson (1972). The prediction task, as it appeared to the students, is shown in Table 4. One reason we constructed this particular hypothetical study is so that we could compare the number of schema and encoding concepts that students used to help predict the outcomes. Table 5 shows how the eight target concepts can be applied to the prediction task. The quotes are from participants in a pilot study.

If actively analyzing the cases prepares students to understand the lecture deeply, then there should be a crossover interaction on the prediction task. Students who analyze the doctor visit cases should make more predictions about schema effects than students who read about the doctor visit studies. A prediction

TABLE 4
A Copy of the Prediction Task Given to Assess Student Understanding of the Target Concepts

Below you will find the description of an experiment. Your task is to make predictions about the participants' patterns of recall in each condition. You should also predict how the recall patterns in the two conditions will differ or be similar. Make your predictions specific enough that it will be possible to see if you were right or wrong. If it helps to clarify your predictions, give an example of what a subject might recall. If possible, give reasons for your predictions. Make as many predictions as you think will really happen.

Twenty people will participate in the following study. They will read five numbered passages. Each passage is on a separate page. The passage of experimental interest is the third passage about washing clothes. Here is the passage.

PASSAGE 3

The procedure was actually quite simple. First he collected all the items into one group. He might have had to use another place due to lack of facilities. But the usual facility was going to be enough. He arranged the items into different groups. Of course one pile might have been sufficient depending on how much he had to do. It is better to do too few things at once than too many. In the short run this may not seem important but expensive complications can easily arise. Red problems are the worst. He combined each group with the usual brand. At first, the whole procedure had seemed complicated. However, it had become just another facet of his life. He cannot foresee any end to the necessity for this task in the immediate future. But then, one can never tell. After the first procedures were completed, he moved all the groups and used the usual setting. When he arranged the finished materials into new groups, he was careful. He did not want to put an extra wrinkle into his work. He would just have more work later on. Finally, he put them into their appropriate places. Eventually they will be used once more and the whole cycle will then have to be repeated. However, that is part of life.

Ten of the individuals will read the passage just as it is shown. This is the *no-title* condition. The other ten individuals will read the same passage but with the title, "Washing Clothes," next to the passage number. This is the *title* condition. After the participants read all five passages, they will be asked to write down everything they can recall from Passage 3.

Please place your predictions on the rest of this page. You may use the back of the sheet if necessary.

Note. See text for a description of how the task was administered and scored.

based on schema effects relies on one of the four target concepts covered in the doctor visit cases and lecture. Reciprocally, students who analyze the balloon passage cases should make more predictions about encoding effects than students who read about the balloon passage studies.

One reason we used a prediction task is that the ability to predict the outcome from an experiment is an ecologically valid test of knowledge; this is what practicing research psychologists do. If students gain a deep, expert-like understanding of the target concepts, then they should make relevant predictions. A second reason is that it enables us to differentiate deep and shallow knowledge. Our claim is not that novices can only learn from a lecture if they have first analyzed a set of contrasting cases. Assuming that a lecture is complete and the students attentive, novices who do not analyze contrasting cases should still learn about the target concepts. It is just that their understanding may not have the expert-like differentiation that comes from analyzing contrasting cases (Michael, Klee, Bransford, & Warren, 1993). To evaluate different levels of understanding, we complemented the prediction task with the verification task shown in Table 6. Students answered eight true–false items that covered the target concepts.

Our prediction was that students who did not analyze the appropriate cases could still answer the verification items. Their understanding should be sufficient for situations in which they do not need to notice distinctive features to bring the appropriate concept to mind. In the Discussion section, we develop our rationale in more detail. For now, one may view the verification task as serving a methodological and a pragmatic purpose. Methodologically, if the students do well on the verification task, this means that we should not attribute weak performances on the prediction task to a general lack of attention to the lecture. Pragmatically, if students do well on the verification task but not the prediction task, it will show that some of the ways that we typically assess student learning (e.g., true–false and multiple-choice questions) may not capture deeper levels of understanding. In-

TABLE 5
Examples of How Students Apply Target Concepts to the Prediction Task

Concept Labels	*Examples of Concepts as Applied to Prediction Task*
Schema concepts	
Stereotypical recall	“Subjects in title condition best remember usual steps in washing clothes”
Script intrusions	“Titled passage: Incorrectly put things they often do washing clothes”
Ordered recall	“Write down in order of steps of washing clothes in title condition”
Obstacle recall	“Remember ‘red’ sentence ’cause it ruins the point of washing clothes”
Encoding concepts	
Total recall	“People recall less for no-title version because it will be meaningless”
Primacy and recency	“The first sentence will be remembered the best”
Gist and verbatim	“The no-title subjects tended to write word for word”
Inference intrusions	“People with title will put things they figured out, like sorting *clothes*”

TABLE 6
A Copy of the Verification Test

1. When people understand a text, they will remember more than when they do not understand a text.
   True False (circle one)
2. When people understand a text, they tend to recall exact sentences from the text.
   True False (circle one)
3. When tested after reading a passage, people tend to forget the first sentence.
   True False (circle one)
4. When people understand what they are reading, they later have trouble distinguishing their inferences from the actual text.
   True False (circle one)
5. When recalling a passage about a familiar event, like going to the doctor, people remember the most stereotypical events.
   True False (circle one)
6. When recalling a passage about a familiar event, people will tend to remember the passage in the order of the steps of the event.
   True False (circle one)
7. When people recall a passage about a typical event, they usually do not add details from their own experiences with that type of event.
   True False (circle one)
8. When people recall a passage about a typical event, they remember, at a high rate, obstacles to successful event completion.
   True False (circle one)

structors may be satisfied with lectures because their assessments yield evidence of learning. However, the assessments they use may not be suited to diagnosing deep understanding.

Method

Participants. Twenty-four students from an undergraduate course at Vanderbilt University were randomly assigned to the two conditions. Three students were absent on the day of the posttest. As a result, 10 students analyzed the doctor visit cases, and 11 students analyzed the balloon passage cases. None of the students had previously studied cognitive psychology or had worked with these types of contrasting cases.

Materials. To summarize the previous descriptions, there were five types of materials:

1. Two sets of contrasting cases—the doctor visit and the balloon passage (see Table 2 and the Appendix)—both adapted from cognitive psychology experiments. Each included brief descriptions of the experimental conditions and simplified data sets. They did not include any rationale for the experiments (e.g., a hypothesis or issue at hand).

2. Read-only versions of the doctor visit and balloon passage cases that covered the same experiments as the contrasting cases, but instead of including the raw data, they included graphs and texts that summarized the results (see Table 3).

3. A 1-hr lecture that covered both the four schema and the four encoding concepts. It used the experiments from the cases as examples of concepts in action.

4. A prediction task in which students had to transfer their learning of the target concepts to make predictions about the results of a hypothetical study (see Table 4).

5. A verification task in which students had to judge whether a claim about each target concept was correct (see Table 6).

Design. The experiment employed a crossover design. A within-subject factor compared the effects of analyzing versus reading the cases. A between-subjects factor tested for effects of the particular set of target concepts. A student in the *analyze doctor visit* condition analyzed the doctor visit cases and read about the balloon passage cases. A student in the *analyze balloon passage* condition analyzed the balloon passage cases and read about the doctor visit cases. There were two dependent measures: a prediction task measuring transfer of deep understanding and a verification task measuring recognition memory.

Procedure. At the end of a regular class, the students received packets corresponding to their condition. The packets had the read-only materials first and the analysis materials second. The students were told to complete the reading assignment prior to analyzing the contrasting cases. For the reading assignment, they were instructed to read and understand the studies. For the analysis assignment, the students were directed to find and graph the significant patterns in the data. They were allowed to draw graphs by hand or computer. The students completed their packets as a homework assignment. At the next class session, 5 days later, the instructor collected the homework and delivered a lecture to all students from both conditions simultaneously. Seven days (two classes) later, as start-of-class seat work and without any mention of a connection to the previous lessons, the students completed the paper-and-pencil prediction task (approximately 15 min) and then the verification task (approximately 5 min).

Coding. To score performance on the prediction task, coders identified the use of the target concepts in the students' predictions. Table 5, drawing from a pilot study, provides a representative sample of predictions as identified by their respective concepts. The maximum possible score is an 8; there was 1 point for each target concept reflected in the predictions. To count as a prediction based on a target concept, there are three requirements:

1. A prediction has to take an operational form. One reason for this requirement is that the prediction task explicitly states that the predictions should be about "patterns of recall" and should be "specific enough to see if you were right or wrong." A second reason is that we have found that, although teachers may easily see the behavioral implication of a student's claim, the student often does not. For example, given the student claim, "The titled passage would be more meaningful," experts would easily conclude, "Therefore participants would remember more." Our experience, however, has shown us that such operationalizations are not always implicit in the claims of students. Consequently, we have chosen to be conservative and not to assign credit unless there is an explicit prediction about observable human behavior.
2. A prediction has to be unambiguously assignable to a concept. In particular, we found in pilot work that lists of likely to be remembered sentences are insufficient for determining what students have in mind. For example, one pilot student simply wrote down six sentences, one of which was the first sentence of the passage. Although the first sentence is likely to be recalled owing to primacy, it is unknown whether this is why the student included it. Consequently, when there is doubt, coders do not read "intent" into a list of sentences.
3. A prediction has to be a correct application of a target concept. In general, there were very few misapplications. One or two students in each of the following experiments incorrectly thought there would be a recency effect.

In this and the following experiments, there were two coders, one blind to condition. In each case, the coders had at least 92% overlap in the identification of the target concepts in the students' predictions. Disagreements were settled by coding conservatively with respect to the leading hypothesis that the instructional combination of analyzing contrasting cases and receiving a verbal exposition would yield the most predictions.

Results

Figure 2 displays the average number of understood concepts as measured by the prediction and verification tasks. The left-hand side of the graph shows performance on the four possible schema concepts, and the right-hand side shows performance on the four possible encoding concepts. The top of the figure shows that students in both conditions were able to verify sentences about both sets of target concepts at near ceiling levels (93% average accuracy). Consequently, one may conclude that the students did understand the target concepts at one level.

