Iterative Refinement Cycles for Videotape Analyses of Conceptual Change

Richard Lesh
Purdue University

Richard Lehrer
University of Wisconsin–Madison

Note: Before reading this chapter on iterative videotape analyses, readers may find it useful to read the transcript that is given in the Appendix to the chapter. The transcript is one of those that is referred to in the chapter; and, many of the chapter’s main points make reference to it. In this regard, the following comments are especially relevant:

• The students’ goal for this problem solving session was to produce a description of the information that was given; and, the purpose of this description was to enable the students to rank workers whose summer jobs involved selling things in a park.

• The students’ initial way of thinking about the information was quite barren and distorted; that is, they neglected to notice a great deal of relevant information, and, at the same time, they made assumptions that were not valid.

• To gradually arrive at a final way of thinking about the situation, the students went through a series of modeling cycles in which they interpreted givens and the goals in systematically different ways that focused on different facts, quantities, relationships, and patterns.

• The researchers’ goals were to produce a description of how the students’ ways of thinking evolved during the problem solving session. Consequently, the researchers also arrived at a final interpretation of the session by going through a series of modeling cycles in which their early ways of thinking were remarkably barren and distorted compared with those that evolved later.

The following issues are useful to think about when reading the transcript. What do you think was the researcher’s perceived purpose for the description that he was generating? Is it likely that another purpose or another perspective might have emphasized different principles for selecting and organizing information, or for interpreting their significance? In what ways did the researchers test the usefulness of alternative interpretations? What methods did they use to determine whether a given interpretation needed to be refined, modified, or extended? If the researchers undertook the job of producing a movie to convey to audiences what was going on in the problem solving session, do you think that the best kind of movie would consist of nothing more than the videotape of the session?

This chapter about iterative videotape analyses is a companion to two others in this book. The first is about model-eliciting activities (Lesh, Hoover, Hole, Kelly, & Post, chap. 21, this volume); and, the second is about multitiered teaching experiments (Lesh & Kelly, chap. 9, this volume). In general, all three chapters focus on research whose central purpose is to investigate the nature of the constructs that students (or teachers or groups) develop to make sense of a targeted class of problem solving or decision-making situations. Also, all three chapters emphasize research designs that are potentially shareable; that is, it is possible to coordinate the work of several researchers who are working at multiple sites using diverse practical or theoretical perspectives.

Whereas the chapter on model-eliciting activities (chap. 21, this volume) describes principles for designing problem solving situations in which observations are to be made, the chapter on multitiered teaching experiments (chap. 9, this volume) focuses on principles for designing the data collection stages of research, and this chapter on iterative videotape analyses focuses on the stages of research that involve data analysis. Also, the chapter on multitiered teaching experiments concentrates on studies in which conclusions based on today’s sessions influence tomorrow’s data collection activities; whereas, this chapter on iterative videotape analyses focuses on research settings in which the most significant stages of data collection must be completed before the most significant stages of data interpretation begin. Therefore, in this latter case, there exist no severe constraints on the turnaround times for data interpretation.
FIG. 23.1 focuses on the procedures that are used to test, refine, and extend their interpretations of their data. It suggests one important way that videotape analyses sometimes need to be quite different depending on the types of context in which they occur. For example, an essential characteristic of many teaching experiments is that it is for each successive videotaped session to be designed partly to test the validity of researchers’ interpretations of students’ responses during similar sessions that occurred earlier. Therefore, an iterative series of hypothesis-testing or decision-making cycles may enable researchers to validate (test, refine, and extend) their own constructs gradually.\(^1\)

In contrast to the preceding kinds of situations, an essential characteristic of many studies involving videotape analyses is that significant portions of the data interpretation stages of the project cannot begin until after many of the most significant data-gathering stages have been completed. Therefore, tentative interpretations of one videotaped session cannot influence the problems that students confront in succeeding videotaped sessions, and there may exist no effective ways for researchers to test their interpretations by making predictions from one episode to the next. Consequently, researchers’ interpretations must be tested and refined in other ways. One goal of this chapter is to examine some of the most important of these other possible forms of construct validation; and, a second goal is to describe some ways that it is possible for a single

---

\(^1\) Similarly, an essential characteristic of many clinical interviews is that researchers’ interpretations of students’ responses to one question often influence the next question that is asked. Again, a series of such questions provides a series of rapid hypothesis-testing and decision-making cycles in which researchers test, refine, and extend their own constructs gradually.
research project to integrate procedures that associated with teaching experiments, clinical interviews, ethnographic observations, or videotape analyses. For example, FIG. 23.2 describes one way that iterative videotape analyses may be integrated with approaches to research that are emphasized in other chapters of this book.

FIG. 23.2. Activity in one area of study often stimulates action in another area, and all of them can occur in the same study.

Because the research design issues that we want to emphasize in this chapter are easier to describe in the context of specific studies, we’ll refer to the problem solving session that’s transcribed in the appendix to this chapter. It involves three-person teams of students who are working together on 60-minute “thought-revealing activities” of the type described in chapter 21 in this book. Nonetheless, the research design principles that are discussed are not restricted to this context; they are relevant to videotape analyses of a broad range of teaching and learning situations.2

THE ROLE OF VIDEOTAPES DEPENDS ON OTHER APPROACHES USED WITH THEM

Sometimes, it is said that a researcher who is viewing videotapes may not see the forest because of the trees. Alternatively, if the researcher is too narrowly predisposed to view situations through one particular theoretical window, then other types of details or perspectives may be ignored systematically. For some of these kinds of prejudices, negative effects can be minimized if videotape analyses occur in research designs that also include ethnographic observations, clinical interviews, teaching experiments, and other methods of data collection and/or data analysis. Furthermore, the roles that videotaping is expected to play depends partly on the strengths and weaknesses of the other approaches that are employed. By seeking agreement among the interpretations that are generated from these alternative approaches, it sometimes is possible to avoid biases that would result if only a single method were used. For example:

- Videotaping may be part of a strategy to focus on perspectives or details whose significance can be anticipated. Suppose that videotape analyses are used in conjunction

2 Studies involving model-eliciting activities are of interest mainly because, in such problems, the objective for students is not merely to produce answers to questions in which the givens and goals are stated in a form that is intended to require little interpretation. Instead, when teams of students work on model-eliciting activities, a large part of their objective is to construct an explicit mathematical interpretation (description, explanation, and justification) that typically must go through several refinement cycles in which trial constructs are tested, revised, and improved repeatedly for a given purpose. Therefore, because constructs and thought processes are externalized and made explicit, students tend to reveal the nature of their mathematical constructs in a straightforward manner. Videotapes of such sessions often provide direct evidence about the nature of the constructs that students use as well as the mechanisms that contribute to the development of those constructs.
with ethnographic observations of complex situations in which the range of possible responses cannot be anticipated. In such situations, it is often wise for the attention of an on-the-scene observer to be free to adjust to circumstances, to focus attention here or there as needs arise, and to zoom in and out in ways that are judged appropriate. But, in such circumstances, even though the collection and coding of data cannot be reduced to an algorithm, it is nonetheless important for the holistic judgments of experts to be governed by ground rules that enable later judgments to be made about why one type of information was emphasized and another type was ignored. Also, in such circumstances, videotapes may be useful to provide stable and unchanging perspectives that focus on aspects of the situation that can be predicted to be important (such as particular details or particular points of view). For instance, in group problem solving activities, one preplanned videorecording may focus on a close-up view of individual students’ desktop work, whereas another may focus on body language and other factors for the group as a whole. In this manner, alternative “windows on reality” may enable several diverse perspectives to be juxtaposed for easy comparison during after-the-fact analyses.

- **Videotaping may be part of a strategy to broaden perspectives or details beyond those whose significance can be anticipated.** For instance, suppose that videotape analyses are used in conjunction with clinical interviews in which carefully preplanned and highly structured sequences of standardized questions are used. In such circumstances, videotapes may enable a variety of unforeseen factors to be recorded and considered that might not have been recognized otherwise. For example, videotapes of children’s faces frequently provide clear evidence about the importance of certain aspects of their responses, even in cases where these factors were not anticipated in preplanned questioning and recording schemes.

