Making Mathematics on Paper:
Constructing Representations of Stories About Related Linear Functions

Rogers Pierpoint Hall

Department of Information & Computer Science
University of California, Irvine, CA 92717

Technical Report 90—17

June 26, 1990

This doctoral dissertation was supported by grants from the Cognitive Science Program of the Office of Naval Research (contract N00014-85-K-0373), a University of California Regents Fellowship, a Joseph J. Fischer Fellowship, a Pease Barker Dissertation Fellowship from the University of California, and a student internship at the Institute for Research on Learning in Palo Alto, California.

List of Figures

List of Tables

Acknowledgements

Completing this dissertation was a larger struggle than I originally anticipated. Without the daily support of my partner, Kären Wieckert, I would not likely have survived the process. While the encouragement and guidance of my committee members provided focus and structure to the project, Kären did much of the "invisible work" that allowed me to fill in that structure. Well, I can see that work now, and I hope to reciprocate.

I first want to thank my cohort in the Irvine Computational Intelligence Project for an environment in which to learn about cognitive science and to try out an academic life. The list of friends, co-conspirators, and sometimes-adversaries should be a proper history of the Project, but the following folks were ever present and important to me: David Aha, Marc Albert, Jack Beusmans, Greg Bolcer, Dan Easterlin, Kurt Eiselt, Doug Fisher, David Foster, Steve Hampson, Jennifer Holbrook, Jim Kipps, Linda Palmer, Bruce Porter, Don Rose, David Ruby, Jeffrey Schlimmer, Karl Schwamb, Dave Schulenburg, Chris Truxaw, Etienne Wenger, and Rika Yoshii. As the more recent members of this list know, the environment would be much poorer if not for the work of Caroline Ehrlich. Thanks to you all, and I hope I gave somewhere near as good as I got.

While finishing up, I had the good fortune of working with members of the Grounded Theory seminar, which produced hours of illuminating discussion, careful review of materials I brought to the group, and repeated personal affirmations that kept my work life going. My thanks and the very best to each of you: Werner Beuschel, Bill Brigham, Bruce Ferguson, Jeanne Pickering, Karen Ruhleder, Ed Shi, Leigh Star, Clark Turner, and Kären Wieckert. I hope that we all go on to make qualitative studies of representational practice a reality.

Two lines of influence on this dissertation arose outside my department, converged for a time at the Institute for Research on Learning, and have continued to lead me into very exciting intellectual terrain. The first appeared unexpectedly in a very provocative talk given by Jean Lave about her work with the Adult Math Project (AMP). To my delight, Jean agreed to sit on my examining committee for advancement to candidacy, where she convinced me to pursue the verbal interview studies, looking closely at participants' activities for their personal construction of "problem solving." Chapter 4 is my attempt to follow that suggestion. Also an AMP member, Olivia de la Roche read early versions of Chapter 3 and contributed long discussions over beer, chips, and salsa about the material circumstances of cognition in general and mathematical cognition specifically.

The second external influence began with a review by Jim Greeno of the study reported in Chapter 3. Dealing with his original comments required that I begin rethinking the relation between "knowledge" and "competence," and Chapter 5 grew directly out of this process. Fortunately, Jim agreed to stay engaged in the project, supporting me as a summer intern at the Institute for Research on Learning and then sitting on the dissertation committee. What started as a paper review has turned into an exciting intellectual exploration that I hope will move well beyond the dissertation.

Three people have anchored my learning about cognitive and computer science during graduate studies. First, I want to thank Etienne Wenger for early, formative collaboration on problem structure and later, challenging discussions of problem-solving activities. Etienne more generally has been an intellectual and human companion over the years. Second, I have learned a great deal from Leigh Star in all too short a time together here at Irvine. "Work" is a very compact point of entry into the world of human activity, and this camel thanks you for pulling him through it. Third, I want to thank Dennis Kibler for my introduction to this business and for hours of illuminating discussion about artificial intelligence, cognitive psychology, research methodology, mathematics, and living. It has been a long, sometimes strange, but usually exciting journey. Thanks for your commitment to scholarship and generous patience, particularly when I had lost it for myself.

This dissertation would not have been possible without various forms of institutional support: grants from the Cognitive Science Program of the Office of Naval Research (contract N00014-85-K-0373), a University of California Regents Fellowship, a Joseph J. Fischer Fellowship, a Pease Barker Dissertation Fellowship from the University of California, and a student internship at the Institute for Research on Learning in Palo Alto, California. Finally, my thanks for help and advice from other people in the department: Mary Day, Pat Harris, Dan Hirschberg, Susan Hyatt, Phyllis Siegal, and all the folks in computing support.

Abstract of the Dissertation

Making Mathematics on Paper
Constructing Representations of Stories About Related Linear Functions by
Rogers Pierpoint Hall
Doctor of Philosophy in Information and Computer Science
University of California, Irvine, 1990
Professor Dennis Kibler, Chair

This dissertation takes up the problem of applied quantitative inference as a central question for cognitive science, asking what must happen during problem solving for people to obtain a meaningful and effective representation of the problem. The core of the dissertation reports exploratory empirical studies that seek to answer the descriptive question of how quantitative inferences are generated, pursued, and evaluated by problem solvers with different mathematical backgrounds. These are framed against a controversy, described in Chapter 2, over the theoretical and empirical validity of current cognitive science accounts of problems, solutions, knowledge, and competent human activity outside of laboratory or school settings.

Chapter 3 describes a written protocol study of algebra story problem solving among advanced undergraduates in computer science. A relatively open-ended interpretive framework for "problem-solving episodes" is developed and applied to their written solution attempts. The resulting description of problem-solving activities gives a surprising image of competence among an important occupational target for standard mathematics instruction.

Chapter 4 follows these results into detailed verbal problem-solving interviews with algebra students and teachers. These provide a comparison across settings and levels of competence for the same set of problems. The results corroborate similar generative activities *outside* the standard formalism of algebra across levels of competence. Notable among these nonalgebraic problem-solving activities are "model-based reasoning tactics," in which people reason about quantitative relations in terms of the dimensional structure of functional relations described in the problem. These tactics support different activities within surrounding solution attempts and usually describe "states" in the problem's situational structure.

Chapter 5 holds these activities accountable to local combinations of notation and quantity, reinterpreting results for model-based reasoning in an ecological analysis of material designs for constructing and evaluating quantitative inferences. This analysis

brings forward important relations between what material designs afford problem solvers and the complexity of episodic structure observed in their solution attempts. The dissertation closes with a reappraisal of the relationship between knowledge, person, and setting and, I will argue, puts us on a more promising track for a descriptively adequate theoretical account of constructing mathematical representations that support applied quantitative inference.

Chapter 1
Making and Using Mathematical Representations

1.1 The problem

How do people in school, at work, and in their daily lives reason about quantitative relationships? There is no shortage of answers under particular views of how people *should* reason about mathematics, but these idealized accounts of quantitative inference generally fall short of telling us much about what actually happens as quantitative problems arise and are resolved among individuals. In keeping with prescriptive views, we might better ask what does not happen when people solve these kinds of problems, particularly in light of evidence for stationary or even falling levels of achievement among American students in the classroom and after entry into the working world (e.g., National Commission on Educational Excellence, 1983). Quantitative inference has become part of a high profile social agenda, but we know relatively little about how it occurs outside of formal assessments.

1.2 Ill–structured problems

Confronted with an algebra story problem, a student faces a fundamental sort of "ill–structured problem" (Newell, 1969; Simon, 1973; Star, 1989b). The problem text gives information about initial and goal states, but state–transition operators taking the text into a quantitative solution are hardly well–defined. Even assuming the student has an adequate grasp of mathematical principles and operators within the formalisms of arithmetic and algebra (e.g., the distributive property of multiplication over addition or using algebraic substitution), a solution to the presented problem is often obvious only in retrospect. Rather than searching for a solution path in a well–defined space of representational states, the problem solver is more likely to be attempting to construct a representation in which the problem becomes routine or familiar. Omitted or incorrectly introduced constraints within the problem representation can lead to prolonged and often meaningless calculations, and may encourage otherwise sophisticated problem solvers to give up entirely.

In many respects, the idea of solving ill–structured problems is oxymoronic for traditional studies of artificial intelligence, where problems need to be reasonably well–structured before even "weak methods" can be applied to search for a solution. Granted that problem solving can be improved by a change in an existing problem space (e.g., chunking parts of the space with macro–operators), what makes ill–structured problems difficult is that a representation sufficient to support search in a problem space (even if horribly uninformed search) is required before the computational sense of problem solving can even begin. It is little wonder, then, that information–processing models of ill–structured problem solving remain elusive.

This state of affairs might be puzzling but acceptable if algebra story problems were transient disturbances in the secondary school curriculum. However, these problems recur as a general task throughout the mathematics curriculum and are even found in the quantitative sections of entrance examinations for professional schools. If prevalence alone is an insufficient basis for study, the prescribed role of these problems in bringing mathematical formalisms into contact with "everyday experience" recommends them highly. Viewed from within the classroom, story problems are thought to provide students with an opportunity to apply acquired mathematical abstractions in more familiar domains (e.g., traveling or shopping). Viewed in a wider context, these problems also provide a curricular microcosm of a central pedagogical problem: transfer of training from the algebra classroom to students' later educational or life experiences.

Interpretations derived from either vantage are controversial. For example, I have anecdotal evidence that these problems are avoided by some teachers as being too difficult for both students and teachers. On the other hand, studies of mathematics in practice suggest that "real–world" curricular materials may have little correspondence with mathematical problems or their solution in "real life" (Lave, 1986, 1988a). For cognitive and educational theorists alike, the problem is to determine how applied problems are solved by competent problem solvers and how acquisition of that competence might be supported.

Algebra story problems of the sort shown in Table 1 have been studied extensively by cognitive and educational psychologists, both as a representative task for mathematical problem solving (e.g., Hinsley, Hayes, and Simon, 1977; Mayer, Larkin, and Kadane, 1984; and Paige and Simon, 1966) and as experimental materials for studies of transfer (e.g., Dellarosa, 1985; Reed, 1987; Reed, Dempster, and Ettinger, 1985; and Silver, 1979, 1981). Many studies treat problem solving as an opaque process with an inspectable output (i.e., correct or incorrect) and duration. Manipulations in problem content or presentation are introduced, performance data are collected, and inferences are drawn concerning hypothetical problem–solving mechanisms. In contrast, much as in Kilpatrick's early work (1967) and subsequent studies of mathematical problem solving by Lucas (1980) and Schoenfeld (1985), I have chosen instead to present people with representative problems and then to observe and analyze their uninterrupted responses in some detail. This approach trades experimental control over the problem–solving setting for a richer interpretive view of problem–solving activities.

Table 1.1: Representative algebra story problems.

Motion: Opposite direction (MOD).

> Two trains leave the same station at the same time. They travel in opposite directions. One train travels 60 km/h and the other 100 km/h. In how many hours will they be 880 km apart?

Motion: Round trip (MRT).

> George rode out of town on the bus at an average speed of 24 miles per hour and walked back at an average speed of 3 miles per hour. How far did he go if he was gone for six hours?

Work: Together absolute (WT).

> Mary can do a job in 5 hours and Jane can do the job in 4 hours. If they work together, how long will it take to do the job?

Work: Competitive (WC).

> Randy can fill a box with stamped envelopes in 5 minutes. His boss, Jo, can check a box of stamped envelopes in 2 minutes. Randy works filling boxes. When he is done, Jo starts checking his work. How many boxes were filled and checked if the entire project took 56 minutes?

In addition to finding whether or not a subject has gotten a problem "right," this approach allows exploration of the solution strategies that subjects select and their tactical course in achieving solutions, right or wrong. This is useful for characterizing what competent problem solvers actually do when solving these problems (i.e., a succession of strategic and tactical efforts) and is a necessary first step towards finding methods for supporting acquisition of competent problem–solving behaviors.

1.3 Constructing well–structured problems

Given an algebra story problem text, a problem solver must somehow convert the ill–structured task of finding a precise value from a written description of a situation into a task that is familiar enough to support inferences about quantitative relations between given and unknown quantities. The conventional instructional approach, one also adopted by early computational studies of language comprehension (reviewed in Chapter 2), is somehow to "translate from words to equations" and then to manipulate the resulting equations to find a precise value for the requested quantitative unknown. While a cursory examination of algebra textbooks shows that this method is still taught in beginning algebra classes, most would agree that this kind of activity is a rather narrow form of competent quantitative inference.