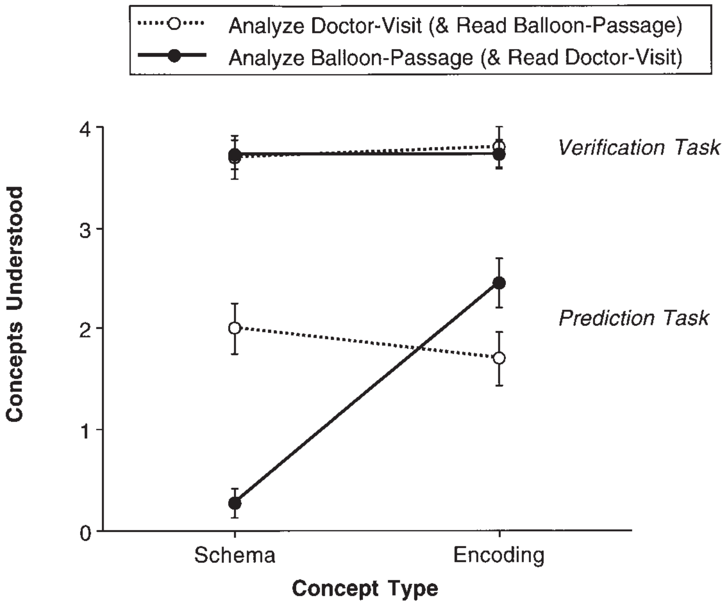


FIGURE 2 Evidence of concept understanding broken out by concept type, condition, and dependent measure (Experiment 1). Students in the analyze doctor visit condition analyzed the doctor visit cases and read about the balloon passage cases. Students in the analyze balloon passage condition analyzed the balloon passage cases and read about the doctor visit cases. Students in both conditions heard the same lecture that explained the schema and encoding concepts.

The prediction task presents a different story. The students tended to make predictions based on the concepts that appeared in the contrasting cases they had analyzed. Students in the analyze balloon passage condition made more predictions based on encoding concepts than did students who read about those cases. Conversely, students in the analyze doctor visit condition made more predictions based on schema concepts than did students who read about those cases.

To confirm the major patterns in Figure 2 statistically, the number of schema predictions and the number of encoding predictions served as measures for the within-subject factor of prediction type. The contrasting cases that a student analyzed, doctor visit or balloon passage, served as a between-subjects factor. The number of correct schema and encoding verifications served as covariates. There was a strong interaction between the cases that a student analyzed and the types of prediction she or he made, $F(1, 18) = 29.3$, $MSE = 0.54$, $p < .01$. Students primarily predicted using concepts that were exemplified in the cases they had analyzed. The verification covariates did not account for a significant portion of the variance on the prediction task, $F(1, 18) = 0.21$, $MSE = 0.54$, $p > .6$. This indicates that students' success at answering the verification questions for a set of concepts was not predictive of their success at using those concepts in the prediction task. Although not theoretically central, there was also a main effect of prediction type indicating that students made more encoding predictions than schema predictions overall, $F(1, 18) = 17.28$, $MSE = 0.54$, $p < .01$. Perhaps the effects of meaningfulness on memory are more easily understood than the more subtle script effects.

Discussion

According to the verification task, students knew the concepts even when they had not analyzed the cases. Because there was no pretest, we cannot know whether the students actually learned the concepts here, but we do know that, at the time of posttest, they understood them sufficiently well to score at near ceiling levels on the verification task. Nonetheless, by another more ecological task—the prediction task—the students could not perform effectively unless they had analyzed the cases.

An important question, one that is unlikely to have an answer found in a single underlying mechanism, is why analyzing the cases helped students on the prediction task, whereas reading about the cases did not. Experiments 2 and 3 are designed to help answer this question; for example, they exclude time on task as a likely explanation. First, we consider some possible theoretical accounts for the results of Experiment 1.

An important difference between the analyze contrasting cases condition and the read-only condition is that the former involved more generation on the part of the learner. A clear example of the benefits of generativity comes from work on the "generation effect" (Slamecka & Graf, 1978). If people generate the partial word in the synonym pair, FAST::R_P_D, they will remember the word RAPID better than if they simply read the pair FAST::RAPID. The generation effect has been used to explain memory advantages in domains substantially more complex than paired associates. Needham and Begg (1991), for example, evoked the generation effect to explain why learning a procedure through problem solving, as opposed to memorizing the procedure, improved subsequent problem-solving performance.

To claim simply that the contrasting cases condition was more active or more generative seems insufficiently precise. There are many different things one might generate (cf. Stein & Bransford, 1979); hence, it is worthwhile to try to identify what type of activity led to these results. For example, the generation effect, as traditionally studied, examines the strengthening of memory and not transfer. In Experiment 1, students in the contrasting cases condition excelled on prediction items that were different from those that they generated while analyzing the cases. Therefore, these results appear to go beyond those typically studied in work on the generation effect.

Learning by doing. One possible generative aspect of the contrasting cases activity that may have caused the prediction results is that students produced an overt response in their graphs. For example, behaviorist theories of acquired distinctiveness argued that differentiation can only occur when individuals produce an external behavior (e.g., Miller & Dollard, 1941). The idea that effective generation involves overt production reflects a common interpretation of constructivism in which "learning by doing" means some form of hands-on or external activity. Therefore, for example, one might propose that "doing" the graphs was responsible for helping the students learn more deeply. Perhaps students who read about the cases (instead of producing something externally) had a weaker or less reinforced memory trace and, therefore, were less able to exercise the relevant memory during the more difficult prediction task.

If external production is the active ingredient of the contrasting cases, then we should expect that, when students do not produce a particular pattern from the cases, they will not be prepared to learn the relevant target concept during a lecture. Alternatively, it is feasible that, even if a student does not generate or notice a given feature, analyzing the contrasting cases could still produce a halo effect for that feature. Much as psychology experts bring to bear their differentiated knowledge to understand the experiments in a journal article, students who analyze the cases may develop knowledge that is sufficiently differentiated that it enables them to fill in the missing details with the aid of the lecture. Students who read about the cases may not learn enough distinctions to enable understanding further distinctions presented in the lecture.

Knowledge assembly. Another class of constructivist explanation for these results draws on the idea that learning best occurs when individuals put forth the effort and attention to assemble their ideas with meaningful connections. This "knowledge assembly" viewpoint is reflected in theories of elaboration that describe learning as a process of connecting ideas (e.g., Bradshaw & Anderson, 1982). By the elaboration account, analyzing the cases encouraged students to assemble relations that connected the case information to other pockets of prior knowledge. Conceivably, this elaboration increased the number of possible retrieval paths (connections) to the target concepts. The multiple retrieval paths increased the chances of recovering the relevant concepts during the prediction task.

As in the case of learning by doing, knowledge assembly is likely to account for some of the effects of the contrasting cases, but it may not provide a complete explanation. If the cause of these results resides solely in whether students elaborate or assemble the concepts, then other methods of leading students to elaborate the target concepts should lead to equivalent performance on the prediction task. If, however, the effect of analyzing the cases is to develop differentiated knowledge, then encouraging elaboration without the contrasting cases should not be as effective.

Discovery as discernment. A third class of constructivist explanation for the results of Experiment 1 involves the idea of discovery. Discovery, as used here, is intended as a psychological description of knowledge growth rather than as an instructional prescription. The "discovery as discernment" position is that individuals learn well when they have generatively discerned features and structures that differentiate relevant aspects of the world. We believe that the discovery of distinctive features is why students who analyzed the cases did better in Experiment 1.

By the discovery as discernment account, reading and hearing about the cases left the students' knowledge overly general and undifferentiated. This general knowledge was sufficient for the verification task because each verification item provided a general statement that would presumably match the general concepts the students had learned. Therefore, the items were effective retrieval cues for this more general understanding. However, for the prediction task, the general knowledge did not help the students notice specific features embedded in the hypothetical study—features that could cue their concept knowledge. This is why analyzing the contrasting cases yielded an advantage on the prediction task; it helped the students discover the characteristic features of the target phenomena. As a result, students could notice those features in the prediction task, and the features could serve as a set of cues that reminded the students of the target concepts (cf. Ross, 1989).

To summarize, in all three of the preceding accounts, learner generativity plays an important role. These constructivist accounts differ, however, with respect to the claims that they make about what was generated in the analyze conditions that led to these results. In the learning by doing version of constructivism, generating an external product was the key ingredient. In the knowledge assembly version, connecting and elaborating ideas was the key ingredient. In the discovery as discernment version, differentiating relevant distinctive features was the key ingredient. The following experiment examines the role of these possibilities in preparing students to learn by being told.

EXPERIMENT 2

Experiment 2 compared students' abilities to learn from a text (rather than a lecture) when they either actively summarized the text or simply read the text after analyzing the contrasting cases. The text involved two "book chapters," the *schema chapter* and the *encoding chapter,* which were written especially for the experiment (available from the authors). The chapters cover their respective concepts and are similar to, but more extensive than, the combined lecture and read-only cases of Experiment 1. Each chapter, including figures, would be about eight pages in an average textbook. They were written so that one could read them separately or together in either order.

As before, students analyzed either the doctor visit or balloon passage cases. Afterward, they received the two book chapters. The students first read the chapter that was relevant to the cases they had analyzed (e.g., schema chapter for doctor visit cases). Then, the students read the chapter that was relevant to the cases they had not analyzed and wrote a two- to three-page summary. Thus, students in the analyze doctor visit condition analyzed the doctor visit cases, read the schema chapter, read the encoding chapter, and then wrote a summary of the encoding chapter. Students in the analyze balloon passage condition analyzed the balloon passage cases, read the encoding chapter, read the schema chapter, and then wrote a summary of the schema chapter. After a 1-week delay, everyone completed the prediction task.

One reason for asking students to write a chapter summary is that it can address a pragmatic concern: Does the typical school assignment of reading and summarizing a text match the benefits of analyzing cases and then simply reading a text? A second reason for including the summarizing condition is that it can help evaluate elaboration as an account for the effect of contrasting cases. There are many different ways to elaborate on a text (e.g., thinking of personal examples). Consequently, summarization may not be representative of other styles of elaboration. Nonetheless, choosing which knowledge to summarize and then assembling that knowledge into a summary requires a construction and evaluation of different possible connections among ideas.

If elaboration is a sufficient explanation for the effects of analyzing the contrasting cases, then we should expect that deciding on and summarizing the important concepts in a text would provide some benefit for the prediction task. If, however, deep understanding requires the differentiated knowledge that comes from analyzing the contrasting cases, then summarizing the text should provide only minimal benefits for the prediction task. We hypothesize that summarizing the chapter will not match the benefits of analyzing the contrasting cases. Before people can productively elaborate on features of the world, they need to discern those features. This discernment is the unique contribution of analyzing the cases.

This study also helps evaluate overt production as an account for the effects of contrasting cases. The students produce graphical documents when analyzing the contrasting cases, and they produce textual documents when summarizing. By looking at these end products of learning by doing, we can determine the extent to which the students' predictions depend on whether they actually produce the relevant concepts or features in their assignments. If the benefits of analyzing the contrasting cases extend beyond those ideas the students actually produce, then the relation between analyzing the contrasting cases and successful prediction cannot be attributed solely to external production. We hypothesize that the benefits of analyzing the contrasting cases extends somewhat beyond those features the individuals actually discover and document. Discovering a subset of distinctive features may provide students with a sufficient critical mass of differentiated knowledge to enable them to understand related distinctions presented in a text.