The point of view of research design, one danger with either of the preceding focusing or broadening strategies is that it is easy for researchers to imagine that videotapes represent complete and unbiased recordings of everything that occurred during given sessions. The truth is otherwise. Every time a videocamera focuses on one thing, it tends to de-emphasize or ignore something else, and, in general, videotapes are poorly suited to record certain types of information. For example, when three-person teams of students are engaged in 60-minute group problem solving sessions, experience shows that an on-the-scene human observer almost always notices some things (such as an uncomfortable temperature in the room or tension among members of the group) that may not be apparent in the videotapes of the session. Conversely, even when on-the-scene experts are available during videotaped sessions, they nearly always fail to notice some things that are recognized later to be very significant in the videotapes of the same sessions (perhaps because the significance of an event is not clear until its consequences are apparent). The problem of judging importance becomes even more complex for long-term studies of conceptual change. This is partly due to the difficulty of judging the significance of an event until its consequences become clear, as they unfold and emerge over time. Therefore, as information emerges about the history that is relevant to a given classroom activity, video backtracking may play an important role in establishing an adequate framework for interpretation.

Sometimes, a picture is worth a thousand words; and a moving picture sometimes has the power to reveal patterns that would not be apparent in the snapshots that they contain. For instance, videotapes often make it possible to observe changes across time (by enabling researchers to go back and forth among related segments, even in cases where the relatedness could not have been apparent to on-the-scene observers). Also, videotapes sometimes offer the possibility of viewing segments multiple times from multiple perspectives, as well as the possibility of juxtaposing segments that occur at different times (so that patterns, regularities, or trends may become apparent that would tend to be overlooked if only isolated sessions were viewed). However, even though videotapes sometimes approach the ideal of preserving raw data so that it can be analyzed multiple times from multiple perspectives, it also is true that some data interpretation always occurs each time that information is filtered, selected, simplified, or organized. Consequently, when information is filtered out and noncontiguous segments are juxtaposed, the rationales that are used to make selections, deletions, or comparisons should be explained. There is a need to make explicit the interpretation framework that is used to fill in gaps or to focus on potential relationships in the filtered data. Otherwise, concealed biases are inevitable.

To accomplish the preceding goals in a given investigation, one important step is to specify which purposes and functions videotape analyses are intended to fill. Then, steps should be taken to address these functions. For example, simply because multiple viewing is possible does not ensure that it will take place; and, simply because videotape analyses make it possible for subsequent examinations to reveal diverse and
unforeseen factors in students’ complex performances sometimes, this does not eliminate the need for careful preplanning.

To identify several types of preplanning that are needed in the case of videotape analyses, consider the kind of videotaping that is used habitually in assessments of preservice teachers’ undergraduate, student-teaching assignments. In such settings, the following issues arise commonly:

- Which episodes should be recorded? For example, the characterization that emerges of a given teacher’s abilities often depends on whether the videotaped episodes emphasize only (i) teacher-centered lectures (in which the goal may be for students to try to follow the teacher’s thinking) or (ii) student-centered discussions (in which the goal may be for the teachers to follow the students’ thinking).
- What perspective(s) or viewpoints should be emphasized? For example, a bird’s-eye view of a classroom situation often provides very different information than a teacher’s-eye view (or a given student’s-eye view) of the same situation.
- Which details should be emphasized in any given perspective? For example, a video segment in which the teacher is not visible may provide information about teaching ability that is more revealing than a video segment in which the teacher is the center of attention.
- What “grain size” or unit of analysis is most useful for seeing patterns and regularities that underlie discrete pieces of information? For example, when a video segment “zooms in” on one student who is being “turned on” during a given student–teacher interaction, it may neglect to notice that 29 other students are being “turned off” by the same interaction.

By considering the preceding kinds of issues, it is clear that, to formulate the design component of a research proposal, it is not enough merely to state that an investigation will involve videotape analyses (or clinical interviews or ethnographic observations) of a particular type of situation. Furthermore, issues of validity, reliability, and generalizability do not disappear because qualitative analysis procedures are used. For example, the meaning of a given videotaped session often depends on other sessions that surround it. In this context, the following questions need to be answered:

- How typical is a session that is chosen for special attention?
- How is one session influenced by experiences in earlier sessions?
- What aspects of an isolated session only become significant in the light of sessions that followed?

Unless answers are available to these kinds of questions, analyses of isolated sessions often are misleading. Furthermore, this may be true regardless of how thoroughly and conscientiously isolated videotapes are analyzed as independent entities.

**SOME ISSUES TO CONSIDER TO IMPROVE THE QUALITY OF VIDEOTAPE ANALYSES**

At their worst, proposals to conduct videotape analyses sometimes are accused of being “trust me” studies, with little effort made to assure reliability and validity. The claim that video vignettes “speak for themselves” is often given as a justification for such approaches. But, the apparent ability of video vignettes to speak for themselves tends to be an illusion; video draws its power from the interpretive framework established by researchers. In particular, researchers may find it useful to consider the following questions.

- What steps were taken to ensure that videotaped segments that are given special attention constitute a representative or unbiased sample?
- Did the researchers’ interpretations go through any testing and revision cycles? If so, what types of tests were used?
- What refinement cycles were used to filter, select, weigh, organize, and code the data that were collected?
- What cross-checking or triangulation techniques were used to ensure that information from alternative viewpoints fit together to produce stable and consistent interpretations of results?

To capitalize on the potential strengths of videotape analyses, as well as to avoid their potential weaknesses, researchers should plan acceptable procedures to document and cross-check their responses to decision-making issues that arise during data collection and data interpretation. Yet, because the transcription
and analysis of videotapes are expensive and time consuming, many projects end up collecting far too many videotapes and spending far too little time interpreting the data they yield. The result is that the ratio of videotaping hours to video-analysis hours becomes many-to-one rather than one-to-many, and the lure of capturing the ultimate video clip sometimes seems to become a fever for researchers that is similar to the lure of gold in the Klondike. Therefore, the next section of this chapter describes some specific techniques to avoid such occurrences.

**PROCEDURES FOR TESTING, REFINING, AND EXTENDING INTERPRETATIONS OF VIDEOTAPES**

To describe appropriate procedures for testing, refining, and extending interpretations of videotaped research sessions, it is useful to focus on a particular example, and it is useful for this example to be a typical project that involves most of the types of procedures that we want to discuss. However, whenever a particular case is given special attention, it is important to emphasize that this does not imply that the underlying issues and procedures could not have been handled in other ways. On the contrary, nearly always a variety of options is available.

The study described next lasted for an entire semester (10 to 16 weeks); the participants included an entire classroom of 24 to 30 students; and, the videotaped sessions involved 60-minute problem solving sessions in which the students worked in three-person groups that had access to the appropriate tools and resources. This study is typical in the sense that it involves many of the kinds of issues that arise frequently in mathematics and science education research. But, it is not typical in the sense that support from the National Science Foundation made it possible to involve a larger number of levels and types of interacting research staff members than are available in most published studies of mathematics and science education (Lesh, 1979). Table 23.1. illustrates a typical scheme for collecting data within such studies.

**TABLE 23.1**
A Typical Sixteen-Week Study Involving Eight, Three-Person Groups of Students and Eight Pairs of Problem Solving Situations (A$_1$-A$_2$, B$_1$-B$_2$, ..., H$_1$-H$_2$)