1.3.1 Competent algebra story problem solving

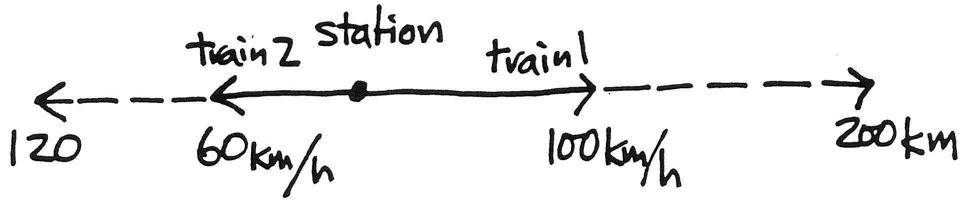
Figure 1.1 shows a written protocol collected with an advanced undergraduate in computer science, taken from a larger study described in Chapter 3. It is important to note that this student is a *competent mathematical problem solver*, in the sense that she has been trained to institutional standards in the University of California system. Given this background, what does the written protocol tell us about constructing an algebraic representation of related linear functions?

First, there are many notations and quantities used in this solution attempt, including a diagram, labels on diagrammatic components, textual annotations recording conjectures or observations, a table organizing related quantities, traditional algebraic expressions showing an algebraic proportion, and an arithmetic expression that clearly violates any conventional interpretation of equality (i.e., "30 min = 80 miles"). Second, we can divide the written protocol into what appears to be a series of coherent problem–solving episodes, starting with a diagram showing part of the specified train separation. Successive extensions to the diagram appear to be used to calculate cumulative distance, leading to a conjecture about relative motion — i.e., "The trains are moving apart at 160 km/h?" The diagram appears to be used as a model that shows the relative position of trains after two hours of travel, based on implicit inferences that distances elapsed can be combined and that the trains travel at the same time.

Only part way into this solution attempt, we find that the student has already constructed a quite elaborate material representation of the events depicted in the problem, that labels in this drawing carry relevant quantities, and that this construction has consequences for generating further quantitative inferences (i.e., the conjecture about combined rates). Next, the student outlines a table and iteratively constructs individual "states" consistent with the model from the first episode. Each successive state organizes a set of simple arithmetic calculations (addition and multiplication), and the recorded values give implicit evidence for a linear relationship between time and individual distances travelled or the distance apart. These state calculations eventually exceed the global constraint that trains end up 880 km apart, but confirm the earlier conjecture about relative motion (i.e., "Yes. It will take between 5 & 6 hours at that rate.") and constrain the desired value both above and below.

This is a dramatic shift in representation, moving from one well–structured account of train separation (e.g., the connected diagram) to another that appears more likely to remain within the confines of the working surface. Furthermore, the table explicitly records values for time that were left implicit in the original diagram. Moving beyond the given limit of 880 km with integral values for time in the table, the first evidence of standard algebraic formalism appears as a simple set of assignments. In the third and final episode, the difference between the distance apart at 5 hr (800 km) and the given distance is found (80 km). Along with the combined rate of travel, these are assembled into an algebraic proportion between the 160 km covered in 1 hr and the distance

Two trains leave the same station at the same time. They travel in opposite directions. One train travels 60 km/h and the other 100 km/h. In how many hours will they be 880 km apart?



~~Are~~ The trains are moving apart at 160 km/h?

ti	distance
0	0
1	100 + 60 = 160
2	200 + 120 = 320
3	300 + 180 = 480
4	400 + 240 = 540
5	500 + 300 = 800
6	600 + 360 = 960

Yes. It will take between 5 & 6 hours at that rate.

5 hr = 800
? hr = 80

$$\frac{160m}{1h} = \frac{80m}{Xh}$$

30 min = 80 miles

total time = 5 hr 30 min

Figure 1.1: Protocol of student W22 on the MOD problem.

remaining after 5 hr. Finding it will take 30 min to travel the remaining 80 km, the student combines partial times and presents a solution.

1.3.2 Applied quantitative inference

Where is the mathematical representation in this solution attempt? The algebraic proportion that finished the work would be the most traditional choice, but this is only an incomplete piece of the algebraic calculation that would be required to solve the entire problem. Thus, the "mathematics" of this solution attempt must also lie further upstream in episodes using nonstandard notations and quantities to construct a model of problem structure. Furthermore, where did these materials come from as the problem solver was working? The experimental setting in this study resembled a typical classroom examination, so these notations must have been constructed by the problem solver. Further, it is unlikely that their particular structure or contents were directly recalled as the solution attempt progressed (e.g., combining motion rates). This dissertation takes up these and related questions about constructing mathematical representations on paper as a form of "applied quantitative inference."

1.4 Overview of the dissertation

Research reported in this dissertation places the problem of applied quantitative inference as a central question for cognitive science, asking what must happen during problem solving for people to obtain a meaningful and effective representation of some problem. As shown in later chapters, this is an intriguing but illusive theoretical issue for traditional cognitive theories of human inference. The core of the dissertation reports exploratory empirical studies that seek to answer the descriptive question of how quantitative inferences are generated, pursued, and evaluated by problem solvers with different mathematical backgrounds. These are framed against a controversy, described in Chapter 2, over the theoretical and empirical validity of current cognitive science accounts of problems, solutions, knowledge, and competent human activity outside of laboratory or school settings.

The starting point for these explorations is that we lack an adequate theoretical language for describing how representations or problem spaces are constructed, short of their being committed to memory as a result of reading verbal instructions. Rather than asking how a representation might be used with increasing efficiency, the question is how to construct a representation of a problem in the first place. This is one excellent reason for adopting "algebra story problems" as an exploratory task, since these problems appear in a form that does not serve as a representation for obtaining quantitative precision. Chapter 2 also presents a prescriptive analysis of the quantitative and situational structure of algebra story problems, asking how these structural aspects might interact to support the construction of a meaningful mathematical representation.

Chapter 3 describes a written protocol study of algebra story problem solving among advanced undergraduates in computer science. In order to examine the activities and content of solution attempts, a relatively open–ended interpretive framework for "problem–solving episodes" is developed. This framework is applied to written protocols in Chapter 3, but also serves to organize exploration of more detailed verbal protocols in Chapters 4 and 5. The problem–solving activities of mathematically sophisticated undergraduates give a surprising image of competence among an important occupational target for standard mathematics instruction.

Chapter 4 follows these results into detailed verbal problem–solving interviews with algebra students and teachers. These provide a comparison across settings and levels of competence for the same set of problems, and the results corroborate similar generative activities *outside* the standard formalism of algebra for each comparison. Notable among these nonalgebraic problem–solving activities are "model–based reasoning tactics," in which the problem solver reasons about quantitative relations in terms of the dimensional structure of functional relations described in the problem. These tactics support different purposes within surrounding solution attempts and appear on paper as nonstandard combinations of notations and quantities that usually describe "states" in the problem's situational structure.

Chapter 5 holds these activities accountable to local combinations of notation and quantity, reinterpreting results for model–based reasoning in an ecological analysis of material designs for constructing and evaluating quantitative inferences. This analysis brings forward important relations between what material designs afford problem solvers and the complexity of episodic structure observed in their solution attempts. These analyses close the dissertation with a reappraisal of the relationship between knowledge, person, and setting and put us on a more promising track for a descriptively adequate theoretical account of constructing mathematical representations that support applied quantitative inference.

Chapter 2
Representations Prescribed: Algebra Story Problems

2.1 Quantitative inference and complex mathematical structure

A theme running throughout this dissertation is that "representations" are designed artifacts, and that to understand mediating "mental representations" we must also look outside the individual, into the material and social setting, in order to adequately describe competent use of these artifacts. In this chapter, the type of problem shown in Figure 1.1 is treated as just such a designed artifact. Algebra story problems, under this view, sit at the boundary of many interest groups, each having quite different relations to the problems and their meaning. These groups include algebra students, school alumni, and their teachers, but also the larger social worlds of professional mathematicians, mathematics educators, and behavioral scientists. Considered together, these different views of algebra story problems contribute to a broader understanding of competent mathematical problem solving.

This chapter focuses primarily on how cognitive and educational research have treated these problems. However, on the basis of interviews (described in Chapter 4) and anecdotal conversations with both algebra alumni and teachers, for many students these problems are the best remembered, the least preferred, and a terminal experience in mathematical instruction. For their teachers, these problems are often a disruption in the delivery of mathematics instruction, managed with difficulty, and sometimes dropped from the curriculum altogether.

In sharp contrast, these problems are perhaps nowhere so important as for researchers in cognitive and educational psychology. First, students' difficulties when solving algebra story problems are one of the more consistent empirical findings in empirical studies of quantitative reasoning (reviewed below). Second, these problems have become an idealized standard for complex human problem solving, partly because they are so difficult, but also because they stand in as a microcosm for a central issue in theories of learning and instruction — i.e., transfer of training beyond the instructional setting. These different perspectives on algebra story problems are crucial for understanding competent quantitative inference, since interpretations of problem–solving activity either implicitly or explicitly reflect views of "problem," "solution," and "competence" that originate in the communities that surround the artifacts themselves.

2.1.1 Chapter overview

This chapter approaches algebra story problems from several points of view. First, a brief "archaeology" of applied mathematics problems is presented, starting with representative problems and pedagogy found in the first mathematics textbook ever published with movable type. This singular historical event is examined against the historical institutionalization of mathematics instruction in this country, both in terms of what counts as realism within the mathematics classroom and shifting designs for competent quantitative inference among school alumni. This is followed by an examination of the current distribution and status of applied problems in mathematics texts, a cultural repository that stands in simultaneously as a central concern for educators and cognitive scientists. The archaeological tour is concluded by reflecting over these diverse materials in terms of the current (and recurrent) reform movement expressed as "standards" for mathematics education.

Second, empirical studies of problem solving conducted with arithmetic and algebra "story problems" are reviewed. Theoretical and methodological parallels between these research programs are used to identify shortcomings in the current stock of ideas for understanding competent quantitative inference, both on ostensibly "real world problems" in the classroom and in the world of human activity outside of school. This comparison points to a profound discontinuity between school mathematics and the quantitative inferences that people construct and reconstruct in their "everyday" lives.

Third, the structure of a particular collection of algebra story problems is analyzed prescriptively at both quantitative and situational levels. This analysis leads to a proposal for interactions between these levels, asking what happens in the gap between "words and equations" when people construct quantitative inferences while attempting to solve these problems. The review of theoretical and empirical literatures on applied quantitative inference provides starting questions for the studies presented in Chapters 3 and 4 of this dissertation. In addition, the analytic view of problem structure developed here is carried through interpretations of problem–solving activity.

2.2 An archaeology of applied mathematical problems

While the origin of mathematical knowledge, either as ideal entities or socially constructed practices, is a topic of philosophical debate (e.g., Kitcher, 1983), there can be little doubt that the genre of applied mathematical problems sometimes used to communicate or work with mathematical knowledge are designed artifacts. This section looks briefly into the history of applied problems typically found in secondary school texts. Despite a continual evolution in preferred views of competent mathematical knowledge, these problems have a remarkably durable form.

2.2.1 Word problems and practical arithmetic

The first "word problems" ever to appear in a popular edition of a mathematics textbook can be found in the *Treviso Arithmetic*, published in 1478 as an manual for training young men to serve as "computers" for commercial transactions in Renaissance Italy (Swetz, 1987). The text contains no formal algebraic notation, and instruction proceeds by presenting procedures for arithmetic calculation followed by a collection of applied techniques (i.e., the "rule of three things") for solving problems typical in commercial transactions. These calculations were primarily oral prior to the widespread adoption of Hindu–Arabic numeration, and they often presented considerable difficulties for intricate transactions. Each technique in the *Treviso* includes several worked examples, and problems cover categories for coin, cost, percent, barter, mixture, motion and work. Many of these problems would be recognizable to arithmetic or algebra students today:

> The Holy Father sent a courier from Rome to Venice, commanding him that he should reach Venice in 7 days. And the most illustrious Signoria of Venice also sent another courier to Rome, who should reach Rome in 9 days. And from Rome to Venice is 250 miles. It happened that by order of these lords the couriers started their journeys at the same time. It is required to find in how many days they will meet, and how many miles each will have traveled (Swetz, 1987, p. 158).

As a textbook, the *Treviso* presents an image both of the value of mathematics for a particular clientele (clerks and apprentices to merchants) and of beliefs about effective teaching. To relate instruction in arithmetic calculation to intended practice, the anonymous author of the text enjoins his students to attend carefully to their studies:

> As a carpenter (wishing to do well in his profession) needs to have his tools very sharp, and to know what tools to use first, and what next to use, etc., to the the end that he may have honor from his work, so it is in the work of the Practica (p. 101).

Thus, the instruction given in this text explicitly offers a mathematical belief system to its readers, in terms of the difficulty of the techniques presented, the amount of work required to master these techniques, and their eventual utility in the surrounding society. These exemplify applied problems as tools for real–world activity, a theme that holds common in instructional texts for the following 500 years. However, as Swetz (1982) points out in an historical survey of this and other practical arithmetics, even on their first printing, these applied problems diverged from commercial practice in significant ways and were little more than preparation for a longer commercial apprenticeship by students to some trading house.