Method

Participants. Eighteen graduate students in an introductory cognitive psychology course were randomly divided into two conditions. The students had previously completed two contrasting cases packets relevant to perception. This removed task novelty as a likely explanation for any positive effects.

Materials. The contrasting cases and prediction task were the same as in Experiment 1. A schema and an encoding chapter were written for the purposes of the experiment. Each chapter develops theoretical models and the target concepts and uses the experiments to exemplify those concepts.

Design. A within-subject factor contrasted the effects of analyzing the cases and reading a relevant chapter with the effects of reading and summarizing a chapter. A between-subjects factor determined which topics a student analyzed and summarized. Students in the analyze doctor visit condition analyzed the doctor visit cases, read the chapter on schemas, and then read and wrote a summary of the encoding chapter. Students in the analyze balloon passage condition followed the same protocol but with the alternative materials. The dependent measure was the prediction task. To examine whether time on task influenced the results, students estimated how long they spent completing each portion of the experiment.

Procedure. The students received their respective contrasting cases to analyze as a homework assignment at the end of a regular class. After turning in their graphs and time-on-task estimates 5 days later, the students received both chapters (ordered depending on condition). The students were told that the chapters were being evaluated for inclusion in a new textbook. They were told to first read the initial chapter and make a note of how long this took. They were also told to read the second chapter and to write a two- to three-page summary of the important concepts in that chapter (and to note the time on task). The summaries were turned in at class 7 days later. Five days later, students completed the prediction task in class as before.

Coding. The homework was coded to determine which concepts or concept relevant patterns the students noted in their documents. The maximum possible score for a single assignment is 4; there was 1 point for each target concept. For the analyses, a concept was considered noted if the student graphed the relevant empirical pattern or included a marginal note indicating the relevant feature. For example, a doctor visit graph showing false recognition rates counted as noting the pattern relevant to the concept of script intrusions. For the summaries, a concept was considered noted if the student described or referred to the concept or its empirical implications, for example, “People remember things from their schemas that don’t happen.” Two coders, working independently, had five disagreements, which were resolved through negotiation.

Results

Before comparing performances on the prediction task, we examined the frequency with which students noted the target concepts in their homework assignments. The scatterplot of Figure 3 shows the rates at which students noted each of the target concepts. The horizontal dimension indicates the percentage of students who summarized a given concept from a chapter. The vertical dimension indicates the percentage of students who noted the concept-relevant pattern in their case analyses. If students had noted concepts at identical rates in the analyze and summarize treatments, the points would make a diagonal between the origin and the upper right cor-

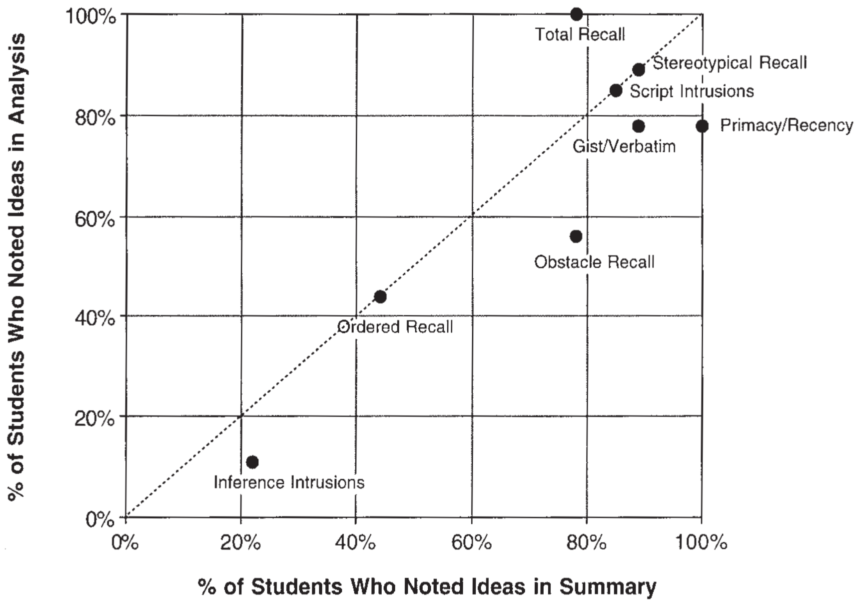


FIGURE 3 Comparison of the rate that target ideas were noticed in summary homework and analysis homework (Experiment 2). The axes represent the percentage of possible notices. For example, 12% of the students who analyzed the balloon passage cases documented inferential intrusions, whereas 22% of the students who summarized the encoding chapter documented inferential intrusions. The comparison is necessarily between subjects.

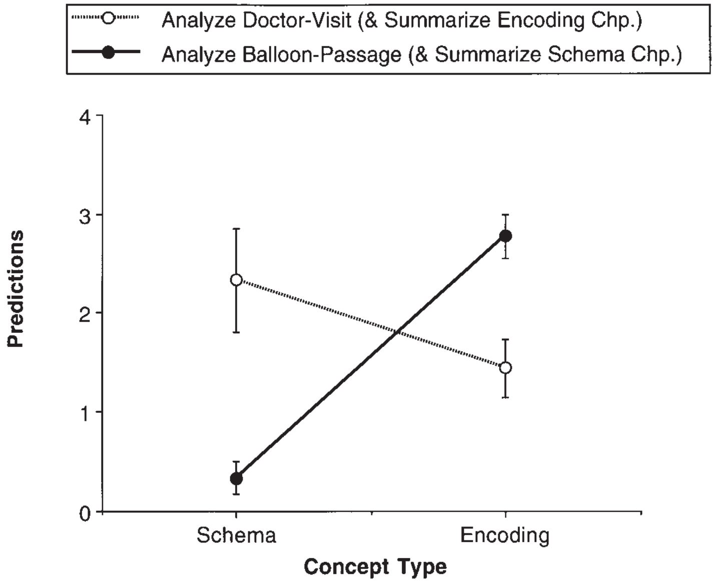


FIGURE 4 Evidence of concept understanding on the prediction task broken out by concept type and condition (Experiment 2).

ner. Figure 3 shows that the points approximate this idealized diagonal. This is a fortuitous outcome. It means that treatment effects on the prediction task cannot be attributed to students overlooking the target concepts in one treatment more than the other. For the students who summarized the schema chapter, an average of 3.0 (*SD* = 1.3) of the four possible target ideas were noted. For the students who analyzed the doctor visit cases, an average of 2.8 (*SD* = 1.1) patterns were found. For the encoding chapter, summarizing yielded 2.9 (*SD* = 0.3) of the target concepts, and analysis of the balloon passage cases yielded 2.7 (*SD* = 0.9) patterns. There were no significant differences or interactions, all *F*s < 1.

Even though the students noted the concepts in their homework at the same rate across treatments and materials, Figure 4 shows that, for both sets of concepts, analyzing the cases led to more predictions than summarizing the chapter. The number of schema predictions and encoding predictions served as within-subject measures. The cases that a student analyzed served as a between-subjects factor. There was a strong interaction whereby students who analyzed the doctor visit cases made more schema predictions than encoding predictions, and students who analyzed the balloon passage cases made more encoding predictions than schema predictions, $F(1, 16) = 17.73$, $MSE = 1.41$, $p < .01$. There was a marginal main effect of prediction type indicating that, as before, encoding predictions were more frequent overall, $F(1, 16) = 3.86$, $MSE = 1.41$, $p < .07$.

The benefits of analyzing the cases extended beyond those concepts that students noted in their homework. Table 7 shows the probability that a given concept would be used for a prediction depending on whether it was noted or missed in the homework. We collapsed the results for the schema and encoding concepts because there were no appreciable differences. It is noteworthy that, even if a given concept was missed when analyzing the contrasting cases, the students still had a .44 probability of applying that concept to the prediction task. This is well above the .11 conditional probability found for the summarizations. Compared to summarizing, the beneficial effects of analysis were not confined to those concepts that were overtly produced in the homework.

An important issue is whether time on task played a role in the students' prediction performance. Perhaps students spent a great deal of time analyzing and graphing the cases and this led to their superior learning and prediction performance. Attributing the superior results of analyzing the cases to time on task would require a difference in the amount of time students spent on the tasks of analyzing and summarizing plus evidence that time on task covaries with performance on the prediction task. There was no evidence supporting either point. Students' self-reports of time on task indicated that they spent about the same amount of time summarizing and analyzing (schema concepts: read and summarize, $M = 89$ min, $SD = 64$; analyze and read, $M = 113$ min, $SD = 61$; encoding concepts: read and summarize, $M = 122$ min, $SD = 83$; analyze and read, $M = 118$ min, $SD = 83$). There were no significant differences by treatment, $F(1, 16) = 0.19$, $MSE = 9{,}458.7$, $p > .6$; by concept cluster, $F(1, 16) = 1.84$, $MSE = 1{,}690.3$, $p > .15$; or by their interaction, $F(1, 16) = 0.5$, $MSE = 1{,}690.3$, $p > .4$. The reports also indicate that time on task did not have a measurable effect on prediction performance. The number of schema predictions and encoding predictions served as within-subject measures. The cases that a student analyzed served as a between-subjects factor. Each individual's two time-on-task estimates served as covariates. The time-on-task covariates do not explain a significant portion of the prediction task variance, $F(1, 15) = 0.96$, $MSE = 1.41$, $p > .34$. The benefits of analyzing over summarizing remains reliable, $F(1, 15) = 15.76$, $MSE = 1.41$, $p < .01$.

TABLE 7
Probability That Students Would Make a Prediction as Conditionalized on Whether They Noted or Missed a Concept in Their Homework (Experiment 2)

	Probability of Prediction Based on a Concept	
Type of Homework	*Noted in Homework*	*Missed in Homework*
Analyzing cases	.74	.44
Summarizing chapter	.26	.11

Discussion

Analyzing the contrasting cases and simply reading a chapter helped students on the prediction task more than reading and writing a summary of the same chapter. It is informative to note that the graduate students who summarized a chapter in this experiment did not outperform the undergraduates of Experiment 1 who simply read about the cases and then heard a lecture. Although summarizing may lead students to elaborate and assemble the key concepts within a text, it does not guarantee that the resulting knowledge will have the differentiation needed for application.

A revealing outcome of this study is that students who missed a concept in their chapter summaries had a negligible chance (.11) of making a prediction based on that concept. However, students who missed a concept-relevant pattern in their case analyses still had a decent chance (.44) of making the relevant prediction. In follow-up questioning about the missed patterns in the analysis activity, the students generally stated that they had not noticed the patterns (as opposed to not bothering to graph them). These results are interesting because the analyzing students who missed a pattern in the data were on equal footing with the summarizing students when they read about the relevant concept in the chapter; neither had previously seen the relevant pattern. However, as it turned out, even when the analyzing students missed a pattern, they still had a higher probability of making the relevant prediction than did the students who explicitly summarized the relevant concept in their chapter summaries (.44 vs. .26).