<table>
<thead>
<tr>
<th>Week</th>
<th>Group I</th>
<th>Group II</th>
<th>Group III</th>
<th>Group IV</th>
<th>Group V</th>
<th>Group VI</th>
<th>Group VII</th>
<th>Group VIII</th>
<th>Substitute Groups</th>
<th>Substitute Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Prob $A_1$</td>
<td>Prob $B_1$</td>
<td>Prob $C_1$</td>
<td>Prob $D_1$</td>
<td>Prob $E_1$</td>
<td>Prob $F_1$</td>
<td>Prob $G_1$</td>
<td>Prob $H_1$</td>
<td>Prob $A_1$</td>
<td>Prob $B_2$</td>
</tr>
<tr>
<td>2</td>
<td>Prob $A_2$</td>
<td>Prob $B_2$</td>
<td>Prob $C_2$</td>
<td>Prob $D_2$</td>
<td>Prob $E_2$</td>
<td>Prob $F_2$</td>
<td>Prob $G_2$</td>
<td>Prob $H_2$</td>
<td>Prob $A_2$</td>
<td>Prob $B_2$</td>
</tr>
<tr>
<td>3</td>
<td>Prob $A_3$</td>
<td>Prob $B_3$</td>
<td>Prob $C_3$</td>
<td>Prob $D_3$</td>
<td>Prob $E_3$</td>
<td>Prob $F_3$</td>
<td>Prob $G_3$</td>
<td>Prob $H_3$</td>
<td>Prob $A_3$</td>
<td>Prob $B_3$</td>
</tr>
<tr>
<td>4</td>
<td>Prob $A_4$</td>
<td>Prob $B_4$</td>
<td>Prob $C_4$</td>
<td>Prob $D_4$</td>
<td>Prob $E_4$</td>
<td>Prob $F_4$</td>
<td>Prob $G_4$</td>
<td>Prob $H_4$</td>
<td>Prob $A_4$</td>
<td>Prob $B_4$</td>
</tr>
<tr>
<td>5</td>
<td>Prob $A_5$</td>
<td>Prob $B_5$</td>
<td>Prob $C_5$</td>
<td>Prob $D_5$</td>
<td>Prob $E_5$</td>
<td>Prob $F_5$</td>
<td>Prob $G_5$</td>
<td>Prob $H_5$</td>
<td>Prob $A_5$</td>
<td>Prob $B_5$</td>
</tr>
<tr>
<td>7</td>
<td>Prob $A_7$</td>
<td>Prob $B_7$</td>
<td>Prob $C_7$</td>
<td>Prob $D_7$</td>
<td>Prob $E_7$</td>
<td>Prob $F_7$</td>
<td>Prob $G_7$</td>
<td>Prob $H_7$</td>
<td>Prob $A_7$</td>
<td>Prob $B_7$</td>
</tr>
<tr>
<td>8</td>
<td>Prob $A_8$</td>
<td>Prob $B_8$</td>
<td>Prob $C_8$</td>
<td>Prob $D_8$</td>
<td>Prob $E_8$</td>
<td>Prob $F_8$</td>
<td>Prob $G_8$</td>
<td>Prob $H_8$</td>
<td>Prob $A_8$</td>
<td>Prob $B_8$</td>
</tr>
<tr>
<td>9</td>
<td>Prob $A_9$</td>
<td>Prob $B_9$</td>
<td>Prob $C_9$</td>
<td>Prob $D_9$</td>
<td>Prob $E_9$</td>
<td>Prob $F_9$</td>
<td>Prob $G_9$</td>
<td>Prob $H_9$</td>
<td>Prob $A_9$</td>
<td>Prob $B_9$</td>
</tr>
<tr>
<td>10</td>
<td>Prob $A_{10}$</td>
<td>Prob $B_{10}$</td>
<td>Prob $C_{10}$</td>
<td>Prob $D_{10}$</td>
<td>Prob $E_{10}$</td>
<td>Prob $F_{10}$</td>
<td>Prob $G_{10}$</td>
<td>Prob $H_{10}$</td>
<td>Prob $A_{10}$</td>
<td>Prob $B_{10}$</td>
</tr>
<tr>
<td>11</td>
<td>Prob $A_{11}$</td>
<td>Prob $B_{11}$</td>
<td>Prob $C_{11}$</td>
<td>Prob $D_{11}$</td>
<td>Prob $E_{11}$</td>
<td>Prob $F_{11}$</td>
<td>Prob $G_{11}$</td>
<td>Prob $H_{11}$</td>
<td>Prob $A_{11}$</td>
<td>Prob $B_{11}$</td>
</tr>
<tr>
<td>12</td>
<td>Prob $A_{12}$</td>
<td>Prob $B_{12}$</td>
<td>Prob $C_{12}$</td>
<td>Prob $D_{12}$</td>
<td>Prob $E_{12}$</td>
<td>Prob $F_{12}$</td>
<td>Prob $G_{12}$</td>
<td>Prob $H_{12}$</td>
<td>Prob $A_{12}$</td>
<td>Prob $B_{12}$</td>
</tr>
<tr>
<td>13</td>
<td>Prob $A_{13}$</td>
<td>Prob $B_{13}$</td>
<td>Prob $C_{13}$</td>
<td>Prob $D_{13}$</td>
<td>Prob $E_{13}$</td>
<td>Prob $F_{13}$</td>
<td>Prob $G_{13}$</td>
<td>Prob $H_{13}$</td>
<td>Prob $A_{13}$</td>
<td>Prob $B_{13}$</td>
</tr>
<tr>
<td>14</td>
<td>Prob $A_{14}$</td>
<td>Prob $B_{14}$</td>
<td>Prob $C_{14}$</td>
<td>Prob $D_{14}$</td>
<td>Prob $E_{14}$</td>
<td>Prob $F_{14}$</td>
<td>Prob $G_{14}$</td>
<td>Prob $H_{14}$</td>
<td>Prob $A_{14}$</td>
<td>Prob $B_{14}$</td>
</tr>
<tr>
<td>15</td>
<td>Prob $A_{15}$</td>
<td>Prob $B_{15}$</td>
<td>Prob $C_{15}$</td>
<td>Prob $D_{15}$</td>
<td>Prob $E_{15}$</td>
<td>Prob $F_{15}$</td>
<td>Prob $G_{15}$</td>
<td>Prob $H_{15}$</td>
<td>Prob $A_{15}$</td>
<td>Prob $B_{15}$</td>
</tr>
<tr>
<td>16</td>
<td>Prob $A_{16}$</td>
<td>Prob $B_{16}$</td>
<td>Prob $C_{16}$</td>
<td>Prob $D_{16}$</td>
<td>Prob $E_{16}$</td>
<td>Prob $F_{16}$</td>
<td>Prob $G_{16}$</td>
<td>Prob $H_{16}$</td>
<td>Prob $A_{16}$</td>
<td>Prob $B_{16}$</td>
</tr>
</tbody>
</table>

In the study depicted in Table 23.1, it was possible to compare the behaviors of multiple groups students working on each of the problems: $A_1, B_1, \ldots, H_1; A_2, B_2, \ldots, H_2$. Furthermore, because several different teams of students worked on each problem, it also was possible to compare the behaviors of students who worked on a given problem at different times during the sixteen-week study; and, because the 16 problems
consisted of 8 pairs of structurally similar problems, and because the teams of students remained invariant over the entire 16-week study, it was possible to observe the development of groups across time.

Of course, in similar research investigations, more or fewer groups might have been involved; the study might have been conducted over a period of more or fewer weeks; the number of students in each group might have been more or fewer; and, other factors (such as the nature and length of the sessions that were videotaped) might have varied. But, for the purposes of this chapter, the basic structure of the study would remain essentially unchanged regardless of whether the episodes being videotaped consisted of clinical interviews involving individual students, problem solving sessions involving small groups of students, or teacher-led activities involving entire classrooms of students.

When planning a similar investigations, researchers must make choices about the following kinds of issues.

- **Practical** considerations involve the following kinds of issues. What is the purpose of the observations that are being made? For example, an on-the-scene observer might concentrate on collecting small scraps of paper, or on other physical or anecdotal evidence, that would be unlikely to be captured by other recording devices. An analysis of a typed transcript (which is based on a videotape) might focus on trends and patterns that are difficult to observe when videotapes are viewed sequentially. A meta-analysis might look for generalizations that occur in the behavior of a given group of students across a number of related videotaped sessions, or it might look for similarities and differences among the behaviors of several different groups across similar situations.

- **Physical** considerations involve the following kinds of issues. How many video cameras will be used? What will each camera focus on? For example, if small-group problem solving sessions are being videotaped, one camera might focus on students’ desktop work and another on faces or gestures for the group as a whole. Still another perspective might be embodied in the notes of on-the-scene observers.

- **Theoretical** considerations involve the following kinds of issues. What kinds of entities or events and what kinds of relationships and patterns will be given special attention? For example, in a given physical perspective, a psychological perspective might concentrate on the changing roles of individuals within groups, a teaching perspective might converge on facts and skills that promote or inhibit progress, and a problem solving perspective might highlight the representational issues or the modeling cycles that students go through during the solution of given problems.

- **Temporal** considerations involve the following kinds of issues. What counts as an episode? A lesson? What is the period of time that is of interest? If the point of the analysis is to understand the evolution of students’ epistemology or significant changes in ontological structure, then often a long time span is indicated. If, however, the interest is in characterizing local conceptual development, especially changes in strategies or slight revisions in models, then the sampling of episodes important for understanding change may be from minute to minute or problem to problem (e.g., Siegler & Jenkins, 1989).

As FIG. 23.3 suggests, it is productive to analyze videotapes from multiple standpoints such as: theoretical, physical, and practical. Usually, each perspective yields somewhat different information. Therefore, a primary way to test and refine interpretations is to go through a series of triangulation and consensus-building cycles so that descriptions based on multiple viewpoints can be compared to identify points of agreement and disagreement, and efforts can be made to reach a stable consensus.

---

1. During a given session, half of the teams worked on one problem from the pair, and half of the teams worked on the other problem from the pair. Then, later in the project, each group worked on the problem that they had not addressed earlier.