2.2.2 Institutionalizing the mathematics curriculum

Mathematics in secondary and higher education has had a varied history in the United States, emerging relatively suddenly between 1820 and 1830 (Cohen, 1982) and continuing to grow into a societal norm of "numeracy" by the beginning of the 20th century. This historical period saw an increasing the number of students remaining in school through adolescence, accompanied by a standardized curriculum for mass education designed to convey "routine abilities: simple computation, reading predictable texts, reciting religious or civic codes" (Resnick, 1987a, p. 5). Against this historical development, more ambitious goals for the acquisition of diverse areas of mathematical knowledge are a more recent and still contentious proposal.

Hirstein, Weinzweig, Fey and Travers (1982) argue that the elementary algebra curricula from the turn of the century to 1960 was dominated by applied problem solving, the primary motivation being to introduce students to algebra as a tool for solving practical problems. Commercial algebra textbooks and standardized assessments during this period made frequent use of "word" problems, to the extent that most textbooks introduced manipulative operations with these problems rather than in a deductive or axiomatic method. For example, a text in the 1940's introduced operations on signed numbers with a series of stories about savings (Hirstein *et al.*, 1982, pp. 375–376):

> If a man saves \$4 a day, then 3 days from now he will have \$12 more than now. We may write: $(+3) \times (+\$4) = +\$12$

Only after every combination of multiplying signed numbers was shown in this way did the text present the "rule:"

> RULE. The product of two numbers having like signs is positive; the product of two numbers having unlike signs is negative.

Starting in the 1950's, a curriculum reform initiated primarily by professional mathematicians sought to introduce new conceptual content into elementary algebra, including a systematic treatment of variables (versus literal numbers), attention to inequalities and sets, and an axiomatic justification for manipulating traditional notations (unlike the presentation shown above). This "curriculum revolution" took a full ten years to reach implementation in the commercial market for textbooks, standardized assessment for mathematical achievement, and changes in teacher training. After this period of establishment, however, the elementary algebra curriculum contained new symbolic notations and had greatly reduced the use of applied problems, both as exercises and as demonstrations for mathematical concepts. For example, the $| x |$ notation for absolute value allowed an axiomatic definition of multiplying signed numbers, eliminating the need (in theory) for any applied development of the these operations. By incrementally replacing more traditional textbooks and training (or retraining) mathematics teachers in the more rigorous curriculum, this reform movement eventually displaced institutionalized patterns of mathematics teaching that had existed unchanged since the turn of the century.

One of the forces essential to achieving this reform was the expansion of standardized assessments of mathematical achievement, and in the early 1970's these showed falling student achievement in computational skills. These declines were generally associated with the "new math" and its introduction of rigor into the curriculum at the expense of basic skills. In reaction, many local school districts initiated "back to basics" movements that stressed calculation skills over more difficult to measure mathematical concepts. Despite these reactions and owing largely to the replacement of traditional curricular materials and a trained workforce of mathematics teachers, the elementary algebra curriculum of the 1980's has retained much of its formal content.

2.2.3 The current status of applied problems in the mathematics curriculum

These observations on the history of applied mathematical problems and the mathematics curriculum are only intended to show that definitions of numeracy have changed dramatically since the original appearance of a relatively stable genre of arithmetic and algebraic story problems. Despite their varied history of pedagogical intent, these problems have recently been adopted by cognitive and educational psychologists as a representative task for studies of complex human problem solving. In fact, Mayer's (1981) analysis of approximately 1100 story problems appearing in ten California secondary school texts has been taken as a common starting point for cognitive studies, both to describe the structure of problems and to judge their relative prevalence in current algebra textbooks.

A hierarchical taxonomy for contemporary algebra story problems. Mayer classifies story problems in this sample within a hierarchical taxonomy. At the most inclusive level, a problem belongs to a *family* by sharing a source formula with other problems, generally a three term multiplicative relation between quantities. Four of Mayer's families involve rates, and these account for 81.3% of the observed problems.

For example, all of the problems shown in Table 1.1 belong to an "amount–per–time rate" family because they involve a characteristic *output* = *rate* × *time* formula. Other families involve more general quantitative constraints in geometry, physics, or statistics (e.g., perimiters or Ohm's Law). At the next lower level, a problem belongs to either a simple or complex *category* according to its general content domain. For example, problems MOD and MRT are from Mayer's complex "motion" category, while problems WT and WC would be classified in his complex "work" category.

Within families and their categories, Mayer also classifies each problem according to its propositional *template*, which describes semantic relations between problem components (i.e., the "story" presented by the problem) and *variations* in the assigment of given and unknown values to typed quantities. For example, in Mayer, Larkin and Kadane (1984, p. 249), the template of problem MOD in Table 1.1 is given the following collection of propositions:

(RATE FOR A) = ____
(RATE FOR B) = ____
(DISTANCE BETWEEN A & B) = ____
(TIME) = UNK

In contrast, distance and time in problem MRT are rendered using different propositions:

(TIME FOR ENTIRE TRIP) = ____
(DISTANCE FOR ENTIRE TRIP) = ____

Although some of Mayer's (1981) categories appear to confound rate structure and semantic domain (e.g., amount–to–amount rates appear in motion and work domains), they present several interesting aspects of the current applied problem solving curriculum. First, most algebra story problems involve simple rates (13.5% of observed problems) or compound relations between two or more rates (67.8% of observed problems). The most numerous of these are from the amount-per-time family (27.8% of observed problems). Second, these problems reproduce each of the semantic domains for "applications" contained in the *Treviso Arithmetic*, despite a passage of over 500 years. That is, the curricular view of applied mathematical reasoning still consists predominately of commercial transactions around planning work, payment for purchases or sales, currency conversions or making change, and various investment calculations. For example, the addition of more recent scientific calculations (e.g., Ohm's law, permutations, or maximization) accounts for less than 20% of the applied problem–solving curriculum, at least as judged by the prevalence of problem types.

Aside from a comparison with 15th century mathematics texts, the existing distribution of algebra story problems might help to explain what school alumni are learning about applied problem solving. For example, Stigler, Fuson, Ham and Kim

(1986) have shown that the distribution of arithmetic word problems in first, second, and third grade textbooks in the United States tends to over–represent problem structures that are easiest for children to solve — i.e., the arithmetic solution exactly parallels the semantic structure of the problem. Stigler *et al.* hypothesize that the biased distribution of less challenging problems in U.S. texts may be the result of feedback between standardized assessment and textbook development.

In a similar survey of Soviet mathematics texts, they found problems to be distributed more equally by difficulty (as measured among U.S. school children), and the sequencing of problem types was also more varied. Although there does not yet exist a systematic body of research on problem difficulty for algebra story problems, the situation may be similar. For example, Mayer *et al.* (1984) found that frequently occuring algebra story problems have a simpler relational structure than less prevalent problems. Since relational propositions were also least likely to be recalled correctly, it may be that easier problems also predominate in algebra textbooks. In related findings, they reported that ranked frequency of occurrence for algebra story problems in textbooks correlated positively with successful recall and that conversion errors in recall resembled more prevalent textbook problems.

Curricular standards for applied mathematical problem solving. The actual distribution of algebra story problems may also be incompatible with prescriptions for curricular standards. The most recent National Council of Teachers of Mathematics (NCTM) standards for algebra in secondary schooling (NCTM, 1989) give every indication that another curricular "revolution" is underway, this time attempting to integrate aspects of mathematics as an abstract conceptual domain with more intuitive, physically motivated, or real–world curricular approaches. These calls for reform promote using multiple systems of representation to render a particular mathematical idea, including close attention to informal representational systems that students might bring into the classroom from their outside–of–school lives. Another motivating claim for current visions of reform is that algebra should serve as a conventional vocabulary for later mathematical and scientific work, including significant extensions (or transfer) to work sites outside the school setting. As with the incremental reformation of the "new math" described by Hirstein *et al.* (1982), forces of institutional inertia in commercial textbooks, the market structure of educational software, teacher training, and standardized assessment lend a certain stability to the current set of arrangments.

The point of this brief archaeology of applied mathematics problems is simply that they are objects reflecting multiple, sometimes incompatible purposes: real–world accesibility, consistent manipulation, unambiguous communication, and preparation for after–school life. Cognitive studies of complex human problem solving over the past two decades generally accept these problems and their place in the curriculum at face value, treating their structure as an idealized medium for evaluating theories of text comprehension, mathematical problem solving, or analogical inference. Given an explicit curricular rationale of transferring mathematical knowledge outside of school settings, it is ironic that the majority of "applied" problems today are oriented towards an after–school life that existed 500 years ago.

2.3 Empirical studies of applied mathematical problem solving

There is a sizable empirical literature on students' solutions of "applied" or "real–world" mathematical problems, including algebra story problem solving. These studies provide a meaningful starting point by focusing on the kinds of knowledge hypothetically required to solve these problems, the sorts of text comprehension and calculation strategies that appear to be involved in solution attempts, the domain–specific categorizations adopted by expert problem solvers, and the prevalence of analogical inference under different problem–solving conditions. The majority of these studies come out of traditional research methods in cognitive psychology, but an alternative and theoretically challenging line of analysis has developed in the fields of cultural psychology and cognitive anthropology. These studies challenge the theoretical and methodological status of "problem," "knowledge," "setting" and "competence" that are generally taken as nonproblematic assumptions in studies of quantitative inference. The challenge to these assumptions is whether they provide a suitable level of analysis, empirical methodology, or interpretive vocabulary for examining the detailed structure of individual solution attempts and how problem–solving activities interact with the social and material setting. This section reviews and compares major studies in these parallel disciplinary traditions.

2.3.1 Assessments of mathematical achievement

An indirect but important source of information about "applied" quantitative inference comes from standardized assessments of mathematical achievement. In particular, the National Assessment of Educational Progress (NAEP), established in 1969 to track the achievements of United States students in middle and secondary schooling, provides evidence for recurrent difficulties in different areas of mathematical reasoning as well as some indications of how these areas have changed in the past two decades. The brief review in this section is drawn from analytic summaries presented to the mathematics education community at the conclusion of each major assessment (e.g., see Brown, Carpenter, Kouba, Lindquist, Silver and Swafford, 1988 and Kouba, Brown, Carpenter, Lindquist, Silver and Swafford, 1988 for a summary of the fourth NAEP assessment of mathematics).

Although NAEP mathematics assessments cover a broad spectrum of topics, there are several findings in these achievement data that are important for understanding applied quantitative inference, both in the middle school years and in secondary schooling where algebra story problems appear in the traditional curriculum. First, students appear to be relatively successful at arithmetic calculation by grade 7 and at manipulating given algebraic expressions by grade 11. The manipulatory performance of students at these grade levels may even have improved over successive assessments, although arithmetic calculations appear less stable when the numbers involved are non–whole or unfamiliar (e.g., fractions or decimals). On test items designed to assess conceptual understanding of mathematical concepts underlying calculation, performance at both grade levels drops considerably. For example, in the most recent assessment only 20% of 7th graders and 34% of 11th graders chose the correct answer for "9 is what percent of 225?" (Kouba *et al.*, 1988).

Second, and in contrast to students' achievements in manipulating given mathematical expressions, when students are presented with a verbal description of some quantitative situation and asked to identify an algebraic expression that describes the situation (e.g., "the number of chairs (C) is twice the number of students (S)"), approximately half of 11th grade students with two years of algebra answered incorrectly. When verbal situations presented by items are more complex and require that the student find a precise quantitative solution (i.e., traditional arithmetic or algebra story problems), performance is highest on the least complex problems that appear most often in mathematics textbooks (e.g., "Combine" problems, described in a following section). Performance declines when problems present more complex quantitative structures (e.g., 2–step and multiplicative structures) or when problems present unfamiliar situations. For example, on a multi–step multiplication problem:

> Lemonade costs 95 cents for one 56 ounce bottle. At the school fair, Bob sold cups holding 8 ounces for 20 cents. How much money did the school make on each bottle?

only 11% of 13 year–olds and 29% of 17 year–olds were able to choose the correct answer (Carpenter, T.P, Corbitt, M.K. Kepner, H.S., Lindquist, M.M. and Reys, R.E., 1980). There is little evidence that performance on these kinds of problems has improved over successive NAEP assessments.[1]

Finally, assessments of students' attitudes towards mathematics, the instruction they receive, and the utility of both for later life present several puzzling results (Brown *et al.*, 1988).

- The majority of students in 7th and 11th grades agreed that they liked mathematics, that it was an important school subject, and that they were good at mathematics.
- In contrast, 7th graders responded that mathematics was one of their easiest courses, while 11th graders found it one of their hardest courses.