One implication of this finding is that the effects of analyzing the contrasting cases cannot simply be credited to a generation effect. Analyzing the cases yielded benefits for a given concept, even if the concept-relevant pattern was never generated during the analysis. One possible explanation, both for the extension beyond the generation effect and for the aforementioned advantage of analysis over summary, comes from a consideration of what the students were learning in their analyses of the contrasting cases. By our account, they were discerning important distinctive patterns. Discovering the distinctions may have provided the students with sufficiently precise knowledge that they could subsequently understand the chapter’s description of the patterns they had overlooked. Previous research on perceptual learning, for example, has shown that, if people discern distinctions within a domain, these distinctions can facilitate subsequent learning (Arnoult, 1953; Gagné & Baker, 1950; E. J. Gibson, 1940). At a conceptual level, one can recognize this phenomenon among experts who can understand new distinctions described in an article even though they have never analyzed the raw data. The students who analyzed the contrasting cases developed some expert-like differentiation that could support learning by being told, whereas the summarizing students never had the chance. Consequently, they did not understand the chapter’s descriptions and explanations at a deep level. Future research will be needed to explore this interpretation further.

In the meantime, an important remaining question is whether distinctions, once discovered, are sufficient for deep understanding. In the introductory section, we claimed that opportunities to analyze cases can set the stage for further learning but are often not sufficient because students lack an overriding framework that helps them develop a theory or model to explain the distinctions they have discovered. Experiments 1 and 2 did not allow us to test this claim.

Experiment 3 addresses this issue by examining whether analysis without a following lecture leads to strong performance on the prediction task. We hypothesize that it will not. Students will not develop the framework that helps them to generalize their differentiated knowledge to new situations (i.e., the prediction task).

Experiment 3 is also designed to allow more precise control over variables of theoretical importance. In Experiments 1 and 2, we attempted to maintain the ecological validity of the results. The students, for example, completed the assignments as homework. This allows for an easy extension of the experimental manipulations into educational practice, but it leaves open many alternative interpretations of the results. One alternative was that students spent more time on their analyses than their summaries. Although the analyses of the time-on-task reports did not support this alternative, it is possible that the students' reports were inaccurate. A more controlled experimental environment is worthwhile in this regard. There are also other concerns. When students analyzed the cases, they had two exposures to the concepts spread over several days; they analyzed the cases and then read a chapter that rehashed the cases they had examined. In contrast, the summarizing task provided only a single exposure to the relevant information. Another concern is that the chapter summarization always came second in this study, which might put it at a disadvantage. The following experiment uses a more strictly controlled and balanced design to defuse these and other alternative explanations.

EXPERIMENT 3

This study pulls together the main experimental hypotheses into a single experiment. First, it tests the hypothesis that "creating a time for telling by doing more telling" may not always be optimal. This study improves on the demonstration of Experiment 2; students in a *summarize + lecture* condition both summarize a relevant text chapter and hear a subsequent lecture. Second, the experiment tests the hypothesis that analyzing these contrasting cases alone is not sufficient for deep learning. The study includes an *analyze + analyze* condition in which students analyze the same set of contrasting cases twice but never hear a lecture or read a text. We have no a priori basis for predicting how the "double discovery" and "double telling" conditions should fare relative to one another. Finally, the experiment tests the hypothesis that analyzing the contrasting cases prepares students to learn by being told. For this *analyze + lecture* condition, students analyze the cases and hear a lecture. If the hypotheses are correct, this "discovery and telling" condition should lead to the best performance on the prediction task.

Unlike the previous experiments, this study does not counterbalance the use of the schema and encoding materials by instructional treatment. All the students begin by analyzing the balloon passage and hearing the lecture on encoding concepts. It is only afterward that they complete their respective between-subjects treatments by learning the schema concepts in one of the three ways described previously. One reason for having everyone learn the encoding concepts in the same way is that it simplifies the experimental design and procedures as compared to counterbalancing the possible treatment and concept combinations. At the same time, it provides baseline measures for ensuring that the experimental groups are comparable. Presumably, all conditions should apply the encoding concepts equally well to the prediction task.

Method

Thirty-six college sophomores who had no prior courses in psychology participated. The experiment employed a mixed design using the prediction task as the dependent measure. The students completed all the elements of the study within a classroom setting. Students served as their own controls by completing the encoding lesson first. As an in-class assignment, the students analyzed the balloon passage cases for 80 min. Two days later, the students collectively heard the lecture on the encoding concepts. For the schema lesson 5 days later, the students were randomly separated into three between-subjects conditions. In the analyze + lecture and analyze + analyze conditions, students received a packet with the doctor visit contrasting cases and worked with them for 80 min. In the summarize + lecture condition, students received a packet with the schema chapter and read and summarized the chapter for 80 min. This seat work was collected at the end of class. Two days later, the students in the two lecture conditions heard a 30-min lecture on the schema concepts. The students in the analyze + analyze condition spent this 30-min period looking for more doctor visit patterns in another room. Seven days later, the students completed the prediction task as before. Coding of the student work was completed as before. There was 92% agreement for the prediction task and 86% initial agreement for the graphs and summaries. (Because of the in-class time pressures, students' graphs and summaries were not as well composed as before.)

Results

The performances on the prediction task indicate that neither the double discovery nor the double telling were sufficient for deep learning as compared to the combination of discovery and telling. Figure 5 shows that the students in the analyze + lec-

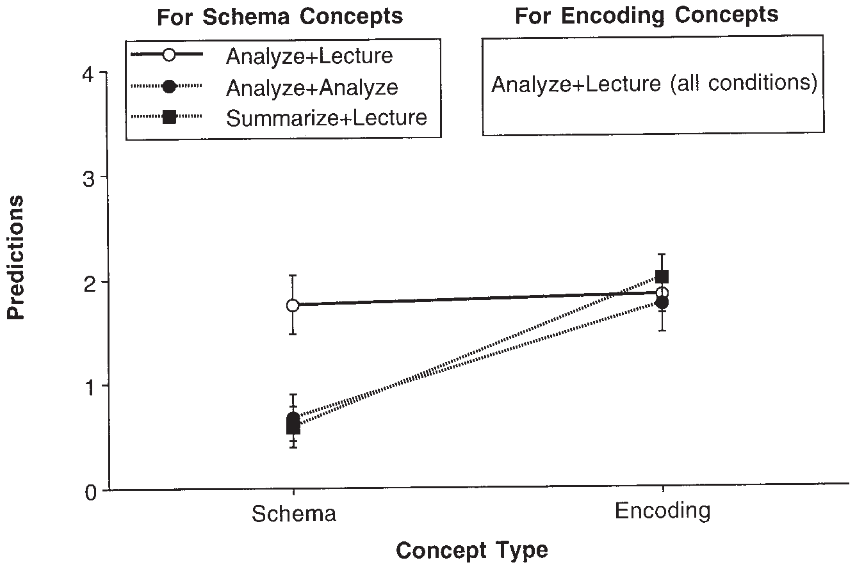


FIGURE 5 Evidence of concept understanding on the prediction task broken out by concept type and condition (Experiment 3). All students analyzed the balloon passage cases and heard the encoding lecture. For the schema concepts, the analyze + lecture condition analyzed the doctor visit cases and heard the schema lecture, the analyze + analyze condition analyzed the doctor visit cases twice, and the summarize + lecture condition summarized the schema chapter and heard the schema lecture.

ture treatment made the most schema predictions. In contrast, students in the summarize + lecture and analyze + analyze treatments made fewer schema predictions and were about the same as one another. Furthermore, the students in the summarize + lecture and analyze + analyze treatments made fewer predictions based on schema concepts than encoding concepts. This makes sense in that they had the opportunity to learn the encoding concepts through discovery and telling.

To test statistically the advantage of the analyze + lecture treatment, the number of schema predictions and the number of encoding predictions served as within-subject measures. The three treatments for the schema concepts served as a between-subjects variable. A significant interaction of Treatment × Prediction Type (schema and encoding), $F(2, 33) = 5.49$, $MSE = 0.03$, $p < .01$, shows that the analyze + lecture condition yielded the same number of schema predictions as encoding predictions, whereas students in the other two treatments made fewer schema predictions than encoding predictions. Evidently, neither the contrasting cases nor a verbal exposition were sufficient in their own right for deep understanding; these novices needed both the discovery and the telling.

Confining the discussion to the schema lessons (which are in experimental focus), the students who summarized the chapter noted 66.7% of the target ideas in their summaries, the analyze + analyze students noted 58.3% of the ideas in their graphs, and the analyze + lecture students noted 52.1% of the ideas in their graphs. (Perhaps due to the 80-min limit, students in the two analyze conditions did not note as many ideas as in Experiment 2.) Even though the analyze + lecture students noted the fewest concepts in their assignments, they performed the best on the prediction task. Table 8 shows the probabilities of making a schema prediction conditionalized on whether the concept was noted during the lesson. As in Experiment 2, the analyze + lecture treatment yielded the highest probability of making a target prediction, regardless of whether the students noted or missed the concept in their analyses.

Discussion

In this study, the double telling students who summarized a chapter on the target concepts and then heard a further lecture on those concepts did not perform well on the prediction task. Similarly, the double analyze students who looked for patterns in the cases without any following exposition did not perform well. This latter group of students showed poor performance even for those concept-relevant patterns that they had discovered in their analyses. In contrast, students who analyzed the cases and heard a lecture did quite well on the prediction task. In addition to demonstrating the beneficial effects of creating a time for telling, these results rule out time on task as an explanatory variable because all students in all treatments worked for the same amount of time. It is interesting to note that, even when combined, the percentage of possible predictions made by the analyze + analyze students (16.7%) and the summarize + lecture students (14.6%) does not add up to the percentage of predictions made by analyze + lecture students (43.8%). This suggests a synergy between the opportunity to differentiate one’s knowledge of the phenomenon at hand and the opportunity to hear a conceptual framework that articulates the significance of those phenomena.