2. Students in the substitute groups were used to fill in for other students who might miss class because of illness or other reasons. Therefore, for the research design that is illustrated in Table 23.1, the 30 students were divided into 24 regular participants and 6 substitutes.
To implement a scheme similar to the one described in Figure 23.3, it often is useful to use a series of telescoping analyses that begin by giving equal attention to all of the sessions (using procedures that are necessarily more superficial and less time consuming) and that end by focusing on deeper and more time consuming analyses of a relatively small number of specially selected sessions (whose criteria for selection and weighting are specified clearly and whose representativeness is described in an organizational scheme that is identified clearly also). Or, in general:

| Begin with analyses of isolated sessions. | Evolve gradually toward analyses that involve a collection of related sessions. |
| Begin by giving equal attention to issues and events that apply to all of the sessions. | Evolve gradually toward special attention to sessions that prove to be exemplary or illuminating for specific purposes that emerge gradually as being especially significant. |
| Begin by emphasizing information that can be judged to be significant without examining patterns and regularities that cut across several sessions (or that cut across several stages within a given session). | Evolve gradually toward emphasizing information whose significance becomes apparent only when attention is directed toward deeper phenomena. |

To conduct the preceding type of telescoping analyses within a 16-week study similar to the one described at the beginning of this section, the following interpretation cycles can be expected to occur.

**Interpretation Cycle 1: Begin With Debriefing Forms From the On-the-Scene Observers.** During each videotaped session, it is useful for an on-the-scene observer to be available to:

- Take notes.
- Keep the session from getting derailed by irrelevant factors (such as squabbles over who uses the calculator or other resources) or going off on unrelated tangents (such as those that occur when students make unintended assumptions about the problem solving situation).
- Ensure that the videotape equipment is working properly.
- Collect information that might get lost.

The on-the-scene observer may be a researcher, a teacher, or a graduate student in mathematics education, psychology, or another relevant field. Experience shows that on-the-scene observers often notice a great many things that go unremarked if only videotaped records are preserved. Therefore, as soon as possible after each
session, it is wise for each on-the-scene observer to prepare a brief report summarizing his or her observations and appending any relevant documentation that might be lost or neglected. To get the most out of records from on-the-scene observers, careful preplanning is needed in order to anticipate the kinds of information that it may be most important to emphasize (e.g., there is no need for on-the-scene observers to focus on factors that will be recorded adequately using videotapes or results that students will produce). As mentioned already, it is important for debriefing sessions to occur as rapidly as possible following the event; otherwise, the kind of information that is distinctive to this perspective often gets misplaced or forgotten.

Of course, there are risks associated with giving too much emphasis to rapidly produced conclusions that are based on factors that are most salient to the on-the-scene observers, but the same could be said of any perspective that is produced at any time. No perspective has exclusive access to “the truth” about a given session. Yet, even if a perspective proves to be highly inaccurate in the light of later analyses, it still may contribute in positive ways to how the session is interpreted later. For example, in research on model-eliciting activities (Lesh & Akerstrom, 1982), it was discovered that the sessions that the on-the-scene observers considered to be most or least interesting at first seldom turned out to be the ones that proved to be most interesting based on deeper and more thoughtful analyses. Nonetheless, fresh firsthand assessments tended to provide powerful counterpoints to impressions that evolved from other later analyses, and they often provided important ways to cross-check the validity of interpretations generated by more remote observers. It is not necessary for later interpretations to agree with the rapidly produced impressions of the on-the-scene observers; however, in general, it does tend to be important for later interpretations to account for these observations and to offer alternative explanations.

**Interpretation Cycle 2: On-the-Scene Observers also Produce Observations of a Videotape for a Second Session.** While the results of a given session are still fresh in the minds of the on-the-scene observers, it is useful for each of them to generate a second debriefing form of at least one other session in addition to the one for which he or she was the on-the-scene observer. Furthermore, it also is useful for each session to be interpreted independently by at least two observers (one interpretation based on the comments produced by an on-the-scene observer and one based on comments produced by an observer of the videotape). When on-the-scene observers produce debriefing forms for at least two separate sessions, they tend to notice similarities and differences that stimulate observations that lead to productive refinements in both reports. Furthermore, when each session is interpreted independently by at least two separate observers, each observer often receives valuable feedback showing that how he or she interpreted students’ actions is seldom the only explanation that could be given. Therefore, this feedback may provide valuable input that can be used as cross-checks for later rounds of transcript interpretation.

**Interpretation Cycle 3: Produce Written Transcripts and an Executive Summary for a Given Session.** In lengthy, complex, videotaped episodes, it may be hard to discern some patterns and trends unless the sweep of an entire session can be skimmed rapidly so that related segments can be compared in ways that tend to be difficult even when it is possible to fast-forward through a videotape. For example, consider the annotated transcript that is given in the Appendix to this chapter. This transcript is based on a 60-minute problem solving session that involved three, average-ability, middle school students. By reading this transcript and the comments that are interspersed, it should be clear how it becomes possible to see certain trends and comparisons that would have been difficult to notice otherwise. For example, note the successive interpretations of the students’ conceptions in the Appendix, beginning with an initial interpretation of their activity as uncoordinated and relatively disjointed, followed by a focus on a single quantity (the total number of hours worked), and so on. In addition, note that the transcript not only is interpreted with respect to the activity of a single group but also is contextualized by comparison to other groups.

In some transcripts of this nature, the first column often ends up looking similar to the kind of script that tells actors what to do and say in a theatrical performance; the second column often includes relevant snippets of students’ work and other relevant information (e.g., pictures, diagrams, tables, and notes). Then, in the later stages of transcript analysis, more columns can be added that give comments and interpretations from researchers representing a variety of perspectives and that focus on such themes as descriptions of:

- The mathematical modeling cycles that the students went through as they generated a response to the given problem.
- The metacognitive roles that various students played involving such activities as planning, monitoring, and assessment.
- Individual differences, such as those associated with tendencies for impulsiveness or excessive avoidance of risk taking.
- Problem solving strategies that the students used, such as those that are emphasized frequently in classroom instruction.
Because annotated transcripts can facilitate researchers to make comparisons involving the preceding kinds of themes, such transcripts often stimulate a level of analysis that would remain obscured to viewers of uninterpreted videotapes otherwise. Unfortunately, typed and annotated transcripts are very time consuming and expensive to produce. Therefore, it is important to take advantage of the fact that it is seldom necessary to use expensive, high-level, staff members to produce transcripts that are useful. Furthermore, it may not be necessary to produce polished and annotated transcripts for all of the sessions in a project. For example, one purpose of using telescoping methods of data analysis is that screening and weighting procedures can be used to enable researchers to focus their attention progressively on sessions that emerge as being the most interesting.

Once typed transcripts of a session have been made and debriefing forms have been produced by an on-the-scene observer and by a videotape observer, it is necessary for a member of the project’s staff to view the relevant videotape and to ensure that times and quotations have been recorded accurately in the transcript, events that seem to be especially significant have been highlighted and described briefly, and relevant pictures, diagrams, and notes have been appended appropriately. This person can be one of the observers who were involved during Cycles 1 or 2 of the analysis. It is useful for this person to write a two-page, executive summary of the session, attaching the debriefing forms from both the on-the-scene observer and the videotape observer and including a brief description of the significant characteristics of the session, together with a tentative assessment of the extent to which the session appears to be worth further detailed analysis.5

Interpretation Cycle 4: Produce a Summary for One Problem Across Multiple Groups. Once executive summaries and typed transcripts have been prepared for all of the sessions that involve a given problem, it is useful for one researcher to be assigned to write a brief, two- to five-page report summarizing the results based on all of the sessions. This report should describe the similarities and the differences among all of the groups who worked on the given problem, describe and illustrate any generalizations that seem appropriate across all of the groups, and select at least one of the transcripts for indepth analysis from a specifically assigned, theoretical perspective. For example, if the researcher who is doing this analysis is a mathematician, then the transcript that is selected could be analyzed from the point of view of the kind of mathematical modeling perspective that is represented in the appendix. A mathematician might concentrate on successive cycles of modeling (e.g., focus on a single quantity to consider derived variables, like trends), whereas a social psychologist might study the events surrounding transitions in roles among the students in the group analyzing, for example, the implications of the sentence in the transcript in the appendix that reads: “Because Carla was calculating sums, she was no longer recording everybody else’s results, as she had done earlier. Therefore, the results of the calculations were not written down. . . . Consequently, at this point, Barb started writing down the results? . . .”

It is beneficial for the report to be presented at a research seminar in which the participants represent a broad range of theoretical and practical perspectives. The discussions of the report should be aimed at trying to reach a tentative consensus about how the session is described and explained. The discussions also should clarify alternative interpretations that could be given of the highlighted session, critique the summary of the entire collection of sessions, and identify issues that should be investigated during additional interpretation cycles. In particular, the discussions should assess the criteria that are emerging for selecting and eliminating sessions for further detailed analysis, identify a group of students whose performance should be investigated across multiple sessions and, it should choose a theoretical perspective that should be emphasized in the next analysis.