[1]Sampling strategies, timing constraints, and item presentation formats have changed over successive NAEP assessments, making detailed comparisons of changes in achievement difficult.

- Half or fewer of 11th grade students agreed that mathematics was important for getting a good job or that they expected to do work after school graduation that required mathematics.
- Half or more of both groups responded that mathematics was based on rules, that learning mathematics primarily required memorization, but that the process of finding solutions was more important than getting a correct answer.

On the one hand, students view mathematics instruction as an important aspect in their school experience, a valued but increasingly difficult curricular subject. On the other hand, they view the constituent activities and meaning of mathematics as having very narrow utility, both in school and out.

The image of school mathematics presented by these assessments, both in terms of students' achievement and attitudes towards mathematics instruction, is not very encouraging for standard visions of education that emphasize a conceptual understanding of mathematics or the development of "higher–order thinking" among school alumni. Instead, there appears to be an inversion between modest gains in skilled calculation and gradual deterioration of conceptual understanding evident in successive assessments (Resnick, 1987a). That is, it may be that students learn to manipulate mathematical notations without learning how to construct or interpret these same notations as meaningful mathematical models of situations presenting quantitative constraints. Still, patterns of performance on standardized testing tell us little about the actual phenomenology of applied quantitative inference, either inside or outside the mathematics classroom. The following sections review an extensive literature on these very issues.

2.3.2 Cognitive studies of applied quantitative inference

For a variety of methodological reasons, mathematical problems presented as a story about world events have become representative tasks for cognitive studies of complex human problem solving. First, competent performance on these problems has some obvious relevance to preferred cultural activities: sophisticated mathematical reasoning is generally thought to require high intelligence and considerable training; mathematics often stands in as a foundational discipline for other sciences and so accrues a generally high cultural regard for diverse scientific activities (i.e., the "queen of the sciences"); and these problems are commonplace in traditional mathematical schooling and its assessment, the institutional arenas where these cultural valuations are reproduced. Second, these problems provide compact, remarkably inexpensive, and relatively malleable experimental materials for traditional studies of memory and inference in complex problem solving. Third, and because of their regular form, these problems appear to provide an unambiguous interpretation of correct or incorrect performance — i.e., whether or not a subject as found a solution and thus comprehended the story and its quantitative implications. Algebra story problems are indeed a "twentieth century fable" (Hinsley, Hayes and Simon, 1977), but for many different constituencies and for many different reasons.

There are two parallel lines of empirical research on applied mathematical problem solving as it appears in school, one dealing with arithmetic word problems involving additive structures, and the other dealing with relatively more complex algebra word or story problems involving both additive and multiplicative structures.[2] Using traditional research methods in cognitive psychology, these studies focus on the kinds of knowledge hypothetically required to solve these problems, the domain–specific categorizations adopted by expert problem solvers, the sorts of text comprehension and calculation strategies that appear to be involved in solution attempts, and the prevalence of analogical inference under different problem–solving conditions.

Knowledge sources and quantitative inference. The origin of interest in algebra story problems for many of the studies reviewed in this section is Bobrow's (1968) dissertation on computer understanding of natural language texts. Rather than taking these problems as a medium for studying mathematical cognition, Bobrow wrote a program, called STUDENT, that could demonstrate an equifinal sense of text comprehension by finding a correct value for a quantitative unknown requested in a traditional textbook word problem. This was accomplished through a systematic, template–driven translation of textual phrases into variables and relations, with the application of more specialized translation templates as needed (e.g., for determining the value of dimes versus quarters in a coin problem). Similar computational techniques were later extended to the solution of word problems in calculus (Charniak, 1969) and physics (Novak, 1976), with the general conclusion that adequate text comprehension required domain–specific knowledge about both "commonsense" aspects of story entities (e.g., that Mary is a person) and their representation as quantitative abstractions (e.g., that Mary can be treated as a point of mass).

Paige and Simon (1966) used Bobrow's STUDENT as a hypothetical information processing model of human problem solving with algebra story problems, but they compared its line–by–line translation processing with think-aloud protocols taken with high school and undergraduate students of high mathematical ability. Students were presented with story problems, some of which had impossible quantitative constraints (e.g., "One piece [of a board] was two thirds as long as the whole board and was exceeded in length by the second piece by 4 feet"), were asked to find an equation using a single unknown, and were only allowed to tell the interviewer what to write on a blackboard as they worked on each problem. There were notable differences between students and the STUDENT program, the most important being students' use of "auxiliary representations" that allowed them to identify implicit relations between problem components. Detecting these constraints helped some subjects to recognize

[2]Treating the mathematical content of these problems as "structures" is an analytic extension by Vergnaud (1981, 1983) to the conceptual analysis that these problems usually receive.

physically impossible problems or to alter a problem's representation so that it was physically possible (i.e., introducing a fortuitous misconception).

Auxiliary representations were described as any sort of information that helped a student comprehend the problem situation. Thus, as in Bobrow's STUDENT, the value of currency in a coin problem was "indispensable as a supplement to the direct translation processes" (p. 85). However, students sometimes instructed the interviewer to write out diagrams that related problem components and quantities, as in a labeled number line for an impossible board cutting problem (i.e., quantities label intervals designating board length) or a variety of labeled containers for a wet mixture problem. Paige and Simon conjectured that these diagrams, or mental representations that were "functionally equivalent" to diagrams, promoted "crucial conservation assumptions [that] can be read directly off the figure" (p. 105). When students were asked to draw diagrams after solving all of the presented problems, materials included in their diagrams corroborated the distinction between syntactic translation and more "physical" approaches to modelling problem structure.

These findings present a paradox for information processing theories of complex human problem solving. On one hand, mathematically capable students sometimes approach algebra story problem solving as if they were syntactic translators, ignoring semantic constraints when asked to generate algebraic expressions. On the other hand, students also produce materials that help them to comprehend the quantitative structure of a problem before generating any equations. Similar discontinuities appear when students are given verbal descriptions of a multiplicative relation and asked to write algebraic equations. For example, asked to use variables S and P in an equation for "There are six times as many students as there are professors at this university," students with widely varying backgrounds produce a characteristic "reversal error," writing $6S = P$ rather than the correct $S = 6P$. Errors on this and similar tasks occur among 25% to 52% of students, depending on their educational background (Clement, 1982; Kaput and Sims–Knight, 1983).

Although some students do appear to adopt a syntactic translation scheme (e.g., "Well, the problem states it right off: '6 times students.' So it will be six times S is equal to professors" in Clement, 1982, p. 19), many students construct "equations" for a sensible quantitative model that compares collections of typed entities (e.g., "Six times as many students as professors... S for students... You got a ratio of six to one... S to P... OK, just ah, 6S equals ah, P, 1P" in Clement, 1982, p. 24). Despite writing algebraic expressions that are incorrect for a standard view of the multiplicative relation, these students are generally able to find precise quantitative solutions without standard algebraic manipulation, even though they are also able to correctly manipulate algebraic expressions that are given to them (Wollman, 1983).

Clearly, various kinds of "auxiliary representations" support quantitative inference among competent problem solvers. They provide an intriguing view of knowledge and activity that often differs from the formal structure of algebraic expressions, yet they may also help to generate and evaluate standard algebraic expressions. With the exception of research on semantic relations within arithmetic word problems (reviewed in a following section), more focused empirical investigations of quantitative inference outside the standard notation of algebra have not been undertaken. Thus, a number of important theoretical and empirical questions can be asked: What kinds of auxiliary representations are observed when people with different mathematical backgrounds solve algebra story problems, how do these representations support quantitative inference, and what are their limitations? In contrast, existing cognitive research has largely taken algebra story problems as a self–contained task for investigating problem categorization or recall in relation to competence or analogical inference. The relevant methodological and theoretical assumptions can be drawn directly from exemplary studies:

> [Algebra story problems provide] rigorous standards for accurate comprehension... there is a standard and widely known representation used for solving algebra word problems — algebraic equations. This representation is useful for tracing the process of comprehension. (Hinsley *et al.*, 1977, p. 89)

> People's performances differ because people possess differing information processing systems that differ and because people possess differing amounts and kinds of knowledge. (Mayer, Larkin and Kadane, 1984, p. 233)

These assumptions are certainly not incompatible with hypotheses about "auxiliary representations" that allow "conservation assumptions" to be directly constructed, but by relying on a prescriptive view of mathematical structure as a primary knowledge source and underlying mental representation, these assumptions beg fundamental questions about representation and the construction of quantitative inference.

Knowledge, problem categorization, and memory for texts. If domain–specific and mathematical knowledge are critical for solving algebra story problems, then differences in the quantity and quality of these knowledge sources should be observed when novice and expert problem solvers are compared. This has been taken up as a theoretical hypothesis in a variety of studies of problem categorization and recall, where the rationale is to show that differences in problem categorization or memory for representative problems are related to objective measures of competence.

Hinsley *et al.* (1977) presented textbook algebra story problems to high school, undergraduate, and graduate students who were successful alumni of traditional algebra instruction. When asked to sort problems by "type," these people reliably produced approximately 16 clusters that resembled Mayer's (1981) later classification of problems by category (e.g., DRT, river current, or interest). In a similar task, participants were asked to judge categories while reading problems line–by–line. Most were able to produce a correct category judgement before reading more than a quarter of the text, and some participants gave detailed predictions about unseen problem materials and about their likely solution strategies. Think–aloud protocol studies with several more advanced problem solvers (e.g., psychology graduate students) indicated that early categorization and recalled strategies were actually used when these people were asked to solve the problems. Hinsley *et al.* concluded that these findings corroborated their hypothesis that competent problem solvers possessed "problem–solving schemata":

> information about the problem categories which is useful for formulating problems for solution. This information includes knowledge about useful equations and diagrams and appropriate procedures for making relevance judgements (p. 104).

Subsequent studies have examined the relation between educational experience and problem categorization more systematically. Silver (1979) asked high ability prealgebra students to sort relatively difficult verbal mathematics problems that were designed to cross mathematical structure and content similarity. Sortings were collected before and after students both attempted to solve these problems and studied correct solutions that were given by their instructor. Sorting on the basis of similar mathematical structure was positively correlated with measures of mathematics ability and achievement, likewise with students' success in solving the problems after the first sort, and these relationships were stronger for the second problem sorting after students had studied and discussed correct solutions for each problem they attempted to solve. The relationship between problem–solving performance and perception of structural similarity was corroborated in a subsequent study by Silver (1981), and Dellarosa (1985, Experiment 1) has found differences in sorting behavior when students were required to answer schema "orienting" questions about the texts to be sorted. These studies show that students with different mathematical backgrounds or exposure to problem–specific training perceive problem structure differently, at least as measured by the way they sort problems on a brief reading. Although the sorting task itself does not allow students to solve problems, the explicit experimental rationale is to look for evidence of knowledge structures that more or less competent problem solvers would have available when doing so.

There have been similar studies of students' memory for selected aspects of algebra story problems, under the assumption that pre–existing knowledge sources would influence the way in which story problems were read and committed to memory. Mayer *et al.* (1984) reported that when undergraduates were asked to remember but not to solve representative problems, errors in cued recall were more common for propositions involving questions or relations (e.g., the rate in still water is 12 mph more than the rate of the current) than for simple value assignments, and errors of commission tended to convert relation propositions into assignment propositions (e.g., the rate in still water is 12 mph). Errors in cued recall were also more frequent for propositional materials that were irrelevant to the problem's quantitative structure (e.g., the name of a character or type of vehicle). These recall errors were interpreted as evidence that (a) relational propositions are difficult for students to represent and (b) irrelevant problem materials are not assimilated to problem schemata used during text comprehension.

These studies of categorization, memory and expertise break problem–solving activities into hypothesized stages for translation, comprehension, planning, and calculation (e.g., see Mayer *et al.*, 1984, p. 235). Under this analytic decomposition, problem sorting and recall can be taken as evidence that problem–specific sources of knowledge influence translation and comprehension processes. From a cross–sectional view, students with higher mathematical abilities sort problems in a way that corresponds to a normative account of quantitative structure, they are less likely to misrecall relevant aspects of problem texts they have read, and they are more likely to find correct solutions when later asked to solve these problems. From a longitudinal view, when students are exposed to problem–specific instruction that focuses on quantitative structure, their problem sorting begins to reflect this structure.

While there indeed appears to a relation between sorting, memory for problem texts, and problem solving performance, this relationship cannot be unambiguously attributed to differences in the knowledge structures held by problem solvers. Since students do not solve the problems they are asked to sort or recall, they may construct a differently elaborated representation of problem structure when reading or sorting. Rather than direct evidence for what students know before the experimental task begins, these relationships may also reflect the nature of categorization or recall tasks as settings for activity. For example, this would be consistent with the uninterpreted observation in Mayer *et al.* (1984) that question propositions were equally likely to be misrecalled as relational propositions when problem "comprehension" was decoupled from "planning" and "execution" of solution strategies. I return to these alternative interpretations at the end of this chapter.