TABLE 8
Probability That Students Would Make a Schema Prediction as Conditionalized on Whether They Noted or Missed a Concept in Their Analyses or Summaries (Experiment 3)

	Probability of Prediction Based on a Schema Concept	
Condition	*Noted in Study Work*	*Missed in Study Work*
Analyze + lecture	.60	.26
Analyze + analyze	.18	.15
Summarize + lecture	.23	.06

GENERAL DISCUSSION

In three experiments, students analyzed contrasting cases that brought to light memory phenomena. Afterward, they received a lecture or text on memory. Approximately 1 week later they predicted outcomes for a hypothetical experiment that could be interpreted in light of the concepts they had studied. The students made more accurate predictions compared to the control treatments of (a) reading about the features in the cases and hearing a lecture, (b) summarizing a relevant text and hearing a lecture, and (c) analyzing the cases twice without receiving a lecture. Interestingly, the benefits of analyzing the cases extended to concepts that students did not actually discover in their analyses. This occurred, however, only when they received a subsequent exposition. In combination, the results indicate that teaching by telling can play a significant role in deepening students' understanding if the students have had a chance to acquire appropriate prior knowledge. In these studies, contrasting cases helped the students generate this prior knowledge. Our interpretation of the psychological mechanism behind these results is that analyzing the contrasting cases provided students with the differentiated knowledge structures necessary to understand a subsequent explanation at a deep level.

Comparisons to Other Common Methods of Instruction

As with any new instructional approach, there are a proliferation of questions about the active ingredients that make the instruction work. For example, would the effects be as strong if students read the text first and then analyzed the cases? How important was the instruction to graph the results of the analyses, and would nonvisual production work as well? Does the initial analysis of the contrasting cases lead students to attend to some topics at the expense of others during a subsequent telling? The empirical effort to decompose all the possible ingredients and outcomes is beyond the scope of this article. Moreover, to maintain the practical force of our claims, we have not moved to tightly controlled laboratory settings in which it would be more feasible to single out underlying psychological processes. Whole-scale comparisons of instructional methods cannot easily isolate underlying psychological causes. They do, however, provide evidence that new instructional approaches are worth pursuing. In addition, they can provide a chance to contrast the effectiveness of different psychological theories when applied to an authentic learning context.

In these three experiments, we implemented our basic claim that deep understanding requires both a differentiated knowledge structure (as develops when discerning the contrasts among cases) and an explanatory knowledge structure (as often comes through telling). The studies allowed us to explore several alternative psychological proposals for how to enhance understanding through generative activity, including elaboration (knowledge assembly) and overt production (learning by doing). Other instructional implementations of elaboration theories and external production would make for fitting comparisons, as would prescriptive pedagogical proposals such as case-based learning (Williams, 1992) and project-based learning (Barron et al., 1998). We consider four possibilities to help frame our main instructional points. Our intent in the following comparisons is not to suggest that contrasting cases are the only way to create a time for telling. Rather, our intent is to clarify our claims about why this intervention worked so well by placing those claims in a falsifiable context.

One approach frequently used in classrooms is to have students read a text and think of personal examples. Another approach is to have students think of questions about the text. These activities are both forms of elaboration; students try to make a connection or spot a missing connection between the information in the text and their knowledge. Would these forms of elaboration work better than the contrasting cases approach? We suspect not. Elaborative approaches assume that students already have the differentiated knowledge that would enable them to assemble appropriate examples or questions. In many cases, however, the personal examples or questions that novices generate for a given concept may have only a vague relation to the distinctive properties of the concept. As a consequence, there is little in the elaborations or questions that can help the novices ascertain the truly distinctive features in future situations. This example highlights the first of our claims. Assembling ideas is important for understanding, but it is also important that people discern the distinctive features of the ideas they are assembling.

A second common instructional approach, in psychology at least, is to have students participate in classic experiments. The idea is that students will be active participants in the phenomena; they will be learning by doing. Although worthwhile, there are two reasons that we do not embrace this specific approach as a way to teach concepts, even if it is coupled with a subsequent lecture or text. The first is pragmatic—participating in experiments is cumbersome and becomes an occasional event rather than a staple of pedagogy. The method of contrasting cases, however, easily scales up to a weekly homework assignment that can be completed prior to attendance in a large lecture hall. The second concern is that doing does not necessarily result in "doing with understanding." Experimental participation puts the student in the role of subject rather than psychologist. The opportunities for noticing features as a subject are quite different from when analyzing the cases. For example, students may focus on completing the experiment rather than on noticing distinctions in their own thoughts and behavior. If the goal is to have students think like a psychologist, then they should participate like psychologists during learning (cf. Morris, Bransford, & Franks, 1977). This example highlights our second point. Generative activity is central to understanding, but the right things must be generated.

A third method of instruction is to use a “telling case.” Students, for example, may explore a videotape of a psychological phenomenon in action, or they may read a detailed example. This case-based learning grounds otherwise overly abstract knowledge. Coupled with a principled explanation, this instruction can be very effective (Hmelo, 1994; White, 1993). A potential risk with this approach, however, is that it assumes that students will notice the relevant features of the case or example. Without explicit contrasts, this may only occur if students have sufficient background knowledge to discover the important features of the case. An expert teacher, for example, viewing a videotape of an elementary classroom may notice that the chairs are too close to the door and could cause chaos at the bell. A novice, however, is unlikely to observe this classroom feature.

One way around limited prior knowledge is to point out important features. Experiments 2 and 3, for example, showed that students could make predictions based on features that were told to them but not noticed on their own. These predictions, however, occurred only when students had already differentiated some of the domain’s distinctive features. Moreover, Experiment 1 showed that telling students about features was not as effective as helping students actively discover them. Further support for the importance of helping novices generate important features came from Gagné and Brown (1961). They found that neither instruction relying on unscaffolded discovery from an example nor instruction using a single example and an explanatory rule are as effective as instruction that takes the form of guided discovery. In these experiments, one may think of the contrasting cases as a way to guide the students’ discovery of significant features, highlighting our third point. Contrasting cases provide a powerful way to help students differentiate their knowledge of a domain.

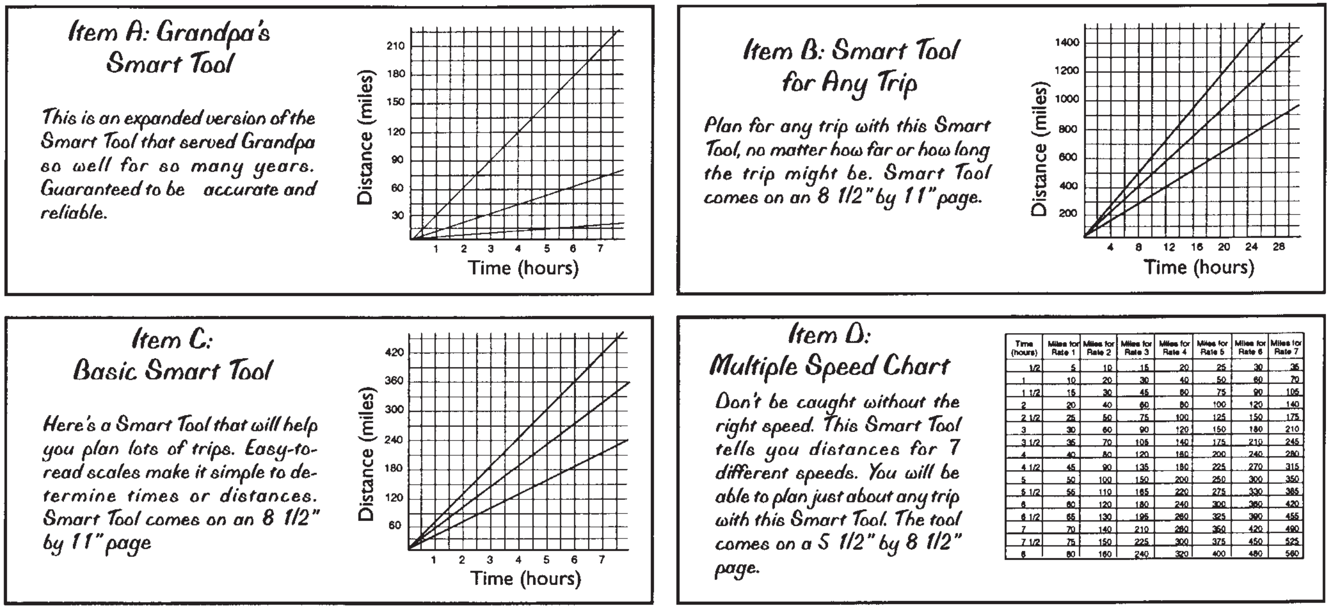
A fourth instructional approach asks students to design and conduct their own experiments. In this form of project-based learning by doing, students are both engaged and generative, and they explore the contrasts inherent to their own experimental designs. Assuming that novices receive appropriate support so they can design revealing experiments, this approach can be a powerful learning vehicle (Barron et al., 1998). The project, however, is often the culminating event in the instructional sequence rather than a preparation for further learning. An opportunity is missed when project activities create a time for telling but do not take advantage of it. Experiment 3 demonstrated that discovery without subsequent explanation does not always lead to deep understanding. This highlights our fourth point. Deep understanding requires both differentiated knowledge about phenomena and an understanding of the significance of those differences.

Features of Good Contrasting Cases

Our experiments provided no formal tests of the optimal elements of a contrasting case, but a few examples can nonetheless move the discussion forward. The contrasts used here, the doctor visit and balloon passage cases, included multiple contrasts and the background noise of variable data. The students were able to handle this complexity, in part, because they had sufficient domain knowledge and psychological insight to recognize worthwhile contrasts. In domains in which students have less prior experience, less complex contrasting cases may be more appropriate lest students get lost in the little contrasts. For example, we have asked prealgebra students to choose among tools that can help solve classes of rate–time–distance problems (Bransford et al., in press). Figure 6 shows some of the contrasting tools from which the students may choose. The contrasts between the tools are less "cluttered" compared to the contrasting cases of these studies. This makes it so students with limited algebra knowledge can still locate the important contrasts. Students who worked with these contrasting cases, for example, discussed the meaning of the axes and their scales, whereas students who worked with a single correct graph tended to interpret single points plotted on the graph. The former case represents an increased preparedness for algebra instruction because the children considered general relations rather than a single, arithmetic instance.