Interpretation Cycle 5: Produce a Summary for One Group of Students Across Multiple Problem Solving Sessions. Once the results are available from discussion groups focusing on a series of problems, for another researcher should write a brief, two- to five-page report summarizing the performance of a particular group of students across a series of problem solving sessions. Again, this report should describe the trends across multiple sessions and select at least one of the transcripts for indepth analysis from a new, specifically assigned,

---

5 At this early stage of transcript analysis, it is premature to eliminate a given transcript from further examination. Furthermore, premature screening is especially hazardous if the person doing the screening represents a narrow theoretical or practical perspective. For example, a session that may appear to be uninteresting to a teacher may be very interesting to a psychologist; a session that may appear to be uninteresting to a psychologist may appear to be very interesting to a mathematician. So, in general, it is sensible to follow the “triangle rule”; that is, do not screen out transcripts until at least three distinct perspectives have been represented in the analysis.
theoretical perspective. For example, if the transcript has been analyzed already from the perspective of mathematical modeling, the new analysis might concentrate on a psychological perspective and metacognitive roles and activities. The result should be a transcript with at least three columns: one dealing with what the students did and said, one interpreting the transcript from one theoretical perspective, and one interpreting the transcript from a second theoretical perspective. Again, it is helpful for the report to be presented at a research seminar in which the participants represent a broad range of theoretical and practical viewpoints and for the discussions to emphasize issues similar to those featured during Interpretation Cycle 4.

Interpretation Cycle 6: Produce Additional Analyses of the Sessions, Groups, and Problems That Emerge as Being the Most Interesting and Promising for Further Study. By the time Interpretation Cycle 5 occurs, a small number of sessions, groups, and problems are likely to emerge as being those most worthy of further investigation; also, complex dossiers will have been produced that include interpretations of these sessions from multiple practical and theoretical perspectives. Consequently, because alternative theoretical perspectives tend to result in mismatches and in new issues that need to be explained, one of the main goals of analysis cycles is to clarify which issues and perspectives should be highlighted in follow-up analyses. Furthermore, because highly specialized analyses may be required, it may be necessary to include researchers who were not involved in earlier analyses. Because detailed summaries are available for a small number of sessions, it often is possible for these new researchers to use advanced technologies and participate from remote sites.

SUMMARY

The use of videotape is not a royal road to unfettered evidence about conceptual change. Like any other methodological tool, researchers must establish an interpretive framework within which the video serves as one form of evidence. The validity of any such framework is likely to be enhanced if the researchers make decisions during the design of the study (before collecting any videotape!) about:

- The primary function of the videotape, especially when a researcher can anticipate focusing on a particular issue or on obtaining additional contributions from it to fill in or broaden the potential field of inquiry.
- Likely targets of the video, such as small groups, individuals, or teacher-led discussions.
- Methods used for triangulation, including the roles and functions of additional observers, and other sources of evidence, such as students’ artifacts, that could be used to buttress claims about students made in the light of the restricted lens of the camera.
- The methods used to organize and train a research team to ensure that videotape analysis is conducted in cycles of interpretation, with sufficient opportunities for different (but relevant) perspectives, cross-checking of particular episodes as well as collections of related sessions, and perhaps most important opportunities for changing and revising the interpretive framework so established painstakingly.
APPENDIX
The Summer Jobs Problem

Last summer Maya started a concession business at Wild Days Amusement Park. Her vendors carry popcorn and drinks around the park, selling wherever they can find customers. Maya needs your help deciding which workers to rehire next summer. Last year Maya had nine vendors. This summer, she can have only six – three full-time and three half-time. She wants to rehire the vendors who will make the most money for her. But she doesn’t know how to compare them because they worked different numbers of hours. Also, when they worked makes a big difference. After all, it is easier to sell more on a crowded Friday night than on a rainy afternoon.

Maya reviewed her records from last year. For each vendor, she totaled the number of hours worked and the money collected – when business in the park was busy (high attendance), steady, and slow (low attendance). (See the table.) Please evaluate how well the different vendors did last year for the business and decide which three she should rehire full-time and which three she should rehire half-time.

Write a letter to Maya giving your results. In your letter describe how you evaluated the vendors. Give details so Maya can check your work, and give a clear explanation so she can decide whether your method is a good one for her to use.

<table>
<thead>
<tr>
<th>TABLE A1</th>
<th>Hours Worked Last Summer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>JUNE</strong></td>
<td><strong>July</strong></td>
</tr>
<tr>
<td><strong>Busy</strong></td>
<td><strong>Steady</strong></td>
</tr>
<tr>
<td>MARIA</td>
<td>12.5</td>
</tr>
<tr>
<td>KIM</td>
<td>5.5</td>
</tr>
<tr>
<td>TERRY</td>
<td>12</td>
</tr>
<tr>
<td>JOSE</td>
<td>19.5</td>
</tr>
<tr>
<td>CHAD</td>
<td>19.5</td>
</tr>
<tr>
<td>CHERI</td>
<td>13</td>
</tr>
<tr>
<td>ROBIN</td>
<td>26.5</td>
</tr>
<tr>
<td>TONY</td>
<td>7.5</td>
</tr>
<tr>
<td>WILLY</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE A2</th>
<th>Money Collected Last Summer (In Dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>JUNE</strong></td>
<td><strong>July</strong></td>
</tr>
<tr>
<td><strong>Busy</strong></td>
<td><strong>Steady</strong></td>
</tr>
<tr>
<td>MARIA</td>
<td>690</td>
</tr>
<tr>
<td>KIM</td>
<td>474</td>
</tr>
<tr>
<td>TERRY</td>
<td>1047</td>
</tr>
<tr>
<td>JOSE</td>
<td>1263</td>
</tr>
<tr>
<td>CHAD</td>
<td>1264</td>
</tr>
<tr>
<td>CHERI</td>
<td>1115</td>
</tr>
<tr>
<td>ROBIN</td>
<td>2253</td>
</tr>
<tr>
<td>TONY</td>
<td>550</td>
</tr>
</tbody>
</table>
The students in Ms. Barnes’ seventh-grade class worked in three-person teams, with the members of the teams being assigned by Ms. Barnes. All the members of the team whose work is described herein were all considered to be average-ability mathematics students in an average-ability inner-city classroom. However, because of Ms. Barnes’ emphasis on portfolio-based assessment, the students had had considerable prior experience working on at least 10 projects similar in size to the summer jobs problem. The summer jobs problem was based on a context that was described in a “math-rich” newspaper article that was discussed by the class as a whole on the day before the summer jobs problem was presented.

The students worked at small tables where a tool kit was available that included three TI-92 calculators and other standard classroom tools. The work station included a Macintosh computer with a 12-inch color monitor and software for word processing, spreadsheets, drawing, and making geometric constructions.

After Ms. Barnes had passed out the summer jobs problem and after each group had read the problem carefully, the class as a whole answered the following questions:

- Who is the client? (Answer: Maya)
- What does Maya need you to produce? (Answer: A letter describing a procedure for deciding whom to rehire.)
- Why does she need this product from you? (Answer: So that she will know whom to rehire.)
- When does she need this letter? (Answer: At the end of tomorrow’s class.)

The solution process that follows includes significant segments from a transcript for a group of students whose names are Alan, Barb, and Carla. Most of the graphs shown were produced originally using TI-83 calculators. However, when the team presented their work in class, they used posters that contained redrawn versions of their favorite graphs. Generally, these graphs were constructed using a computer-based, graphing spreadsheet and a color printer. One of these posters is shown in Interpretation 14 at the end of this Appendix. The graph that is appended to the transcript was taken from one the posters. Only in those cases where it is so indicated were the graphs produced by Alan, Barb, and Carla.

Approximately 5 minutes passed as the students read the problem and discussed it.

Alan: Oh, God. We've gotta add up all this stuff? . . . You got a calculator?
Barb: They're in here [the toolbox]? . . . Here. [She finds two TI-83 calculators in the toolbox.]

Approximately 5 more minutes passed while Alan, Barb, and Carla added the numbers in various rows or columns of Table A1. Because the three students had made no attempt to coordinate their efforts, everybody went off in a slightly different direction. For example, Barb and Carla both added the numbers in the first row of Table A1 (which shows the number of hours that Maria worked), whereas Alan added the numbers in the first column (which shows the number of hours that all of the students worked during the busy periods in June).