Situation models, text comprehension, and quantitative inference. While studies of problem categorization and recall provide indirect evidence for students' knowledge of problem–specific information, a diverse literature focuses on the role of "situation models" in text comprehension and quantitative inference on arithmetic word problems typical of mathematics instruction at the elementary school level. These problems have a semantic structure involving combinations or comparisons of quantified sets, and these structures are managed differently depending upon the age of the problem solver (Carpenter and Moser, 1984; Riley and Greeno, 1988). These problems and the attending literature focus on the development of particular types of quantitative inference that involve the selection and procedural execution of one or two arithmetic operations. The primary factors influencing problem difficulty appear to be students' age and level of schooling, the complexity of problem structure, the types and distribution of quantities carried in the problem statement (e.g., given, unknown, or implied), the specific linguistic forms used to describe quantities and quantitative relationships, and the availability of concrete materials during problem solving.

The structure, recipients, and theoretical status of algebra story problem solving differs from arithmetic problems in several important ways. First, students asked to solve algebra story problems are older and so are generally assumed to have reached a stage of cognitive development that supports formal reasoning about quantitative relationships. This assumption can be traced to Piaget's clinical investigations of children's inferences about motion and speed (Piaget, 1970) and about functional relationships (Piaget, 1977), where the ability to reason formally about proportional relations appeared stable at 12 to 13 years of age. Subsequent investigations have challenged these findings for stable cognitive stages and transitions between them, even for familiar multiplicative structures involving speed, time, and distance (e.g., Richards and Siegler, 1979; Wilkening, 1981). Nonetheless, research on algebra story problem solving has generally assumed that formal reasoning about algebraic structures lies within the range of abilities of high school or college students. Although comparisons are made between "novices" and "experts," both groups are generally assumed to engage in formal quantitative reasoning.

Second, the structure of algebra story problems is more diverse than that found in typical arithmetic word problems. When this complexity is coupled with the explicit aim of having students use algebraic expressions as a standard notation for representing and solving these problems, the content and activity of problem–solving outcomes are often narrowed to include only these expected pedagogical materials. Despite these differences, studies of arithmetic word problem solving provide a useful comparison. Content and activities identified as important in their solution may also be important for competent inference about algebraic structures (e.g., reasoning about relations between quantified sets). In addition, general issues about the relationship between situational and quantitative structure may also apply to algebra story problem solving (e.g., the use of concrete models to identify quantitative relations or to organize calculation).

Arithmetic word problems have been shown to have a regular semantic and quantitative structure around the arithmetic operations of addition and subtraction, and this structure appears to influence problem–solving performance across levels of ability. For example, Nesher and Teubal (1974) showed that children's choice of operations for solving arithmetic word problems were biased by keywords appearing in the text (i.e., "less" or "more"), but then argued that direct translation could not be a generally sufficient process for solution because these problems were easily modified to make keywords either facilitating or misleading verbal cues. To account for sources of problem difficulty, general semantic categories for arithmetic word problems have been identified as hypotheses about levels of representation and processing intermediate that are intermediate between problem texts and symbolic operations (Carpenter and Moser, 1981; Riley, Greeno, and Heller, 1983; Vernaud, 1981).

These were summarized by Nesher, Greeno and Riley (1982) as "Combine," "Change," and "Compare" problems, each presenting a different semantic relation between quantities given in the word problem text. A Combine problem involves the composition of two quantified sets to obtain a resulting set, a Change problem involves a transformation to a starting set that yields a resulting set, and a Compare problem involves a quantitative comparison between two sets:

> (Combine) Joe has 3 marbles. Tom has 5 marbles. How many marbles do they have altogether?

(Change) Joe had 3 marbles. Then Tom gave him 5 marbles. How many marbles does Joe have now?

(Compare) Joe has 5 marbles. Tom has 8 marbles. How many marbles does Tom have more than Joe? (Riley and Greeno, 1988)

As experimental tasks, these general categories have been manipulated by changing the distribution of given and unknown quantities (e.g., an unknown starting set for a Change problem), the sequence in which collections of objects are mentioned (e.g., presenting composite before component sets in a Combine problem), and the direction of change or comparison (e.g., taking away marbles in a Change problem).

Performance on these problems improves as children become older and progress through schooling (Riley, Greeno, and Heller, 1983). Despite their simple and relatively homogeneous quantitative structure, these problems (a) were less difficult when the unknown quantity was the result of the activity described in the problem statement (i.e., a combination, change, or comparison), (b) were more difficult when a relational comparison was involved (i.e., Compare problems), (c) were more difficult when the arithmetic operation required was different from the activity described in the problem statement (e.g., when giving away marbles requires addition to find an unknown starting set in a Change problem) and (d) were more likely to be solved when children used concrete tokens to carry out counting strategies that mirrored the semantic structure of the problem. These findings have generally been interpreted as evidence that students possess or must learn to use conceptual schemata that organize different quantitative roles in the story. Furthermore, since solutions usually did not involve any written materials, children clearly were not translating directly from words into arithmetic expressions.

Many of these studies provided children with concrete materials that could be used to designate given or unknown quantities and to manipulate extensive models of these quantities when finding a solution (i.e., placing tokens in sets and counting or matching the members of sets). Comparisons of performance with and without concrete materials (Riley and Greeno, 1988) showed that concrete materials often enabled younger students to find solutions to problems they could not solve otherwise. The use of concrete or mental models in arithmetic word problem solving was taken up directly in a series of experiments by Johnson (1988). Children in kindergarten who were allowed to use concrete tokens solved Change problems at a level comparable to or below that found in previous studies (Riley and Greeno, 1988), but their performance was facilitated by changes in problem wording that made set relations more explicit (e.g., from "2 trees burned down" to "Then 2 of the trees burned down"). When a comparable group of children were asked to solve the same problems without access to concrete materials, performance fell by as much as 50%, and changes in wording no longer had a facilitating effect.

In another experiment, Johnson asked 2nd and 3rd graders to model problems with concrete materials (puppets and tokens) in one setting, and then later asked them to solve the same problems without concrete materials. Children demonstrated the correct meaning of "more than" in the modeling task (i.e., comparative difference as a set found by comparing two original sets) but not in the problem solving task, where they most often added the given numbers together regardless of problem structure. When debriefed, they reported that this was the way that "more problems" were solved in their arithmetic classes. Thus, at least for young children, access to concrete materials facilitated problem solving and appeared to enable use of more explicit linguistic information regarding quantification. Among older children with more exposure to mathematics instruction, the use of concrete materials in a modeling task enabled inferences about complex quantitative entities (i.e., comparative differences), but these inferences were not evident when children were asked to solve problems without concrete materials. Instead, the children appeared to suspend competent quantitative reasoning in favor of their understanding of "school math."

The central finding in these empirical studies is that some interaction between conceptual and quantitative understanding is necessary for competent problem solving. This interaction appears to depend upon the construction of concrete (i.e., material) or mental models of quantitative relations as a primary form of inference. Furthermore, these interactions between modeling, understanding, and calculating relate to the interpersonal setting for problem solvers' activities in ways that are not yet well understood. To explain conceptual understanding in solving arithmetic word problems, two related projects have proposed various kinds of situational and quantitative knowledge, organized as schemata that support the integration of text comprehension and problem–solving strategies. These proposals are more fully developed than the problem schemata hypothesized to support categorization and recall of algebra story problems, and they begin to address interesting questions about the content and structure of "situation models," either as mediating mental representations or as expressed in the material structure of the setting (e.g., concrete collections of tokens).

Riley and colleagues (Riley, Greeno and Heller, 1983; Riley and Greeno, 1988) proposed three stages of conceptual competence for solving these problems: the first included schemata for constructing and counting quantified sets, the second added conceptual schemata that organize quantified sets (e.g., a Set–Compare) specific to problem types, and the third added more general schemata for quantitative relations (i.e., a Part–Whole schema). Each level was implemented as a computational simulation within a production system framework, and the performance of these simulations was compared to problem–solving performance by children at different ages. The higher–order schemata contained problem–specific and quantitative relations that served to organize lower–level set schemata, and this hierarchical arrangement allowed a relatively simple planning process to coordinate between local sets, their relations, and a collection of counting procedures for simple arithmetic calculations.

For example, at the first level of competence, individual sets were constructed as successive phrases of an arithmetic word problem were read, and these sets were to correspond to the child's arrangement of concrete materials while attempting to solve the problem. Without higher–order schemata allowing a coherent interpretation of indeterminate sets (e.g., Tom has some marbles), a child at this stage of competence might disregard this phrase in the problem text and simply deliver a given quantity as the solution. The hypothesis that learners pass through these stages as they acquire specific forms of knowledge was partly supported by a comparison of empirical data on problem difficulty with the performance of computer simulations designed to reflect the influence of increasing schematic knowledge at each stage (Riley and Greeno, 1988). The fit of empirically derived and simulated performance was accurate for young children (kindergarten and first grade), but was less complete for older children and on problems involving a difference between two sets (i.e., Compare problems).

Kintsch and Greeno (1985) and Dellarosa Cummins, Kintsch, Reusser and Weimer (1988) described similar process models for different stages of competence in solving arithmetic word problems. However, they argued for a wider influence of text comprehension strategies, following the theoretical framework of van Dijk and Kintsch (1983) to include the construction of a "propositional textbase" from which logical (i.e., sets), conceptual, and quantitative schemata could be recognized and instantiated.[3] While Riley and Greeno (1988) concentrated on childrens' possession of these schemata as a key predictor of competent problem solving, Dellarosa Cummins *et al.* argued that competence with various linguistic forms determined childrens' ability to connect already well–established conceptual schemata with the specific demands of arithmetic word problems. This hypothesis was based on empirical findings that even young children were able to make complex quantitative inferences when problem texts were written to facilitate these inferences. For example, Hudson (1983) found that most preschool children were unable to solve a problem in which "There are 5 birds and 3 worms. How many more birds are there than worms?" However, most of these children found a correct solution when the question was changed to "How many birds won't get a worm?"

Dellarosa Cummins *et al.* found that structural errors in childrens' recall of arithmetic word problems corresponded to conceptual errors in their solutions to the same problems. As an explanation, they argued that children were forming an incorrect conceptual understanding of these problems and then correctly solving the misunderstood problem. Thus, developmental patterns in solution errors might be based on systematic miscomprehension of problem structure caused by linguistic difficulties rather than by deficits in conceptual knowledge. To explain these findings, they compared versions of their computer simulation in which different areas of knowledge were ablated or changed: either higher–order quantitative schemata were removed (e.g., a "SUPERSET" part–whole schema), problem–specific schemata were removed (e.g., a "TRANSFER" schema for Change problems), or explicit linkages between propositions and other schemata were changed. In the latter version, the

[3]Kintsch (1988) has since described a construction–integration model in which inferences about quantitative relations are more strongly guided by local lexical coherence, achieved by parallel retrieval and competitive inhibition during construction of a problem model.

mapping between lexical items in the textbase and problem-specific or more abstract quantitative schemata was changed to produce many of the errors characteristic of children's incorrect solutions. For example, when the mapping from "some" to an unquantified set was changed so that "some" was encoded as a simple modification, the simulation did not construct a set for the indeterminate reference, and it consequently used "default strategies" to produce an incorrect solution (e.g., reporting the cardinality of the most recently constructed set). Since changes to mappings between language terms and conceptual schemata accounted for more observed conceptual errors than removing either quantitative or problem-specific schemata, Dellarosa Cummins *et al.* argued that childrens' knowledge of this mapping could explain developmental patterns in solution difficulty.

Despite differences in the complexity of problems and the mathematical experiences of problem solvers, two findings in the literature on arithmetic word problem solving are relevant both to the specific case of algebra story problem solving and to the more general issue of describing competent quantitative inference. First, these inferences depend on a complex interaction between comprehending relatively abstract forms of language and constructing concrete or mental representations of problem structure during a solution attempt. Second, directly recalling problem-specific mathematical forms serves poorly as an explanation of problem-solving activity, even on problems presenting relatively simple arithmetic relations (i.e., additive structures). At the very least, a theory of applied mathematical problem-solving needs to account for the active sense of constructing a "situation model" as a representation that holds together diverse materials, including sources of knowledge that arise outside the pedagogical ideal of mathematical structure.