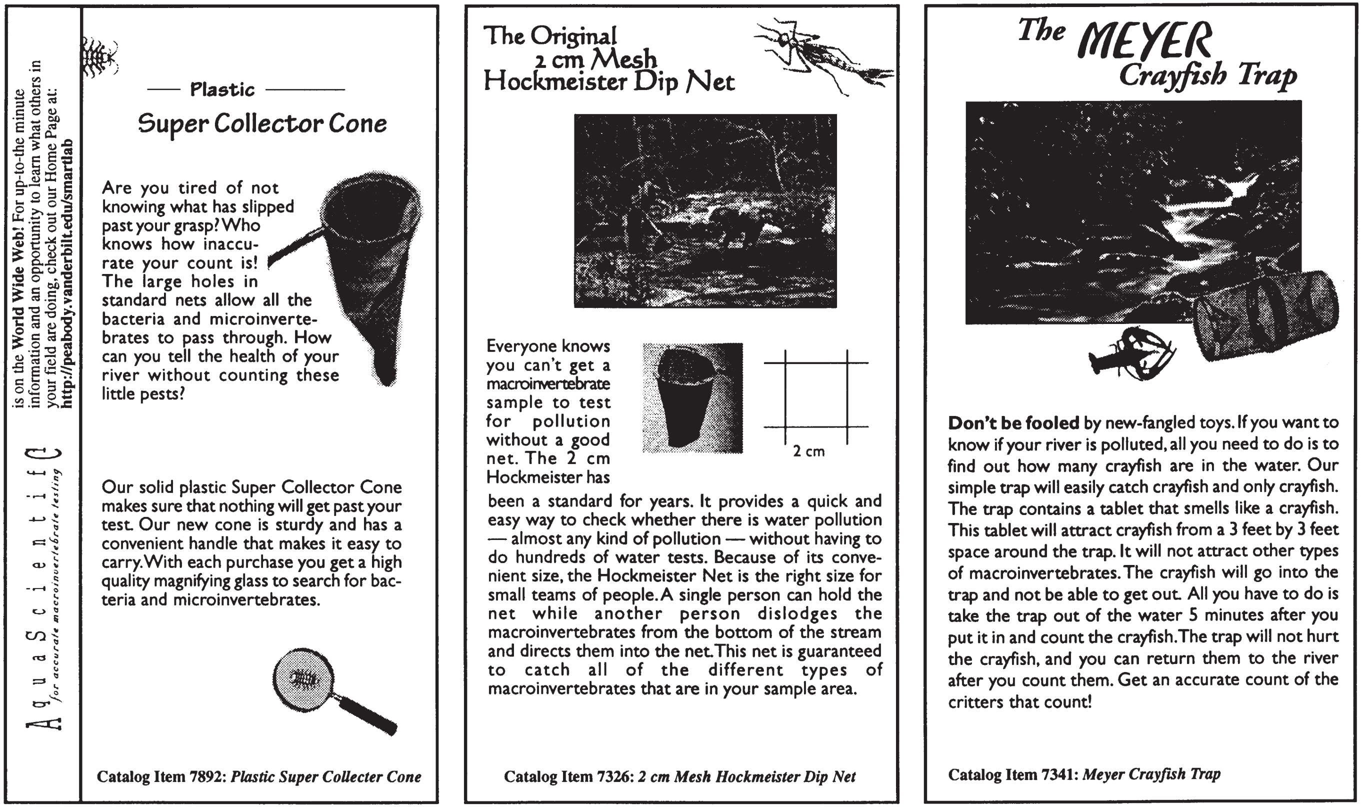
Sometimes, there are reasons to include multiple contrasts in noisy settings, for example, to help students develop the ability to recognize important features and data embedded in future settings (CTGV, 1992). In cases of complexity and limited knowledge, some framing can bring the intended contrasts into relief. For example, we have used an anchoring video story about monitoring for water pollution (CTGV, 1997) in conjunction with the catalog shown in Figure 7. The students "order" the tool that best tests for pollution in a river. The anchor helps students pick out the functional issues, whereas the contrasts among catalog items help students develop questions that differentiate specific alternatives (e.g., should they test for the total number of organisms or the number of types of organisms; Vye et al., 1998).

Perhaps the key feature to a good set of contrasting cases is that they are aligned to the learning goals. For example, if giving a lecture on psychological principles, the contrasts should differentiate empirical features that exemplify the principles in action. It is important to note, however, that contrasting cases are not limited to discriminating empirical phenomena. If the goal is to have people learn about different theories and their distinctive contributions, one might rely on contrasting viewpoints. Multiple perspectives on a single case can be very effective in this regard (Schwartz, Brophy, Lin, & Bransford, in press). For example, we recorded the brief observations of eight psychology faculty who were shown a videotape of an infant's first encounter with a voice-activated mobile. Faculty with different specialties naturally noticed different things. For example, one person wondered how to determine whether the child has learned that her voice activates the mobile, another commented on the instrumentality of voice and language, and a third brought up issues of trust and response consistency. When juxtaposed, these multiple perspectives prepared students to learn how different theories notice and explain different elements of the same situation.



Time (hours)	Miles for Rate 1	Miles for Rate 2	Miles for Rate 3	Miles for Rate 4	Miles for Rate 5	Miles for Rate 6	Miles for Rate 7
1/2	5	10	15	20	25	30	35
1	10	20	30	40	50	60	70
1 1/2	15	30	45	60	75	90	105
2	20	40	60	80	100	120	140
2 1/2	25	50	75	100	125	150	175
3	30	60	90	120	150	180	210
3 1/2	35	70	105	140	175	210	245
4	40	80	120	160	200	240	280
4 1/2	45	90	135	180	225	270	315
5	50	100	150	200	250	300	350
5 1/2	55	110	165	220	275	330	385
6	60	120	180	240	300	360	420
6 1/2	65	130	195	260	325	390	455
7	70	140	210	280	350	420	490
7 1/2	75	150	225	300	375	450	525
8	80	160	240	320	400	480	560

FIGURE 6 Contrasting smart tools in the domain of prealgebra. Students choose which tool(s) would be most helpful for solving various rate–time–distance problems.



is on the **World Wide Web!** For up-to-the minute information and an opportunity to learn what others in your field are doing, check out our Home Page at: **http://peabody.vanderbilt.edu/smartlab**

AquaScientific
for accurate macroinvertebrate testing

— **Plastic** —

Super Collector Cone

Are you tired of not knowing what has slipped past your grasp? Who knows how inaccurate your count is! The large holes in standard nets allow all the bacteria and microinvertebrates to pass through. How can you tell the health of your river without counting these little pests?

Our solid plastic Super Collector Cone makes sure that nothing will get past your test. Our new cone is sturdy and has a convenient handle that makes it easy to carry. With each purchase you get a high quality magnifying glass to search for bacteria and microinvertebrates.

Catalog Item 7892: ***Plastic Super Collecter Cone***

The Original 2 cm Mesh Hockmeister Dip Net

2 cm

Everyone knows you can't get a macroinvertebrate sample to test for pollution without a good net. The 2 cm Hockmeister has been a standard for years. It provides a quick and easy way to check whether there is water pollution — almost any kind of pollution — without having to do hundreds of water tests. Because of its convenient size, the Hockmeister Net is the right size for small teams of people. A single person can hold the net while another person dislodges the macroinvertebrates from the bottom of the stream and directs them into the net. This net is guaranteed to catch all of the different types of macroinvertebrates that are in your sample area.

Catalog Item 7326: ***2 cm Mesh Hockmeister Dip Net***

The MEYER Crayfish Trap

Don't be fooled by new-fangled toys. If you want to know if your river is polluted, all you need to do is to find out how many crayfish are in the water. Our simple trap will easily catch crayfish and only crayfish. The trap contains a tablet that smells like a crayfish. This tablet will attract crayfish from a 3 feet by 3 feet space around the trap. It will not attract other types of macroinvertebrates. The crayfish will go into the trap and not be able to get out. All you have to do is take the trap out of the water 5 minutes after you put it in and count the crayfish. The trap will not hurt the crayfish, and you can return them to the river after you count them. Get an accurate count of the critters that count!

Catalog Item 7341: ***Meyer Crayfish Trap***

FIGURE 7 A catalog of contrasting tools for measuring macroinvertebrates.

Roles for Telling

Although contrasting cases are effective at scaffolding the development of differentiated knowledge, there is a limit to what we can reasonably expect people to discover. A simple contrast may help one discover the distinction between a primacy and recency effect, but it seems unlikely that one would spontaneously discover a theory of working memory that could explain this effect. This is where direct teaching can play a valuable role. It can offer a higher level explanation that would be quite difficult or time consuming to discover. A higher level explanation is important because it provides a generative framework that can extend one's understanding beyond the specific cases that have been analyzed and experienced (e.g., Schwartz & Black, 1996; Schwartz & Moore, 1998). Consider, for example, the third study that the students analyzed in the doctor visit cases (see the Appendix). All the participants in the case remembered that the patient forgot his wallet. Many of the students in our experiment noted and graphed the participants' excellent recall of the forgotten wallet, but few had developed an explanation for why this event was remembered. Consequently, students were not in a position to recognize related examples that differed at a surface level. A lecture or text on this topic remedied this problem by pointing out that scripts have a goal-driven, sequential structure and that events interfering with the successful completion of a script's goal are well remembered. Given this information, the students could recognize the same phenomenon in the prediction task; namely, mixing red and white clothes would prevent successful completion of a clothes washing script. Consequently, the students could infer that this part of the washing machine passage is something that participants should remember.

A text, a lecture, and other forms of direct teaching should not be dismissed as requiring passive reception on the part of a learner (Begg, Vinski, Frankovich, & Holgate, 1991). A given text may simply be an effective or ineffective way to encourage the generative processes by which people construct understanding. Reading a text or hearing a lecture may seem like a passive activity because novices often do not have sufficient background knowledge to approach the text generatively. In reading the same article, novices may generate very little, whereas experts can construct many plausible interpretive alternatives. The important question at hand is how to help readers be generative when they read a text on a relatively novel topic, or as we have asked here, how to create a time for telling.

One approach has been to improve the structure of a text to make it easier for readers to process (e.g., T. H. Anderson & Armbruster, 1984). A considerate text frees up resources for the generative processes required to construct understanding. Other approaches, such as ours, emphasize the role of knowledge. The challenge for these approaches is to present the expert's knowledge in a way that novices can use. When writing for other experts, the challenge is not so great because one may presume that the audience has sufficiently differentiated knowledge to take advantage of the written expertise. When writing for novices, however, special measures and cautions need to be taken to create a time for telling, thus avoiding "the expert trap." Distinctions can seem obvious to experts who, therefore, do not bother to illuminate them. Moreover, the expert's ability to discern is often tacit and, hence, goes unrecognized. As a result, neither instructor nor students may recognize that students have missed important distinctions.

Even when there is an attempt to help students differentiate, the task can be difficult when novices are simply told about distinctions they should make. It is relatively easy to tell a distinction to someone, if that person shares the same set of experiences. However, with respect to the content of instruction, shared experiences are exactly what novices and experts are missing. So, for example, we have found that, even though we, as instructors, carefully describe a cued-recall measure and the students use the term *cued-recall,* they are often really thinking with the less differentiated notion, *memory test.*

By considering the relation between a text and the examples it relies on, we can describe two different knowledge-based approaches to help novices be generative during a telling. One approach might be called the "For Example" approach. Expository devices such as examples and analogies help novices use their knowledge of the example or analogy to make sense of the expert's new ideas. By this model, the examples and analogies illuminate the expert's ideas. The generative activity of the reader occurs in trying to find the similarities between their prior knowledge of the example and the expert's explanations.