Alan: I got, let's see, . . . 116.
Barb: You punched them in wrong. . . . Here, you read them [the numbers] and I'll punch 'em in.
Alan: [Pointing to the numbers in the table.] 12.5, 5.5, 12, 19.9, 19.5 . . .
Carla: Huh! . . . Not those, you dummy!
Alan: Why?
Barb: Here, . . . read these [pointing to the first row of Table A1].

Interpretation #1: A Hodgepodge of Several Unstated and Uncoordinated Ways of Thinking Is Used Inconsistently. This team's first interpretation of the summer jobs problem was similar to those generated by most of the other groups. That is, when the students first began to work on the problem:

- They tended to worry most about “What should I do?” rather
than “What does this information mean?” Therefore, their first interpretations focused on computation, and the information that was given was treated as if no data interpretation or mathematization were necessary. Also, when computation was done, it nearly always involved only two-item combinations; it did not involve computations of whole rows or whole columns of numbers.

They tended to focus on a small subset of the information only, and they tended to concentrate on isolated pieces of it rather than on searching for underlying patterns and regularities. For example, Alan, Barb, and Carla seized on the first information that they noticed, or on the information that impressed them most. In other words, they focused on only the rows or columns in the table that showed the number of hours that each vendor worked. This emphasis was not based on the thoughtful selection of which items were most important. They merely converged on the first details that came to their attention.

- Their early interpretations seldom consisted of a single coherent way of thinking about givens, goals, and possible solution procedures; instead, they usually involved a hodgepodge of several unarticulated and undifferentiated points of view. That is, different students think in different ways; even the same individual may switch sometimes (without noticing) from one way of looking at the problem to another way. For example, in the transcript that is given here, when Alan finished adding the first column of numbers in Table A1, he began to add the first column of numbers in Table A2; there was no evidence that he noticed that Table A1 dealt with hours worked and that Table A2 dealt with money collected. In fact, later in the session, Alan tried to subtract data in Table A1 from data in Table A2; that is, he tried to subtract hours from dollars (e.g., 690 dollars – 12.5 hours = ?).

- They tended to focus on numbers only, and ignored quantity types. For example, the quantity “12.5 hours” usually was read as “twelve point five,” emphasizing “how much” but ignoring “of what.”

Barb: Here, read these [again pointing to the first row of Table A1].
Alan: OK. . . . 12.5, 15, 9, 10, 14, 17.5 12.5, 33.5, and 35. . . . [pause]. . . . What is it? [pause]
Barb: It's 159.

As Alan was reading the numbers, Carla was checking them off in her table. When Barb gave the result, Carla recorded it in a new column on the right side of her table of data.

One noteworthy fact about this session with Alan, Barb, and Carla is that they began to work as a team much earlier than many of their peers. For example, the preceding differentiation of their roles (as readers, calculators, and recorders) is an indication that the students were beginning to work together as a team, rather than working as independent individuals.

Next, Alan, Barb, and Carla spent approximately 5 minutes calculating the total amount of time that the other vendors worked. Carla recorded the results in the last column of her table. The table of sums that they produced corresponds to the graph shown in FIG. A1.
Interpretation #2: Focusing on Total Number of Hours For Each Worker. The graph in FIG. A1 and Table A3 focused on only the total number of hours that each worker worked. In presentations of their results, the notions of “seniority” or “willingness to work” were common justifications that students used for emphasizing “hours worked.” Unlike many other groups that produced the graph in FIG. A1 as part of their final presentations, Alan, Barb, and Carla did not bother to produce the graph. They only produced the table of sums (see Table A3) that would have led to this graph. This seemed to be true for several reasons. First, the table of sums that Alan, Barb, and Carla produced was, in itself, enough to enable them to go on to a new and improved way of thinking about the information that was given. Second, at this point in the session, Alan, Barb, and Carla were only using their calculators to operate on pairs of numbers; they were not operating on whole lists of numbers. Therefore, they were not entering data into their calculators (or their computer) in a form that made it easy for them to produce automatic graphs.

Alan: OK, so who should we hire? [Alan was looking at Carla's table of sums.]
Barb: Robin looks good. . . . [pause] So does Tony.
Alan: Maybe Kim.

Carla: Hey! We ought to look at money, not hours. . . . Money is down here [pointing to Table A2, which shows the amount of money each student earned].

Alan: Yep, money.

[Approximately 1 minute passed while the students thought and looked at the table.]
Barb: OK, let's add these [pointing to the rows in the second half of the table].
Barb: Here, you do Maria [gesturing to Alan]. You do Kim [gesturing to Carla]. And, I'll do Terry.

Again, it is noteworthy that Alan, Barb, and Carla worked as a team much better than many of their peers. For example, here, they divided the task into several different pieces, with different students working on different pieces. Therefore, more planning, monitoring, cross-checking, and rethinking were likely to occur. Alan seemed to be insecure about using a calculator, so he added only one row of numbers; then, he began watching the other two students as they worked. After that, he began to act as a facilitator and a monitor for the group, rather than as a person was doing calculations.
Alan: Here, I'll read the numbers for you [looking at Carla].

Next, approximately 3 minutes passed as the students calculated sums in the second half of the table.

Alan: So, let's see what we've got... Who made the most?

Because Carla was calculating sums, she was no longer recording everybody else's results, as she had done earlier. Therefore, the results of the calculations were not written down in an orderly fashion; they were written on scraps of paper. Consequently, at this point, Barb started writing down the results of her own calculations in a new column at the right of her table.

Barb: Where's the answer for Tony? Where did I put it? Darn, I'm going to have to do it again... Oh, here is it... No... Oh, I'm not sure... I'd better do it again.

Carla: My numbers are: “fourteen nine twenty-one” for Kim, “fifteen two seventy-one” for Robin, “nine three oh eight” for Willy, and “thirteen nine sixty-four” for Tony.

Carla: Hey, I did Tony. It's “thirteen nine sixty-four.” You don't need to do it.

Barb has become the temporary recorder for the group. Here, she takes several minutes to collect the results from the group and to record them in a column (like the one that Carla had constructed in the group's earlier work).

Interpretation #3: Focusing on the Total Number of Dollars That Each Worker Earned: Some teams essentially quit working on the problem at this point. For these groups, their presentations often included a graph like the one shown in FIG. A2. Again, probably for the same kinds of reasons as for Interpretation #2, Alan, Barb, and Carla used only a *table* of sums; they did not bother to construct the graph shown in FIG. A2.

![FIG. A2. Total dollars each vendor earned.](image)

<table>
<thead>
<tr>
<th>TABLE A4</th>
<th>Total Dollars Each Worker Earned</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARIA</td>
<td>$8,196</td>
</tr>
<tr>
<td>KIM</td>
<td>$14,921</td>
</tr>
<tr>
<td>TERR</td>
<td>$7,000</td>
</tr>
<tr>
<td>JOSE</td>
<td>$11,373</td>
</tr>
<tr>
<td>CHAD</td>
<td>$9,284</td>
</tr>
<tr>
<td>CHERI</td>
<td>$11,062</td>
</tr>
<tr>
<td>ROBIN</td>
<td>$15,271</td>
</tr>
<tr>
<td>TONY</td>
<td>$13,964</td>
</tr>
<tr>
<td>WILLY</td>
<td>$9,308</td>
</tr>
</tbody>
</table>

Alan: So, who's the best?... Robin's best. She got “fifteen two seventy-one.”

And, Kim got “fourteen nine twenty-one.” Who's next?

Carla: Tony... He got “thirteen nine sixty-four.”

Barb: This isn't fair. Some guys got to work a lot more than others... Look at Robin and Tony. They worked more than everybody else. That's why they made more money... If Maria worked that much, she'd have made that much money too. [Mumbling.]
At this point in the session, nobody picks up on Barb's suggestion to investigate the relationship between “dollars earned” and “hours worked.” Nonetheless, later in the session, Barb comes back to this same suggestion; and, at that time, her suggestion leads to the idea of investigating “dollars-per-hour” for each worker. Now, however, the idea of “dollars-per-hour” is not pursued. Instead, Barb's suggestion led the students to investigate changes in the dollars earned across time.

Barb: Look, Willy didn't work at all in June [pointing to the zeros by Willy's name in Table A2.] But, he was doing great in August [pointing to the $3,005 by Willy's name in the August column of Table A2]. . . . Let's just see how much everybody got, totally, in August.