Analogical inference and problem structure. Because algebra story problems have a regular quantitative structure that can be partly decoupled from their "stories," they have frequently been used as task materials for studies of analogical inference. As with more traditional information processing tasks (Hayes and Simon, 1977; Kotovsky, Hayes and Simon, 1985; Simon and Hayes, 1976), the approach has been to manipulate systematically the surface content and deeper relational structure of problems and then to present problem solvers with different opportunities to make comparisons between related problems. Analogical inference would be seen to occur if aspects of problem structure or solution on a "source" problem were used to comprehend or solve a "target" problem.[4]

There have been several studies of the kinds of similarity judgments that problem solvers make between source and target algebra story problems. If people can reliably sort these problems into similar categories (e.g., Hinsley *et al.*, 1977), then they should also be able to judge the similarity of related pairs of problems, particularly whether two problems can be solved in similar ways. Other studies have used more direct

[4]See Hall (1986, 1989) for comparative reviews of analogical inference in problem solving and learning.

interventions to determine whether information from a source problem is actually used in the solution to a target problem.

Reed (1987) asked college algebra students to rate whether the solution to one problem in a pair would be useful for solving the second problem in a pair, without solving or seeing a solution to either problem. Pairs of problems were arranged to be the same or different in story context (i.e., motion or work, as in Mayer's "category") and the same or different in solution procedure (i.e., the distribution of given and unknown quantities). Thus, a pair was either equivalent (same context/same solution procedure), isomorphic (different/same), similar (same/different), or unrelated (different/different). When story contexts were diverse (e.g., mixture versus interest problems), students rated equivalent and similar pairs of problems as most useful, apparently preferring context over solution procedure when making problem comparisons. Only when the range of surface similarity in problem pairs was designed to be minimal (i.e., variations of work problems) were participants able to judge comparative utility on the basis of quantitative structure — i.e., equivalent > similar, isomorphic > unrelated. Unlike studies showing a positive relation between mathematical achievement or ability and categorization (e.g., Silver, 1979), students who were identified as good or poor estimators of solutions on a similar set of problems did not differ in the order of utility judgements.

Weaver and Kintsch (1988) reported a follow on experiment in which undergraduates were asked to rate the similarity of pairs of problems that systematically crossed the kind of multiplicative relation presented (e.g., rate as a relation across measured dimensions versus area as a simple product) with the form of equations that might be used to solve the problem (i.e., variations in the distribution of given and unknown quantities). Students rated pairs of problems with similar multiplicative structures as being more useful than pairs with similar equations, and these judgements improved after brief exposure to a graphical network language showing relevant differences between multiplicative relations. Unlike Reed's study, however, there was no content–level similarity between the problems used, so participants could not have been influenced by solution–irrelevant similarities between problems.

Combining these studies with evidence from the categorization experiments reviewed earlier, empirical support for schema–driven perception of quantitative structure has not been compelling. The categories reported by Hinsley *et al.* (1977) differentiated by the situational materials of story problems as well as algebraic structure. For example, DRT, current, and interest problems all share $Amount = Rate \times Time$ as a quantitative relation, yet they were sorted into different categories. Dellarosa (1985) found strong sorting by quantitative structure only when students engaged in orienting tasks that focused their attention on relevant quantitative relations. Only Silver (1979, 1981) found convincing evidence that good problem solvers, as judged by their teachers or their performance on achievement tests, sorted problems according to their mathematical structure. However, these performances were balanced by relatively poor problem solvers' sortings according to surface materials.

In short, it appears relatively difficult to get students to use quantitative structure (as designed for experimental materials) when rating the utility of one hypothetical solution for another. Only when the availability of surface similarity for problem pairs is reduced (Reed, 1987; Weaver and Kintsch, 1988) or when students are oriented towards relevant aspects of similarity by some form of instruction (Weaver and Kintsch 1988), do they reliably rate similarity in quantitative structure as useful for an analogical comparison. Although statistically reliable, the preferences observed in these studies are small — e.g., a 21% advantage for similar over different multiplicative structures in Weaver and Kintsch (1988). These would seem to be anomalous findings for a theory of applied quantitative inference based on memory for problem classes as generalized mathematical structures. Instead, the empirical findings on perception of problem similarity suggest that, at least when they are not allowed to solve problems or otherwise explore aspects of problem structure, people's judgements often reflect task demands in the experimental setting as much as *a priori* differences in mathematical knowledge.

The more direct question is whether or not people use the structure of a source problem or its solution when attempting to solve a target problem. Silver (1981) asked prealgebra students to judge whether a previously studied source solution was "mathematically related" to a target problem and then asked them to solve the target problem. Those who correctly classified the source as related were most likely to find a correct solution to the target problem. However, these students were generally more successful at solving these kinds of problems, and they did not report using source solutions in their work on target problems. Following this and related work in analogical problem solving (e.g., Gick and Holyoak, 1980, 1983), Reed, Dempster and Ettinger (1985) gave worked examples of source problems to college algebra students and then asked them to solve target problems that were either equivalent, similar (e.g., different given and unknown quantities), or unrelated to the source problem (i.e., different content and quantitative structure). Even when these students were given a hint that the source problem might be useful, correct performance on equivalent target problems only exceeded similar or unrelated problems if explanations for the source solution were elaborated and the students were directed to consult these solutions. None of these manipulations improved performance on similar and unrelated pairs of problems, although there was an increase in the number of matching equations generated on similar problems.

The difficulty of obtaining empirical evidence for analogical inference in problem solving has prompted more elaborate experimental interventions, along with outcome measures carefully designed to detect the effects of these interventions. Dellarosa (1985) hypothesized that college students would induce schematic abstractions for problem categories if they engaged in different "orienting tasks" while reading algebra story problems. On a verification task, students identified relational statements about problem components as being correct or incorrect; on a schema task, students chose between two statements describing the basic structure of a problem; and on an analogy task, students completed proportional analogies that compared elements from two different problems — i.e., a:b::c:{d1 or d2?}, where the a:b terms came from a source problem and the c:{d1 or d2?} terms came from a target problem. After the orienting task, each student was asked to match previously read and new (i.e., transfer) problems with descriptions of algebraic techniques that might be used to solve them. Finally, students were asked to attempt solutions for some of these matched problem/solution pairs.

When matching transfer problems to descriptions of algebraic solutions, students who participated in analogy and schema orienting tasks performed better than those in the verification task. However, none of these groups were better at matching problems to algebraic solutions than a group of control subjects who had simply read the same set of problems. When students worked through the problem/solution pairs they had chosen, there were no differences in reaching correct solutions between controls and students participating in any of the orienting tasks. Still, those participating in the schema task made fewer conceptual errors than analogy, verification, or control students (e.g., assigning an incorrect value to a variable in a given equation). Thus, a series of elaborate interventions designed to facilitate components of a hypothetical model of analogical inference and schema induction (e.g., detailed comparisons of local and global structure across problems) had relatively little effect, even when problem-solving outcomes were designed in accordance with these interventions.

In an attempt to improve the materials available for analogical transfer from a "source solution," Reed and Ettinger (1987) provided college algebra students with source solutions that included tables of given, inferred, and unknown quantities and then asked these students to either fill in a table for the target problem or to use a completed table when constructing target equations. Their hypothesis was that a completed table for the target problem would enable "pattern-matching" heuristics during the construction of target equations. By comparison with a control group (i.e., without tables on target problems), students asked to fill in target tables did not construct more correct equations, but completed target tables did facilitate their construction of equations when target problems were equivalent to source problems. These gains were smaller for isomorphic pairs of source and target problems, and tables did not facilitate the construction of equations on pairs of problems that were similar (i.e., with a different distribution of given and unknown values). As in Reed's other studies (Reed, 1987; Reed *et al.*, 1985), syntactic matching errors increased as target problems were less similar to source problems, suggesting that participants were responding to task instructions (e.g., hints to use source solutions) without understanding the structure of source or target problems.

The empirical results of studies of analogical inference in algebra story problem solving have not been encouraging for a theory of competent quantitative inference based on analogical comparisons between problems. Instead, a variety of interventions have had little effect on the reliable finding that spontaneous analogical comparisons are rare in experimental settings (Gick and Holyoak, 1980, 1983). These include giving a hint that a worked example might be useful, directing students to use particular source solutions, giving elaborated source solutions to insure that students have appropriate materials for transfer, intervening when problems are read to insure that students attend to relevant comparative features, and arranging outcome measures to reduce the complexity of the overall task and to focus on hypothesized effects. However, these kinds of problem comparisons should be an integral component in a theory of applied quantitative inference that depends upon problem categorization and the recall of problem–specific mathematical information (i.e., problem schemata) to explain problem–solving activities.

In summary, the empirical findings for perception of problem structure (i.e., problem sorting or judging the analogical utility of problem pairs) and analogical inference do not provide compelling support for a theory of competent problem solving based solely on recall and instantiation of problem–specific mental schemata. Problem solvers with different mathematical backgrounds indeed vary in their performance on these tasks when given algebra story problems (or comparably difficult mathematical problems), but we do not yet have a descriptive account of problem–solving activities or outcomes that moves beyond proportion correct or how equations that students produce match a normative version of quantitative structure. Unlike the situation for arithmetic word problems, we know very little about how novice or expert problem solvers manage to integrate constraints provided by the problem "situation" with quantitative relations required to find a precise solution — i.e., the very "auxiliary representations" with which Paige and Simon (1966) inaugurated this area of research.

2.3.3 Studies of quantitative inference in everyday life

Findings reviewed in the preceding sections are difficult to put aside as relatively minor anomalies in an otherwise coherent account the role of conceptual knowledge in competent problem solving. With continued research, it might be possible to design settings that bring perceptions of problem structure, solution attempts, and problem comparisons under tighter experimental control. In contrast, the research reviewed in this section moves beyond the relatively narrow confines of a single experimental setting with diverse "problems" to consider diverse settings in which the conceptual knowledge assumed in more conventional studies is ostensibly deployed. First, questions about the nature of problems and the mathematical principles thought to underly their competent solution are examined. These provide an introductory motivation to a representative collection of studies that challenge the "received view" in contemporary cognitive theories of knowledge, solving problems, and transferring knowledge or solution strategies across settings.

The nature of problems and principles. As argued in the introduction to this chapter, algebra story problems, and applied mathematical problems more generally, are boundary objects between very diverse constituencies, each having a particular interest in their structure, meaning, and solution. As pointed out by Kilpatrick (1985), a "mathematical problem" at once describes the activity of an individual, a social–anthropological negotiation between individuals, a constructed and constitutive object in the community of practicing mathematicians, and a pedagogical vehicle for theoretical notions of competent mathematical reasoning. Thus, in the research reviewed above, we are presented with descriptions of problems and solutions as a form of genuine human activity, but the theoretical "constructs" at stake and the methodological choices made to pursue them are constructed in a very specialized culture. As described from a different cultural specialization by Lave (1988a, p. 43):

> Problems of the closed, "truth or consequences" variety are a specialized cultural product, and indeed, a distorted representation of activity in everyday life, in both senses of the term — that is, they are neither common nor do they capture a good likeness of the dilemmas addressed in everyday activity. Such a culturally exotic form is more appropriate a category to *be* explained than a source of analytic terms and relations.

If "problem solving" in conventional empirical studies of school mathematics is suspect in terms of ecological validity, there are two primary sites to look for more appropriate evidence for the nature of problems and their solutions: before children are enrolled into school activities and in the after–school activities of school alumni. The story on quantitative inference before schooling is both revealing and perhaps cautionary, since children apparently reach school well aware of some important mathematical principles that govern things like numbers, sets, and counting (Gelman and Greeno, 1989). However, their ensuing school experiences sometimes disrupt these relatively principled performances, perhaps by displacing them with relatively poorly understood symbolic procedures — e.g., the difference between modeling and solving arithmetic word problems found by Johnson (1988). Although mathematics education is clearly about extending students' limited and largely implicit principles with a wider and more explicit collection of mathematical activities, instruction might be better designed to build upon what its recipients bring to school life.

Quantitative inference at work. An alternative site for ecologically valid studies of problems and their solutions is to move outside of school mathematics or psychological laboratories and into homes, markets, or everyday work sites. Mathematical problem solving occurs not only in schools, but also in kitchens, supermarkets, and a host of traditional work settings, and some of these studies show striking *intra*–individual discontinuities between school and after school competence. A diverse collection of recent studies are reviewed in this section that either move into these settings or make explicit comparisons between school and everyday quantitative reasoning.