A different tact may be called the "Detective Story" approach. In this approach, experts view their explanations as something like the solution to a good detective story. They treat an example as something to be explained rather than as the thing that explains the experts' theories. The generativity of the reader occurs in seeing how the experts' explanations illuminate the examples. The text becomes the solution to a puzzle. Of course, for a text to provide the solution to a puzzle, readers must know the puzzle pieces at a fairly precise level. If readers have only a general knowledge of the pieces, then any method of fitting them together will seem adequate, and there will be little generative activity dedicated to determining whether the expert's explanations join the pieces snugly. If the details are known, however, readers can actively determine whether the explanation joins the edges of the pieces adequately. We believe contrasting cases can do such a good job of creating a time for telling because they help novices to discern the distinctive features of the puzzle pieces. This differentiated knowledge makes the "fitting together" done by the expert detective all the more compelling, and it provides the background for understanding deeply because the appropriate distinctions have been made by the learners. As a consequence, learners are better able to grasp the significance of what the expert has to say.

ACKNOWLEDGMENTS

This work was supported by a University Research Grant from Vanderbilt University, by the Lilly Foundation, and by Department of Education Grant R305F60090. We appreciate the thoughtful comments of Mina Johnson-Glenberg, an anonymous reviewer, Walter Kintsch, Leona Schauble, and Isabel Beck. The opinions are our own and do not necessarily reflect the opinions of the granting agencies.

REFERENCES

Abelson, R. P. (1981). The psychological status of the script concept. *American Psychologist, 36,* 715–729.

Anderson, R. C., & Ortony, A. (1975). On putting apples into bottles—A problem of polysemy. *Cognitive Psychology, 7,* 167–180.

Anderson, T. H., & Armbruster, B. B. (1984). Content area textbooks. In R. C. Anderson, J. Osborn, & R. J. Tierney (Eds.), *Learning to read in American schools* (pp. 195–226). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Arnoult, M. D. (1953). Transfer of predifferentiation training simple and multiple shape discrimination. *Journal of Experimental Psychology, 45,* 401–409.

Ausubel, D. P. (1968). *Educational psychology: A cognitive view.* New York: Holt, Rinehart & Winston.

Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrell, N. S., & Nitsch, K. E. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior, 13,* 471–481.

Barron, B. J., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., Bransford, J. D., & The Cognition and Technology Group at Vanderbilt. (1998). Doing with understanding: Lessons from research on problem- and project-based learning. *The Journal of the Learning Sciences, 7,* 271–311.

Beck, I. (1984). Developing comprehension: The impact of the directed reading lesson. In R. C. Anderson, J. Osborn, & R. J. Tierney (Eds.), *Learning to read in American schools* (pp. 3–20). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Begg, I., Vinski, L., Frankovich, L., & Holgate, B. (1991). Generating makes words memorable, but so does effective reading. *Memory & Cognition, 19,* 487–497.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology, 11,* 177–220.

Bradshaw, G. L., & Anderson, J. R. (1982). Elaborative encoding as an explanation of levels of processing. *Journal of Verbal Learning and Verbal Behavior, 21,* 165–74.

Bransford, J. D., Franks, J. J., Vye, N. J., & Sherwood, R. D. (1989). New approaches to instruction: Because wisdom can’t be told. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 470–497). New York: Cambridge University Press.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717–726.

Bransford, J. D., & McCarrell, N. S. (1974). A sketch of a cognitive approach to comprehension. In W. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes* (pp. 189–229). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bransford, J. D., & Nitsch, K. E. (1978). Coming to understand things we could not previously understand. In J. F. Kavanagh & W. Strange (Eds.), *Speech and language in the laboratory, school and clinic* (pp. 267–307). Cambridge, MA: MIT Press.

Bransford, J. D., & Schwartz, D. L. (in press). Rethinking transfer: A simple proposal with educational implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (Vol. 24). Washington, DC: American Educational Research Association.

Bransford, J. D., Zech, L., Schwartz, D. L., Barron, B. J., Vye, N. J., & The Cognition and Technology Group at Vanderbilt. (in press). Designs for environments that invite and sustain mathematical thinking. In P. Cobb, E. Yackel, & K. McClain (Eds.), *Symbolizing, communicating, and mathematics: Perspectives on discourse, tools, and instructional design.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Britton, B. K., & Graesser, A. C. (Eds.). (1996). *Models of understanding text.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229–272). Cambridge, MA: MIT Press.

Chi, M. T. H. (1976). Short-term memory limitation in children: Capacity or processing deficits? *Memory & Cognition, 4,* 559–572.

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher, 23*(7), 13–20.

Cognition and Technology Group at Vanderbilt. (1992). The Jasper series as an example of anchored instruction: Theory, program description and assessment data. *Educational Psychologist, 27,* 291–315.

Cognition and Technology Group at Vanderbilt. (1996). Looking at technology in context: A framework for understanding technology and education research. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 807–840). New York: MacMillan.

Cognition and Technology Group at Vanderbilt. (1997). *The Jasper project.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Dooling, D. J., & Lachman, R. (1971). Effects of comprehension on retention of prose. *Journal of Experimental Psychology, 88,* 216–222.

Frase, L. T. (1975). Prose processing. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9, pp. 1–47). New York: Academic.

Gagné, R. M., & Baker, K. E. (1950). Transfer of discrimination training to a motor task. *Journal of Experimental Psychology, 40,* 314–328.

Gagné, R. M., & Brown, L. T. (1961). Some factors in the programming of conceptual learning. *Journal of Experimental Psychology, 62,* 313–321.

Gagné, R. M., & Gibson, J. J. (1947). Research on the recognition of aircraft. In J. J. Gibson (Ed.), *Motion picture training and research* (pp. 113–168). Washington, DC: U.S. Government Printing Office.

Garner, W. R. (1974). *The processing of information and structure.* Potomac, MD: Lawrence Erlbaum Associates, Inc.

Gibson, E. J. (1940). A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychological Review, 47,* 196–229.

Gibson, E. J. (1969). *Principles of perceptual learning and development.* New York: Meredith.

Gibson, J. J., & Gibson, E. J. (1957). Perceptual learning: Differentiation or enrichment. *Psychological Review, 62,* 32–51.

Gick, M. L., & Paterson, K. (1992). Do contrasting examples facilitate schema acquisition and analogical transfer? *Canadian Journal of Psychology, 46,* 539–550.

Hmelo, C. (1994). *Development of independent learning and thinking: A study of medical problem-solving and problem-based learning.* Unpublished doctoral dissertation, Vanderbilt University, Nashville, TN.

Karplus, R. (1981). Education and formal thought: A modest proposal. In I. E. Siegel, D. M. Brodzinsky, & R. M. Golinkoff (Eds.), *New directions in Piagetian theory and practice* (pp. 285–314). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Kintsch, W., Britton, B. K., Fletcher, C. R., Kintsch, E., Mannes, S. M., & Nathan, M. J. (1993). A comprehension-based approach to learning and understanding. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 30, pp. 165–214). New York: Academic.

Lindberg, M. (1980). The role of knowledge structures in the ontogeny of learning. *Journal of Experimental Child Psychology, 30,* 401–410.

Littlefield, J., Delclos, V., Lever, S., Clayton, K., Bransford, J., & Franks, J. (1988). Learning LOGO: Method of teaching, transfer of general skills, and attitudes toward school and computers. In R. E. Mayer (Ed.), *Teaching and learning computer programming* (pp. 111–135). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Text coherence, background knowledge and levels of understanding in learning from text. *Cognition and Instruction, 14,* 1–43.

Michael, A. L. (1989). *The transition from language theory to therapy: Test of two instructional methods.* Unpublished doctoral dissertation, Vanderbilt University, Nashville, TN.

Michael, A. L., Klee, T., Bransford, J. D., & Warren, S. (1993). The transition from theory to therapy: Test of two instructional methods. *Applied Cognitive Psychology, 7,* 139–154.

Miller, N. E., & Dollard, J. (1941). *Social learning and imitation.* New Haven, CT: Yale University Press.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519–533.

Needham, D. R., & Begg, I. M. (1991). Problem-oriented training promotes spontaneous analogical transfer: Memory-oriented training promotes memory for training. *Memory & Cognition, 19,* 543–557.

Ross, B. H. (1989). Remindings in learning and instruction. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 438–469). New York: Cambridge University Press.

Scardamalia, M., & 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.

Schwartz, D. L., & Black, J. B. (1996). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science, 20,* 457–497.

Schwartz, D. L., Brophy, S., Lin, X. D., & Bransford, J. D. (in press). Software for managing complex learning: An example from an educational psychology course. *Educational Technology Research and Development.*

Schwartz, D. L., & Moore, J. L. (1998). The role of mathematics in explaining the material world: Mental models for proportional reasoning. *Cognitive Science, 22,* 471–516.

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

Voss, J. F., Vesonder, G. T., & Spilich, G. J. (1980). Text generation and recall by high knowledge and low knowledge individuals. *Journal of Verbal Learning and Verbal Behavior, 19,* 651–667.

Vye, N. J., Schwartz, D. L., Bransford, J. D., Barron, B., Zech, L., & Cognition and Technology Group at Vanderbilt. (1998). SMART environments that support monitoring, reflection, and revision. In D. J. Hacker, J. Dunlosky, & A. Graessar (Eds.), *Metacognition in educational theory and practice* (pp. 305–346). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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.

APPENDIX

The first materials are the doctor visit contrasting cases that the students analyzed, and the second materials are the balloon passage contrasting cases. The students' task was to analyze the data and graph what they thought were the important data patterns for each experiment.

The Doctor Visit Materials Analyzed by the Students

Experiment 1

In the first experiment, researchers asked 6 participants to write down the events that occur when they visit the doctor. The results are shown here. Notice that the participants' prose has been reduced to each main idea each participant came up with. For example, if a participant wrote, "You enter the doctor's office," and another participant wrote, "You walk through the door of the doctor's suite," both would be simplified to "Enter office." Although this simplification may miss some subtle details and differences in the participants' protocols, it makes it much easier to compare ideas across different participants.

Participant 1: Enter office. Check in with receptionist. Sit down. Wait. Name called. Enter exam room. Sit on table. Doctor examines. Doctor asks questions. Make another appointment. Leave office.

Participant 2: Check in with receptionist. Read magazine. Look at other people. Name called. Sit on table. Nurse tests. Doctor examines. Leave office.

Participant 3: Check in with receptionist. Sit down. Read magazine. Talk to nurse. Nurse tests. Talk to doctor about problem. Leave office.

Participant 4: Enter office. Sit down. Read magazine. Enter exam room. Undress. Sit on table. Nurse tests. Doctor examines. Get dressed. Leave office.

Participant 5: Enter office. Check in with receptionist. Sit down. Read magazine. Name called. Follow nurse. Enter exam room. Nurse tests. Doctor enters. Doctor examines.