<table>
<thead>
<tr>
<th>Table A5: Dollars Earned Each Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>MARIA</td>
</tr>
<tr>
<td>KIM</td>
</tr>
<tr>
<td>TERRY</td>
</tr>
<tr>
<td>JOSE</td>
</tr>
<tr>
<td>CHAD</td>
</tr>
<tr>
<td>CHERI</td>
</tr>
<tr>
<td>ROBIN</td>
</tr>
<tr>
<td>TONY</td>
</tr>
<tr>
<td>WILLY</td>
</tr>
</tbody>
</table>

It is significant that Table A5 was put together in a top-down fashion. Earlier tables were simple lists, and even these lists were created by doing the individual calculations first, then putting the results into a well-organized form. The organizational system was not generated first and used to guide the computations that were performed. That is, each of the earlier lists was constructed in a bottom-up fashion.

Alan: Look at old Willy. He's really catching on [at the end of the Summer]. . . . Look, back here [in June] he only made a hundred and eighty-nine bucks, but, out here [in August] he was really humming.

Barb: I think August should count most. Then July. . . . I don't think June should count much. They were just learning.

Alan: How we going to do that?

Barb: I don't know. Just look at them [the numbers in the table], I guess.

[pause]

Barb: Let's see, out here [in August], Kim was best. . . . Then Robin, no Tony....Then Robin....I think they're the top three. Kim, Robin, and Tony. . . . How'd they do in July?

Barb: Wow! Look at Kim. She's still the best. . . . But, uh oh, look at Cheri. She was real good in July.

Alan: Let's line 'em up in July. Who's first?

Barb: Kim. . . . [pause] Then Tony, and Cheri, and Robin. . . . [long pause] . . . Then Willy, Chad, and Jose. . . . [long pause]. . . . And, these guys weren't very good [referring to Maria, and Terry].
Though Barb was doing most of the talking and overt work, Alan was watching and listening closely. But, Carla was off on her own, playing with the computer's spreadsheet program and entering lists of numbers. . . . At this point, Carla reentered the conversation.

Carla: Look you guys, I can make a graph of this stuff.

**Interpretation #4b: Using a Graph to Focus on the Total Number of Dollars Earned Each Month:**

For the next 4 minutes, Carla used the computer to flip back and forth, showing the graphs in FIG. A3 that she had made, explaining how she made the graphs, and pointing out who was the top money earner each month.

**FIG. A3.** Total dollars earned each month.

Barb: OK, let's, like, line 'em up for each month.
Alan: You started doing that.
Barb: OK, you [Alan] read 'em off and I'll write 'em down.
Alan: OK, here's August. . . . We got Kim, Tony, Robin, . . . Cheri, . . . Willy, Chad, and Jose; then, Maria and Terry.

**Interpretation #5: Focusing on Trends in Rank Across Time:**

For approximately 5 minutes, Alan, Barb, and Carla worked together to get a list of “top money makers” each month. Then, they noticed that the rankings were somewhat different each month; so, the “trends” shown here were used as an early attempt to reduce this information to a single list.

Alan: Look, Kim was top in July and August, and so was Tony. . . . Robin was next in August, but she wasn't as good in July. . . . [pause] . . . But, she [Robin] was really good in June. . . . [pause] . . . I think June is most important because some of them were just learning. . . . August is how they'll probably do next summer.
Approximately 5 more minutes passed while each of the three students nominated workers that they believed should be hired, based on the rankings and trends in the Table A6. In most cases, when the students spoke in favor of a given worker, they made up some sort of cover story to account for the “ups” and “downs” in the worker’s performance. These cover stories involved the following kinds of possibilities: Some workers learned and improved, whereas others got bored; some were not able to work as much as others; and some were good during busy periods, but not during slow periods. In these discussions, the students started to pay attention to the fact that the months might not be equally important (e.g., July is the busiest month; August might be the best indicator of current abilities) and that busy, steady, and slow periods might not be equally important (e.g., the half-time workers would not be hired during slow periods). In addition, the students began to express concerns about the fact that they would have liked to have had some additional information that was not available such as: Who really needed a job badly? Who was willing to work when they were called? Finally, as Carla was looking at Table A6 (see Interpretation #5), she got the idea to make a similar graph using the computer; this idea led to Interpretation #6.

Carla: I can make a graph like that [pointing to Table A6] with the computer. Wanna see?
Alan: Sure, um, what’s it look like?
Barb: . . . Let me see.

**Interpretation #6: Focusing on Trends in Money Earned For June, July, and August:** Carla's graph was a line graph showing the total number of dollars that each worker earned for June, July, and August (see FIG. A4).

**FIG. A4. Trends in money earned for June, July, and August.**

Alan: Wow! Neat! How'd you do that?
Carla explained again how she had made the graphs using the computer.

Alan: Now who do we pick . . . Who's this?
Carla: Um, let's see, it's Kim. . . . And this is . . . um . . . Tony.
Alan: Who's this?
Barb: Let me see.
Carla: Oh, it's Robin.
Barb: So, we've got Kim, . . . Tony, and Robin. Who's next?

[pause]

Carla: What about this guy? . . . Who is he? . . . Um, it's Cheri. . . . Look, she was really good here, but, then, she screwed up.
Barb: How we gonna decide which of these guys to hire? They were all good some and bad some. . . . [long pause] . . . How many were we supposed to hire anyway? . . . [pause]. . . . Look at the problem [speaking to Alan]. What does it say?

[long pause]

Alan: We're supposed to hire three full-time and three part-time.

[long pause]

Alan: I think we should hire Willy. He was good here [pointing to July and August] . . . and he didn't get to work much here [pointing to June].

Interpretation #7: Using Telescoping Decision Rules: Up until this point in the session, the students implicitly seemed to assume that the best way to choose workers should be to use a single rule for ranking the workers. Then, if this list was successful in ranking workers from “best” to “worst,” the top three workers could be hired for full time, and the next three workers could be hired for part time. But, unfortunately, life was not this simple. No single rule seemed to work to form a single list. For example, both Barb and Carla suggested the idea of using some sort of average. But, this idea was not considered in detail, because the type of averages that were mentioned didn't seem to involve equally important quantities. Therefore, the students began to consider more sophisticated decision-making rules. For example, one rule involved the following kind of two-step process: First-round decisions about who to hire could be based on the ranking in August alone; then, second-round decisions could be based on the ranking in July alone (or based on busy periods alone).

Barb: Look, you guys. Some of these people got to work a lot more than others. . . . That's not fair. Look, Willy didn't get to work at all back here [in June].

Carla: So, what're we gonna do? [Mumbling.]

More than 1 minute passed.

Alan: Here. I'm trying something. . . . I'm subtracting how much each guy worked. That'll kind of even things out. . . . I worked for a guy who did that once. We were cleaning up trash and he wanted us to work fast.

Interpretation #8: Subtracting Time Scores From Money Scores: The most important characteristic of this new idea is that, for the first time, it took into account a relationship between the amount of money that was earned and the amount of time that was spent working. But, because the numbers in the tables didn't include any unit labels, nobody noticed that it might not make sense to subtract hours from dollars. Nonetheless, neither Barb nor Carla were convinced that the idea made sense. What did make sense to Barb and Carla was to apply lessons they had learned from their own prior real-life experiences to help them make decisions in the case of the summer jobs problem. Therefore, the team didn't pursue Alan's suggestion. Instead, Alan's suggestion was used as a (transitional) way of thinking that led to a better idea that Barb suggested, which would take into account both time and money.

Barb: Hey, that's a good idea! We could figure out dollars per hour. . . . I did that for my jobs last summer.

Interpretation #9: Focusing on Dollars-Per-Hour: Barb wasn't really paying close attention to Alan's idea. The new ideas that she heard was to think about the situation in the same way that she thought about her own past jobs. That is, both Alan and Barb were using past real-life experience to make sense of the current problem. Therefore, Barb thought in terms of dollars-per-hour.
For the remaining minutes of the class, Alan, Barb, and Carla went back to the original data tables of hours worked and dollars collected by each vendor last summer and started calculating dollars per hour. As class ended, they decided that, to prepare for the next day’s class, each student should bring a graph showing the dollars per hour collected by each vendor. Then, they planned to use these graphs to make final decisions about who to rehire. The graphs in Figs. A5 and A6 show what each student brought to class the next day.

**Interpretation #10a: Alan’s Dollars-Per-Hour Graph Based on Sums For the Whole Summer:**
First, Alan calculated the total amount of money that each worker earned for the whole summer. Then, he calculated how much time they worked altogether. Finally, for each worker, he divided total dollars by total time.

**Interpretation #10b: Barb’s Dollars-Per-Hour Graph Based on Sums For Each Month:**
First, Barb calculated the total amount of money that each worker earned for each month. Then, she calculated how much time they worked each month. Finally, for each month, she divided dollars earned by time worked.