Scribner (1984) conducted detailed descriptive and experimental studies of relatively mundane arithmetic calculations required for different types of work in an industrial milk–processing plant. By observing daily activities at different work sites in the dairy, she identified recurrent job tasks requiring calculation, was able to construct simulated tasks that varied the quantitative structure of these recurrent tasks, and then compared the performance of dairy workers, clerks, and local high school students on these simulated tasks. For example, "product assembly" work was performed by relatively unschooled (i.e., to 10th grade), blue collar workers who put together collections of various types of products according to computer–generated order forms. The quantitative structure of this task depended jointly upon the types of quantities specified on the order form and the actual array of products available for making up successive orders in the "icebox" where orders were filled. Forms expressed orders as mixed combinations of cases, which held a fixed number of product units (e.g., 16 quarts of lowfat milk to the case), and units, which appeared when a customer's order required only part of a case. Thus, an order might call for 2 cases + 8 units or 1 case − 3 units, where leftover units were added if less than or equal to half a case and subtracted otherwise. Order forms presented numerals and arithmetic operators without showing case or unit terms (e.g., "1 − 3").

Since the actual array of products available for filling an order included both full and partial cases, it was possible for product assemblers to trade off movement of physical units against mental calculation when dealing with partial cases.

> For example, if an order is 1 − 6 (10) quarts and a preloader has the option of using a full case and removing 6 quarts (the literal strategy) or using a case with 2 quarts already in it and adding 8, the literal strategy is optimal from the point of view of physical effort: it saves 2 moves. If the partial case, however, has 8 quarts and only 2 quarts must be added, filling the order as 8 + 2 is the least–physical–effort solution (the saving is 4 quarts). (Scribner, 1984, pp. 21–22)

Preloaders at work adopted least–effort strategies in 100% of the orders that allowed these savings. In a series of simulated assembly tasks designed to provide opportunities for physical effort savings, preloaders correctly used nonliteral strategies 72% of the time, inventory workers and drivers at the dairy 65%, dairy clerks 47%, and high school students 25% of the time.

Examining videotaped protocols of simulated task performance, Scribner found that students and clerks, both better educated but with less product–specific experience, tended to use a single algorithm for numerical calculation before assembling product orders in a literal and relatively inflexible way. In contrast, those with more direct product experience but less schooling were able to flexibly and efficiently construct orders according to the existing array of products. Furthermore, they gave little evidence of overt or covert numerical calculations, instead reporting "I don't never count when I'm making the order. I do it visual, a visual thing, you know" (p. 26). Similar improvisational calculations were observed among dairy drivers, who used personalized combinations of case and unit prices to reduce mental calculation when totalling the value of order deliveries on their routes. Scribner concluded that "expertise" in the kinds of quantitative inference required for dairy work was very different from the model of "expert" calculation as taught in schools. Quantitative inference at work was both motivated and constrained in ways that school mathematics tasks were not, since speed and precision in the work setting have different and more immediate consequences both for repairing incorrectly completed work products and creating local disruptions in the social organization of the setting (i.e., introducing delays into coordinated activity or requiring more work from others).

Not only can quantitative inference outside of school settings be flexibly rearranged to accommodate work situations, but complex forms of quantitative reasoning can be distributed across the social and physical working environment. Based on observations of routine calculations required for ship navigation Norman and Hutchins (1988) described the working arrangements for solving a recurrent mathematical problem (p. 10):

> Imagine a navigator who has just plotted a position fix and needs to compute the ship's speed based on the distance the ship has moved in the interval of time that elapsed between the current fix and the previous one... [T]he two fix positions are 1,500 yards apart and... 3 minutes of time have elapsed between the fix observations: What is the ship's speed (in knots)?

These were typical conditions under which a navigator might be asked to solve a $Distance = Rate \times Time$ problem. In an algebra classroom, this problem would require written calculations, algebraic manipulation of the given formula, and multiple unit conversions. In the genuinely time–limited work of a navigational bridge, however, specially designed representations — the nomogram, navigation slide rule, and "3–minute rule" — were embedded in the working environment in a way that practically eliminated the need for manipulating algebraic expressions or conducting extended written calculations. Instead, these tools required that navigators draw lines, rotate dials and align numbers, or delete trailing digits in a system of units that matched the quantities actually produced in the environment. Each was a designed "cognitive artifact" that organized common cognitive abilities by transforming the task into a representational system in which answers and methods for achieving them were accessible to workers who might not otherwise have the time or ability to conduct rapid algebraic calculations.

Quantitative inference in and out of school. Naturalistic aspects of quantitative reasoning do not occur only under the constraints of wage labor. In a study of quantitative measurement among members of a Weight Watchers program, de la Rocha (1986) found that dieters used formal measurements (i.e., calculations with fractions and unit conversions) in less than half of the kitchen circumstances calling for precise food quantities. Instead, they either used "package arithmetic" or adopted a collection of "personal inventions" that were less disruptive to the ongoing activities of preparing food for themselves and their families. Package arithmetic enacted calculations directly around the conventional packaging of commercial foodstuffs, as described by one of de la Rocha's participants:

> Well, I had a pound package of veal and I just split it into four equal pieces and made four patties, froze three and just cooked one. (de la Rocha, 1986, p. 226)

As might be expected, package arithmetic was most frequent when the units of measurement provided by commercial products matched exactly, were multiples, or fractions of the precise amounts required by the Weight Watchers program. Measurement based on inventions included personal devices like using a familiar container as a standard measure (e.g., filling a cereal bowl to a particular level), personal rules like recalling the number of spoonfuls in a half cup of corn, or personal units like cutting slices of cheese to cover a piece of bread:

> You know, if you cut three pieces of cheese to fit bread crosswise, it'll exactly weigh the way I cut it. It'll weigh out exactly one ounce of cheese. (de la Rocha, 1986, p. 233)

Use of these inventions varied according to physical properties of the foods being measured, with personal devices most often used on liquids and personal units most often applied to solid materials.

Presented as quantitative constraints embedded in dietary instructions, measurement problems were represented and resolved using materials available in the physical work setting of the kitchen. As with Scribner's (1984) dairy workers, de la Rocha's dieters achieved a form of efficiency in their measurement activities, converting "problems of scale" presented by the need for quantitative precision into "problems of sense" that made use of materials in the local setting. This conversion from scale to sense allowed them to reduce or eliminate disruptions in meal preparation brought about by the need for calculation, with the result that quantitative inferences both produced local materials and were constructed out of these same materials (i.e., personal inventions). The result, as described by Lave (1988c, p. 9), is that:

> When people own problems about quantities and their relations, they act to relate them in ways that make sense within ongoing activity. They do not "pop out" to represent them in mathematical formulas, which furnish only an impoverished representation when the world is available as a "model" of itself.

Carraher and colleagues have carefully compared strategies and performance on school arithmetic tasks with quantitative problems that occur in everyday work settings. Among children who worked as street vendors in Brazilian markets (Carraher, Carraher, and Schliemann, 1985), relatively complex calculations required to make correct change were conducted orally, using the the currency itself to organize an "add on" strategy: change was counted up, starting from the purchase price, by adding the values of currency notes until reaching the amount of money paid by the customer. During interviews conducted in the marketplace, children of various ages seldom made errors when making change (98% correct). However, when identical quantitative problems were posed as traditional school problems, only 74% of these same children gave correct responses in a word problem format, and only 37% succeeded on identical calculation exercises commonplace in school. Thus, there were significant intra–individual differences in competent mathematical reasoning on tasks that could be described as quantitative isomorphs of one another.

Carraher and Schliemann (1987) found similar activities among 3rd grade children who were not street vendors. On closer inspection, differences between oral and written calculation strategies appeared to more important for explaining their success than the actual problem situation: simulated market transaction (80% oral calculation), word problem (50% oral), or numerical calculation (15% oral). In oral calculation for addition and subtraction, a decomposition heuristic was used to partition numbers around decimal place values (e.g., 100's, 10's, 1's), and children then operated on these parts from left to right. For example, in a word problem requiring 200 – 35 a child responded "If it were thirty, then the result would be seventy. But it is thirty–five. So, its sixty–five, one hundred sixty–five" (p. 91). In oral calculation for multiplication and division, a repeated grouping heuristic was used to partition numbers by place value and then to add or subtract these amounts. For example, in a word problem requiring 75 ÷ 5 a child responded "If you give ten marbles to each, that's fifty. There are twenty–five left over. To distribute to five boys, twenty–five, that's hard. That's five more for each. Fifteen each" (p. 93).

Carraher and Schliemann argued that oral calculation was not only different from written calculation in its organization (i.e., adding on, left to right ordering), but that the quantities involved in oral calculation retained their relative value during narrated operations, whereas written calculations generally separated numbers from their value in the place system. For example, the value 10 remained "ten" during oral multiplication but was referred to as "one" during a written multiplication. Although decimal place-value and individual arithmetic operations in oral calculation were similar to the written procedures taught in school, the meaning and control over quantitative operations in each case were quite different.

These patterns are not only true of single–step addition or subtraction problems solved by children in school or the marketplace. Carraher, Schliemann, and Carraher (1988) asked construction foremen with different levels of schooling to answer questions about scale drawings that were different from their routine work in two ways: (a) the drawings did not show the scale used and (b) the implied scale was not one that would be familiar to them. The question was whether foremen would be able to flexibly modify their routine calculations by finding the implicit scale, and whether they would then be able to transfer their routine calculations to novel scales. Two strategies were observed and were unrelated to level of schooling. The majority of foremen (60%) discovered a simplified ratio by dividing quantities given in the drawing and then enacted a form of "rated addition" that incremented each part of the resulting proportion using addition or multiplication. For example, in a drawing were 5 centimeters was shown for 2 meters, a foreman found that "One meter is worth 2 1/2 centimeters," began an iterative series of calculations as "Two meters, 5 centimeters. Three meters, 7 1/2..." (p. 82), and eventually found a precise solution. These strategies showed a flexible adaptation of their conventional calculations that transferred from familiar to unfamiliar scales. In contrast, 34% of the foremen used calculations with given quantities to test whether the drawing used a familiar scale and reported being unable solve problems that involved an unfamiliar scale.

It is also possible to find ecologically interesting activity within the ongoing work of teaching mathematics in the classroom. Lampert (1986) described an innovative approach to teaching multidigit multiplication, using reciprocal telling of stories about quantitative relations by the students and the teacher. These were combined with the collaborative design of pictorial representations, maintained by the teacher at a blackboard and updated at the request of either the teacher or students. At various points during work on a problem, a member of the class would propose a representational approach to organizing quantitative information, often highly personalized and at variance with prescribed arithmetic notations. These proposals were sorted out by the class as a whole, and the teacher acted as a referee to resolve arguments between students about competing quantitative conjectures and solution methods, all with respect to the mathematical principles of multiplication that were an intended part of the curriculum. Lampert's approach used the social structure of the classroom to change the learning task presented to students: from acquiring procedural skills within a prescribed notation to constructing a representational system that conforms to principles the prescribed notation was designed to obey. Although a requirement appeared to be extraordinary teaching performance, at least part of the advantage of this approach was obtained by a collaborative exploration of representational systems for multiplication.[5]

What do studies of teaching multiplication, measuring food in the kitchen, making change in the markets of Brazil, pricing deliveries in a dairy, and finding speeds during ship navigation have in common that poses a challenge to predominate theories of competent mathematical problem solving? All these studies place the generative character of linkages between "sense" and "scale" at the foreground of competence. Rather than receiving as input or recalling a problem space for problem solving, problem solvers faced with nonroutine problems appear to *generate* representational systems that use aspects of the physical setting, sometimes placed in correspondence with various conventional notations. As problems become routine, these representations often become historical artifacts that intentionally or unintentionally embed quantitative inference and calculation directly in the work environment. These become "cognitive artifacts" (Norman and Hutchins, 1988) when their design and use give continuity to the ongoing activities of work. When the routine circumstances of work break down, these artifacts may need to be renegotiated, as in Lave's (1986, pp. 96–97) analysis of one of de la Rocha's dieters managing a seemingly obvious multiplication of fractions to find "three–quarters of two–thirds of a cup of cottage cheese" by directly manipulating the cottage cheese into a circle, marking a cross on the circle, and then

[5]Similar advantages for for collaborative problem solving and learning have been found for teaching concepts in Newtonian physics (Roschelle and Behrend, 1988) and for conventional problem–solving puzzles (Levin, Reil, Cohen, Goeller, Boruta and Miyake, 1986).

cutting away one of the resulting quadrants. Finally, the emergence of physically embedded representational systems when nonroutine problems are solved or become routine challenges theories of quantitative inference that rely on encoding and later recall of problem-specific mathematical forms as a primary explanatory construct. That is, the social and material context of problem solving is involved both in generating representational systems and in "remembering" them for later use.