Participant 6: Check in with receptionist. Sit down. Read magazine. Nurse tests. Wait. Doctor greets. Doctor examines. Get medicine. Leave office.

Experiment 2

In the second experiment, 8 new participants read 12 passages about different topics. The paragraph of interest was the one describing John's visit to the doctor.

The Doctor Visit

John checked in with the doctor's receptionist. While he waited he read magazines. The nurse called his name. John undressed. John talked to the nurse. The doctor came in to the examination room. The doctor was very friendly. The doctor prescribed some pills for John. John left the doctor's office.

Twenty minutes after reading all the paragraphs, the participants were given a recognition test. Their task was to rate 12 sentences on a scale from 1 (*low*) to 7 (*high*) as to how sure they were that they had actually read the sentence. The following are the 12 sentences they had to rate, and the results are shown in Table A1.

A) John checked in with the doctor's receptionist.
B) John sat down.
C) While he waited he read magazines.
D) John followed the nurse.
E) John undressed.
F) John talked to the nurse.
G) The nurse tested John in the examination room.
H) The doctor greeted John.
I) The doctor examined John.
J) The doctor prescribed some pills for John.
K) John made another appointment.
L) John left the doctor's office.

TABLE A1
Participants' Ratings of the 12 Sentences

	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*
Participant 7	7	7	6	3	4	3	6	1	7	4	2	7
Participant 8	7	5	6	1	5	4	7	3	5	3	1	7
Participant 9	6	7	7	4	3	5	5	4	3	5	3	6
Participant 10	7	6	7	3	6	4	4	3	7	2	4	7
Participant 11	7	6	5	3	4	3	7	2	5	6	4	6
Participant 12	7	4	7	2	5	5	6	3	6	3	2	7
Participant 13	7	7	7	2	3	3	5	5	7	5	1	7
Participant 14	5	5	7	4	5	5	4	2	4	4	3	7

Experiment 3

In the third experiment, 7 new participants read 12 passages. Again the paragraph of interest involved going to the doctor. The passage was a bit different from the previous passage.

The Doctor Visit

John checked in with the doctor's receptionist. There were four chairs in the waiting room. While he waited he read magazines. The nurse called his name. John undressed. John put his clothes on a coat hanger. John talked to the nurse. The doctor came in to the examination room. The doctor was very friendly. The doctor prescribed some pills for John. John had forgotten his wallet. John left the doctor's office.

Twenty minutes after finishing the passages, the participants were told that they were to write down everything they could remember from the doctor visit paragraph. This is what the participants wrote down:

Participant 15:	John checked in with the receptionist. He sat down. He read magazines while he waited. The nurse called his name. John talked to the nurse. The doctor examined John. The doctor gave John a prescription. John forgot his money.
Participant 16:	John spoke to the receptionist. While he waited, he read magazines. The doctor examined him. John left his wallet at home. John left the doctor's office.
Participant 17:	John entered the office. He checked in with the secretary. He followed the nurse. He sat on the examination table. The nurse tested him. John did not have his wallet.
Participant 18:	John went to the receptionist. He sat down in the waiting room. The nurse called his name. John got undressed. John was tested by the nurse. John was examined by the doctor. John had forgotten his wallet. John left.
Participant 19:	John went to the doctor. He read magazines. The nurse called John. The nurse gave John some tests. The doctor came in. John did not have his wallet. John left the doctor's office.
Participant 20:	John entered the office. John checked with the receptionist. John sat in one of the four chairs in the office. He read periodicals. The nurse called him into the exam room. The doctor came into the examination room. The doctor examined John. John did not have his wallet.
Participant 21:	John went to the doctor. He forgot his wallet. He left the doctor's.

The Balloon Passage Materials Analyzed by the Students

In each of the three experiments, participants read the following passage and then were asked to remember the passage as well as they could.

> *If the balloons popped, the sound would not be able to carry. Everything would be too far away from the correct floor. A closed window would also prevent the sound. This is because most buildings tend to be well insulated. The whole operation depends on a steady flow of electricity. A break in the middle of the wire would also cause problems. Of course, the fellow could shout. But the human voice is not loud enough to carry that far. An additional problem is that the string could break on the instrument. Then there could be no accompaniment to the message. It is clear that the best situation would involve less distance. Then there would be fewer potential problems. With face to face contact, the least number of things could go wrong.*

Experiment 1

In the first experiment, 8 college students read the passage with 13 sentences. These participants were told that they would be tested for their memory of what they had read. Four of the participants were in the *immediate recall* condition. Immediately after reading the passage, the participants were told to write down everything they could remember from the passage. The other 4 participants were in the *delayed recall* condition. After they read the passage, they had to count back from 100 by 3s (e.g., 100, 97, 94 …). After this delay, these 4 participants were also instructed to write down everything they could recall from the passage. Here is what they remembered.

Immediate recall condition.

Participant 1: With face to face contact, the least number of things could go wrong. Then there would be fewer potential problems. If the balloons popped, the sound would not be able to carry. The human voice is not loud enough to carry. A closed window would also prevent the sound from carrying.

Participant 2: With face to face contact, the least number of things could go wrong. It is clear the best situation would involve less distance. If the balloons popped, the sound would not be able to carry. The whole operation depended on a steady flow of electricity.

Participant 3: With face to face contact, the least number of things could go wrong. If the balloons popped, the sound would not be able to carry. A string could break on the instrument. Most buildings tend to be well-insulated.

Participant 4: Then there would be fewer potential problems. With face to face contact, the least number of things could go wrong. If the balloons popped, the sound would not carry. Then there could be no accompaniment to the message. A break in the wire would cause problems.

Delayed recall condition.

Participant 5: If the balloons popped, the sound would not be able to carry. Everything would be too far away from the correct floor. It depends on a steady flow of electricity.

Participant 6: If the balloons popped, the sound would not be able to carry. Of course, the fellow could shout.

Participant 7: A closed window would prevent the sound from carrying. If the balloon popped, the sound would not carry. With face to face contact, the least number of things could go wrong. The human voice is not loud enough to carry that far.

Participant 8: If the balloons popped, the sound would not be able to carry. Everything would be too far from the correct floor. There would be no accompaniment to the message.

Experiment 2

In the second experiment, 8 new college students read the same passages. There were also two conditions, although one was a bit different from before. Four of the participants read the passage once as before. This is called the *no-context* condition. The other 4 participants also read the passage once. However, after they read the passage, these 4 participants saw the picture shown in Figure A1. This picture gives the passage a context. This is called the *context-after* condition because they saw the picture after reading the passage. After completing their tasks, participants in both conditions tried to solve a division problem in their head (i.e., 2,961 ÷ 9). Then, they were told to write down as much of the passage as they could remember. Here is what the participants remembered.

No-context condition.

Participant 9: If the balloons popped, the sound would not be able to carry. Everything would be too far away from the correct floor.

Participant 10: A closed window would also prevent the sound. With face to face contact, the least number of things could go wrong. If the balloons popped, the sound would not be able to carry.

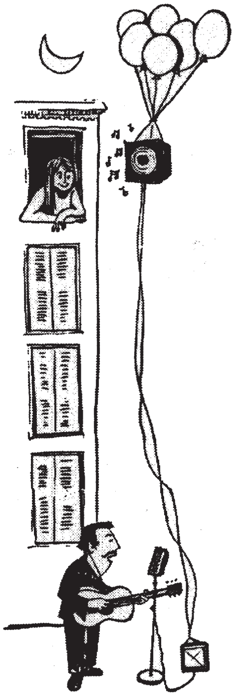


FIGURE A1 The picture seen by participants in the context-after and context-before conditions. (From Bransford & Johnson, 1972. Copyright 1972 by Academic Press.)

Participant 11: The string could break on the instrument. If the balloons popped, the sound would be prevented from carrying. It all depends on a steady flow of electricity. A break in the middle of the wire could cause problems.

Participant 12: The least number of things could go wrong. There would be fewer potential problems. The human voice is not loud enough.

Context-after condition.

Participant 13: If the balloon string breaks, it would also cause problems. The fellow could shout. His voice could not carry up the building.

Participant 14: The fellow is too far from the correct floor. He needed electricity. The wire could not break.

Participant 15: If the balloons popped, there would be problems. The guitar strings might break.

Participant 16: If he could be face to face, less things would go wrong. The electricity might stop. The balloons could pop. He might have to shout.

Experiment 3

In the final experiment, another group of 8 college students read the passage. This time, 4 of the participants read the passage twice without ever seeing the picture. This is called the *double encode* condition because the participants were able to encode (study) the passage twice. The other 4 participants looked at the picture before they read the passage. This is called the *context-before* condition. After the participants finished reading the passages, they counted backwards from 100 by 3s. Then, they were instructed to write down as much of the passage as they could remember. Here is what they recalled.

Double encode condition.

Participant 17: If the balloons popped, the sound would not be able to carry. Everything would be too far away from the correct floor. A closed window would also prevent the sound.

Participant 18: If the balloons popped, the sound would not be able to carry. Of course, the fellow could shout. But the human voice is not loud enough to carry that far. It is clear that the best situation would involve less distance.

Participant 19: With face to face contact, the fewest number of things could go wrong. The whole operation depends on a steady flow of electricity. If the balloons popped, the sound would not carry.

Participant 20: Everything would be too far from the correct floor. If the balloons popped, the sound would not be able to carry. The string could break on the instrument. Most buildings tend to be well insulated.

Context-before condition.

Participant 21: If the balloons pop, the sound won't reach the window. It would be too far away from her floor. Also, she might have a closed window. He also needs a good flow of electricity to the speaker. If the wire breaks, there will be problems. If his guitar strings break, he will also have problems. It would be much better if he could be closer. Face to face contact would reduce the number of things that could go wrong.

Participant 22: It is clear that it would be better if he could get closer to her. If the balloons pop, the sound won't be able to reach from the ground. It would be too far from the correct window. Even if the balloons don't pop, her window might be closed. And the electrical wire might break. Of course, the fellow could shout. But the human voice is not loud enough to carry that far. And what if his instrument's string breaks.

Participant 23: If the balloons pop on his way there, he won't be able to lift the speaker. Then he would be too far away. It would be much safer if he could be at a closer distance. He could try to shout. But his voice won't be loud enough to reach the window. If he could sing to her face, less things go wrong like breaking his guitar.

Participant 24: His balloons might pop and then he couldn't get the sound to her window. Or, the wire up to the speaker might break. The whole thing depends on a steady flow of electricity. Maybe she'll have a closed window. This would prevent the sound since most buildings tend to be insulated. Then he could not shout, although his voice won't be loud enough anyway. If he could be next to her, less things could go wrong. But an additional problem is whether his instrument will break a string. Then there would be no accompaniment for his voice.