**Interpretation #10c: Carla got some help from her brother, who apparently suggested the idea of an average. First, Carla calculated the dollar-per-hour for each cell in the matrix shown in Table A7. Then, for each month, she calculated the average of the rates for the busy, steady, and slow periods. This procedure assumes (incorrectly) that the students intended to treat busy, steady, and slow periods as being equally important!**

**TABLE A7**
Average Dollars-per-Hour Rate Each Month (Where the Average Is Taken Across Busy, Steady, and Slow Periods)

<table>
<thead>
<tr>
<th></th>
<th>June</th>
<th>July</th>
<th>August</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Busy</td>
<td>Steady</td>
<td>Slow</td>
</tr>
<tr>
<td>Maria</td>
<td>$55.20</td>
<td>$52.00</td>
<td>$50.22</td>
</tr>
<tr>
<td>Kim</td>
<td>$86.18</td>
<td>$39.73</td>
<td>$26.19</td>
</tr>
<tr>
<td>Terry</td>
<td>$87.25</td>
<td>$39.24</td>
<td>$19.59</td>
</tr>
<tr>
<td>Jose</td>
<td>$64.77</td>
<td>$38.95</td>
<td>$22.50</td>
</tr>
<tr>
<td>Chad</td>
<td>$64.82</td>
<td>$45.08</td>
<td>$68.81</td>
</tr>
<tr>
<td>Cheri</td>
<td>$85.77</td>
<td>$61.78</td>
<td>$47.83</td>
</tr>
</tbody>
</table>
For approximately the first 20 minutes of the second class, Alan, Barb, and Carla showed one another their rate-per-hour graphs and explained how they were made. Then, for each graph, the team worked together to try to decide which workers should fall into these categories: full-time, part-time, and don’t hire.

For Alan's list, the ranking was easy to read directly from the graph that he had drawn (see FIG. A5), but, for Barb's and Carla's graphs (see Figs. A6 and A7, respectively), it was not so obvious to determine which workers ranked first, second, third, and so on. Therefore, for both of these graphs, the team used telescoping methods of decision-making. That is, first-round (tentative) decisions were based on performances in August alone; then, to make decisions about difficult cases, information was used from July (or June). The results are shown in Table A8.

**Interpretation #11:** Three different lists were generated that ranked workers from lowest to highest based on the three, dollar-per-hour graphs that the students had produced.

<table>
<thead>
<tr>
<th>Alan's List</th>
<th>Barb's List</th>
<th>Carla's List</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheri</td>
<td>Kim</td>
<td>Cheri</td>
</tr>
<tr>
<td>FULL-TIME</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim</td>
<td>Cheri</td>
<td>Jose</td>
</tr>
<tr>
<td>Willy</td>
<td>Willy</td>
<td>Kim</td>
</tr>
<tr>
<td>Maria</td>
<td>Jose</td>
<td>Tony</td>
</tr>
<tr>
<td>PART-TIME</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robin</td>
<td>Chad</td>
<td>Maria</td>
</tr>
<tr>
<td>Chad</td>
<td>Robin</td>
<td>Willy</td>
</tr>
<tr>
<td>Jose</td>
<td>Maria</td>
<td>Chad</td>
</tr>
<tr>
<td>DO NOT HIRE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tony</td>
<td>Tony</td>
<td>Robin</td>
</tr>
<tr>
<td>Terry</td>
<td>Terry</td>
<td>Terry</td>
</tr>
</tbody>
</table>

Because the three lists were somewhat different, Alan, Barb, and Carla tried to make a new list (which they called their “agreement list”), showing points of agreement among the three lists.
Alan: Look, Cheri's in the top for all three lists. . . . Kim, too. We should hire them for sure. Cheri and Kim. . . . What about Willy?...Nope, he's way down on Carla's list. Damn! Now, what're we gonna go?
Barb: We can throw out Terry. He's at the bottom of every list.
Alan: Tony's pretty bad too.
Carla: I don't think he's so bad. . . . Here look at my graph.
Barb: OK, you're right. He did OK.
Carla: OK.

. . . [pause] . . .

Alan: Who else should we hire full-time? Willy maybe. Or Jose. . . . No, Jose's way down here on my list.

[pause]

Carla: Look, on my list, Cheri, Jose, and Kim all got A's. . . . Tony, Robin, and Willy got B's. And Chad, Robin, and Terry got C's. . . . What did they get on your lists?
Alan: What do you mean?
Carla: Give me your list, I'll show you. . . . [pause] . . . See. Cheri got an A, and so did Kim and Willy.
Barb: What are you guys doing?
Carla: Here, watch.

For approximately the next 5 minutes, Carla asked the other two students to give her information to fill in the grading scale shown in Table A9.

**Interpretation #12:** For each list, a “grading scheme” is imposed that is similar to those used for tests in class. Then, the scores are combined (treating each of the rankings as if they were independent ratings).

<table>
<thead>
<tr>
<th></th>
<th>Alan's List</th>
<th>Barb's List</th>
<th>Carla's List</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheri</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Kim</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Willy</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A-</td>
</tr>
<tr>
<td>Jose</td>
<td>C</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Robin</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>B-</td>
</tr>
<tr>
<td>Chad</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>B-</td>
</tr>
<tr>
<td>Maria</td>
<td>B</td>
<td>C</td>
<td>B</td>
<td>B-</td>
</tr>
<tr>
<td>Tony</td>
<td>C</td>
<td>C</td>
<td>B</td>
<td>C+</td>
</tr>
<tr>
<td>Terry</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

Alan: So, it looks like the full-time people should be Cheri, and Kim, and Willy. . . . And part-time should be Jose, and...uh oh! Who should we pick next? Maria, Robin, or Chad.
Barb: Yeah. Tony and Terry are out.
Alan: These other guys are pretty close. . . . It's not fair to just pick one.
Carla: Maybe one of these guys really needs a job. I’d think we should hire guys who really need a job. . . . Maybe Willy doesn't really need a job. Maybe Jose really needs one.
Alan: Some of these guys probably didn't get to work at the good times.

[pause]
Barb: Let's make more graphs like these [pointing to FIG. A7 in Interpretation #10c] for the slow times, and the steady times, and the fast times.
More than 12 minutes passed while Alan, Barb, and Carla worked together to make the following graphs.

**Interpretation #13:** A telescoping series of rules. First round decisions are based on Interpretation #12. Then, second-round decisions are made by comparing dollars-per-hour for busy, steady, and slow periods.

FIG. A8. Dollars-per-hour for busy periods.

FIG. A9. Dollars-per-hour for slow periods.

FIG. A10. Dollars-per-hour for steady periods.

Barb: [Looking at Figs. A8, A9, and A10.] I don't think this helps much.
Carla: [Looking at Figs. A8, A9, and A10.] So, which one should we hire? Maria, Robin, or Chad?
Alan: Look, Maria's only best during slow times, but we don't really care about slow times. We're only going to hire part-time people when things are happening . . . fast times . . . .[pause]
Barb: Wait a minute. Maria's not so bad. Look, um, she's better than Robin during steady times . . . and Chad too.

Approximately 8 minutes passed during which Alan, Barb, and Carla looked back over the graphs that they had brought to class and the work that they had done earlier in the period. In these discussions, they offered “stories” that might explain patterns in the dollars-per-hour collected by the various workers. In the end, they reached an agreement on the following points:

- Slow periods should not be treated as being very important because most of the money would be made during busy or steady periods and part-time workers would not be hired during slow periods.
- Performance in August (and, to a lesser extent, July) should be treated as being most important because it took into account learning and improvement and because it was the most recent indicator of capabilities.

Carla: We've got to write up our report. . . . What should we do?
Barb: I think we should make another graph like the one I made before [i.e., Interpretation #10b, FIG. A6], . . . only this time leave out slow times.
Carla: OK, you do that. . . . I'll get the poster board and stuff.

For the remainder of the class, Alan, Barb, and Carla worked together to produce a large poster like the one shown in FIG. A11.
Dear Maya,

We think you should hire Kim and Cheri and Jose for full time, and we think you should hire Willy and Chad, and Tony for part time. Look at this graph to see why these people are best.

The graph is only about busy times and steady times. You don't make much money during slow times, and you won't hire people for slow times.

Some workers got better at the end of the summer. But, some didn't get better. So, August is most important, and July is also important. July is when you make the most money.

Alan, Barb, and Carla

At a similar point, another group of students produced the graph shown FIG. A12.
At no point in the session did Alan, Barb, or Carla use the spreadsheet as a calculating device. They only used the spreadsheet as a graphing tool. They did all calculations on TI-83 calculators.