2.4 Problem structure

Before presenting exploratory studies, I will examine the domain of algebra story problems at two levels of abstraction: the *quantitative structure* of related mathematical entities and the *situational structure* of related physical entities within a problem. The central issue in problem solving is to find convergent constraints through constructive elaboration of a problem representation. Structural abstractions examined in this section give two basic materials for such a constructive process. Ultimately, these and other levels of analysis may provide a relatively complete domain "ontology" (Greeno, 1983) for algebra story problems and other situations that give rise to mathematical problem solving. For present purposes, I want to identify materials that can provide a descriptive vocabulary for behavioral observations presented in later chapters and can assist intuitions about problem solving, learning, and teaching within this domain. These materials can play several roles: as a description of the task of solving algebra story problems, as a hypothetical account of the representations held by subjects during the solution process, and as an illustrative medium for teaching.

2.4.1 Quantitative structure

By the quantitative structure of algebra story problems, I mean the mathematical entities and relationships presented or implied in the problem text. A particular problem has a "structure" at this level of analysis to the extent that the relationships between mathematical entities take a distinguishable form when compared with other algebra story problems. Of course, there might be many ways of characterizing the quantitative structure of an arbitrary problem or group of ostensibly related problems — e.g., as algebraic equations or as matrices of coefficients. Bobrow (1968) used algebraic equations as a canonical internal representation of meaning for story problem texts, while Reed *et al.* (1985) used equations to define the *a priori* similarity of problems and their solution procedures. The language of algebraic equations may be sufficient for analyzing the task of algebraic manipulation, but it is fundamentally incomplete when the analysis is to include what students actually understand, construct, and use while learning to solve algebra story problems.

A network language of quantitative entities. This analysis starts with a conceptual framework originally proposed by Quintero (1981; Quintero and Schwartz,

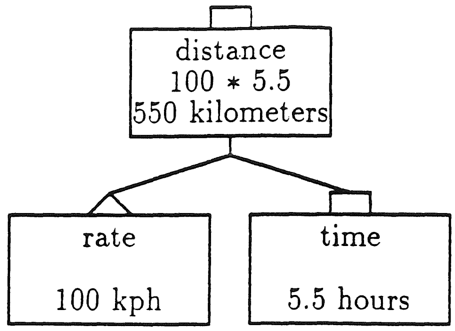


Figure 2.1: A multiplicative relation involving two extensives and a single intensive.

1981) and later extended by Shalin and Bee (1985) and Greeno (1987; Greeno, Brown, Foss, Shalin, Bee, Lewis, and Vitolo, 1986). The framework serves all three roles mentioned above: as an analysis of task structure, as a hypothetical account of subjects' representations of algebra problems, and as an instructional medium. My interest in this work is twofold. First, I will use the framework as a means for describing constraints essential for problem solution, although several additions to the framework will be necessary for it to serve as a representational hypothesis. Second, I will employ some aspects of the framework to describe how an arbitrary pair of problems might be considered similar by a problem solver.

Shalin and Bee (1985) described the mathematical structure of an algebra story problem as a network consisting of quantitative elements, relations over those elements, and compositions of these relations. Quantitative elements are divided into four basic types: an *extensive* element denotes a primary quantity (e.g., some number of miles or hours); an *intensive* element denotes a map between two extensives (e.g., a motion rate relates time and distance); a *difference* element poses an additive contrast of two extensives (e.g., one time interval is 2 hours longer than another); and a *factor* element gives a multiplicative comparison of two extensives (e.g., one distance is twice another). Composing these elements according to their type yields quantitative relations. A quantitative relation is defined as an arithmetic operation (i.e., addition, subtraction, multiplication, or division) relating three quantities. For example, the fact that a train traveling 100 km/h for 5.5 hours covers a distance of 550 km can be expressed as a relational triad over two extensives (550 kilometers and 5.5 hours) and a single intensive (100 km/h) as shown in Figure 2.1. Each element is presented graphically as a box containing several expressions. The shape at the top of the box designates element type — e.g., a rectangular top designates an extensive, a triangular top an intensive.

As an additional level of structure, relational triads can be composed by sharing various quantities to yield "problem structures." These are quantitative networks describing typed quantities and constraints among them. As shown with solid lines[6] in Figure 2.2(a), a single quantitative network can be used to graphically represent the

[6]Portions of the network in dashed lines will be discussed shortly.

problem of trains traveling in opposite directions (problem MOD from Table 1.1). Sharing a common time, two rates combine through multiplicative triads to yield parts of the total distance. These parts are combined in an additive triad to give a single extensive quantity representing the total distance. Figure 2.2(b) shows a quantitative network corresponding to the round trip (MRT) problem. In both networks, the term "output" serves as a generalization over distance and work.

Taken together, quantities, relations and structures provide a language for describing the quantitative form of particular algebra story problems. While a variety of equivalent graphical languages might be used (e.g., parse trees for arithmetic expressions), this language gives explicit representational status to mathematical entities, associates a quantitative type with each, and incorporates a metaphorical sense of storage for holding semantic information (e.g., textual phrases) and intermediate calculations. Constraints on the arithmetic composition of typed quantitative entities restrict the space of possible quantitative relations (Greeno *et al.*, 1986). For example, the multiplicative composition of intensive and extensive quantities (rate and time) in Figure 2.1 is allowed, while an additive composition of the same quantities would be disallowed. Greeno (1987) points out that constraints are also available from compositional restrictions on the units of measurement for quantities,[7] although the network language does not presently incorporate these constraints. Finally, the interconnectivity of a quantitative network supports a form of written algebraic calculus. Expressions can be propagated through the network with the goal of finding convergent constraints on the given unknown.

Quantitative networks provide a visually accessible notation for comparing the structure of different algebra story problems. However, the notation and compositional constraints do not specify which of the permissible quantitative structures a person might generate when solving an algebra story problem. For example, the quantitative network shown with solid lines in Figure 2.2(a) describes the opposite direction problem *after* several crucial inferences have occurred: component distances have been inferred within the total of 880 kilometers, and a single extensive quantity for travel time has been correctly inserted in the network. For the same problem, consider elaborating the quantitative network to include network components shown with dashed lines in Figure 2.2(a). We might imagine a problem solver inferring that the given rates can be added. The resulting combined rate (160 km/h), when multiplied by the unknown time, gives the total distance without adding constituent distances. Empirical studies with this and similar problems (described in Chapters 3 and 4) show considerable variety in the solution approaches taken by different people as well as by individuals within a single problem-solving effort.

[7]An instructional system developed by Schwartz (1982) enforces unit constraints to help users avoid irrelevant calculations, particularly when using intensive quantities. Thompson (1988) combines quantitative networks and unit constraints in another system named "Word Problem Assistant."

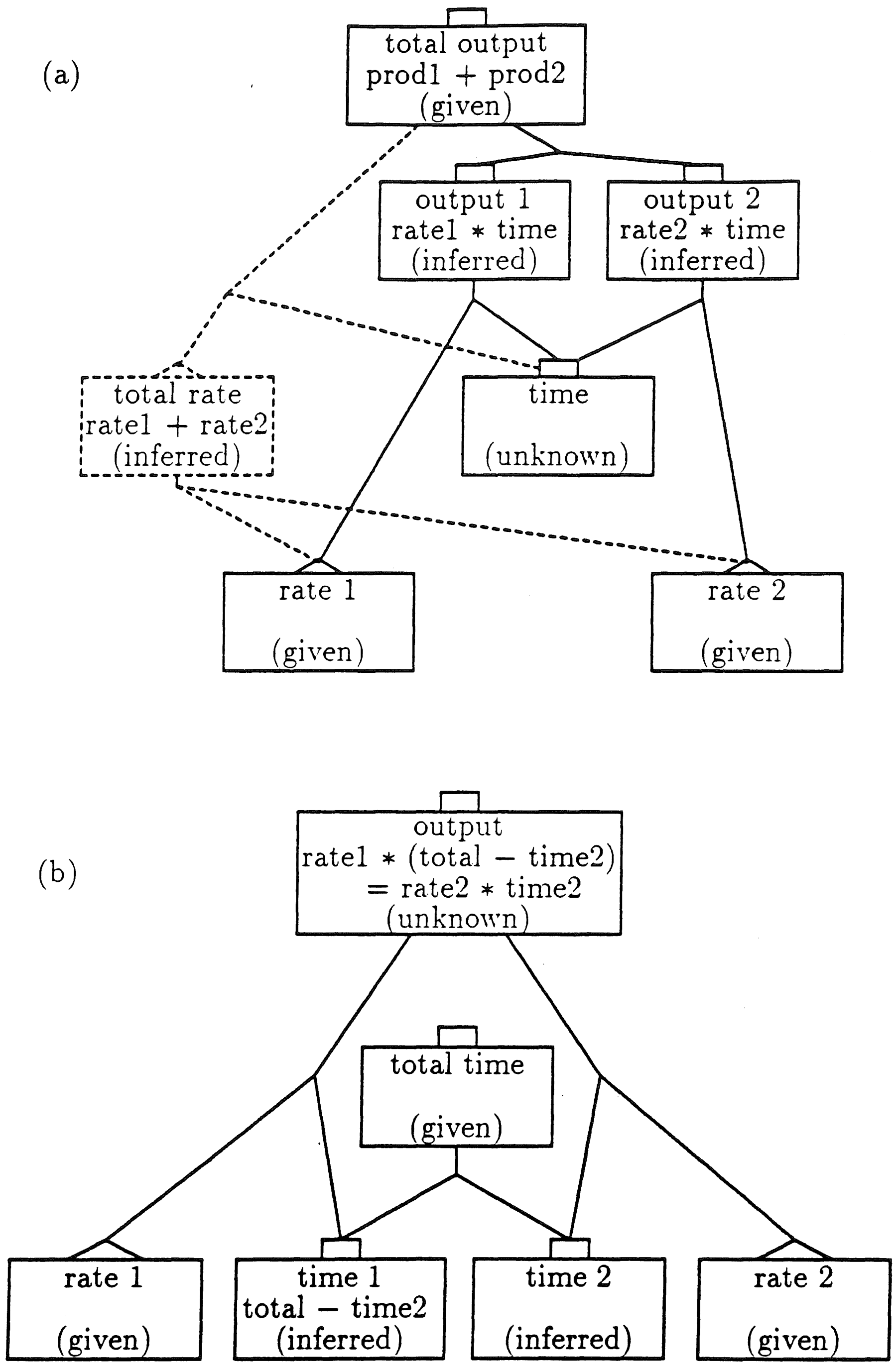


Figure 2.2: The quantitative structure of two problem classes: (a) contains problems MOD, WT while (b) contains MRT, WC.

The networks shown in Figure 2.2 are idealized graphical representations of problem structure that might be constructed by problem solvers who understand the quantitative network language and are able to use the language to comprehend and solve algebra story problems. These networks give a particular quantitative representation, but their content is largely the result of inferential processes that draw on other knowledge sources. These processes may include:

- recognizing quantitative entities directly contained in or implied by the problem text,
- composing these entities into local relational structures,
- composing relational substructures into larger problem structures,
- recognizing familiar substructural arrangements, and
- detecting when constraints are sufficient for solution.

The results of each action lie within the quantitative formalism for which Shalin and Bee's (1985) framework provides a functional description. However, the enablement conditions for these actions or the knowledge sources that support them lie partly outside the formalism. These issues are explored further when I describe the situational structure of problems.

Quantitative networks as problem classes. Quantitative networks provide an analytic tool for examining aspects of quantitative similarity. At the level of entire problems, this approach gives a stronger basis for mathematical similarity than simply noting common equations. At a more fine-grained level, there may be significant areas of substructural isomorphism in globally dissimilar problems.

The problems from Table 1.1 can be grouped into structurally similar pairs as follows: MOD/WT and MRT/WC. Each problem in a pair is a "quantitative isomorph" of the other, as shown graphically in Figure 2.2. In the MOD/WT pair, extensives for kilometers traveled correspond with those for parts of a job completed (output1 and output2). In the MRT/WC pair, a round trip travel extensive corresponds with an extensive for boxes filled and then checked (output). Comparing problems within each pair, extensive and intensive quantities play identical roles in the surrounding network structure. However, when comparing problems across pairs, structural roles of similar quantitative entities change or are even reversed. For example, the additive extensive relation for combined distance (or work output) in Figure 2.2(a) is locally similar to the additive extensive relation for combined time in Figure 2.2(b), but these relations play very different roles in their overall quantitative structures. From a normative view, a specific quantitative network defines an equivalence class of algebraic problems, each of which may have a different situational instantiation. Of course, being directly similar in form does not mean that problems must be solved in the same way. Figure 2.2 presents the quantitative structure of problem materials required for a quantitative solution. We could as well depict the quantitative structure of intermediate representational states apparent in people's solutions, an exercise that sometimes leads to a surprising sequence of graphical images as various conceptual errors are introduced or repaired.

Turning to a finer-grained level of comparison, classes of problems are similar to each other by sharing particular quantitative substructures. A substructure is a subgraph within a larger quantitative network consisting of stated quantities, inferred quantities, and relationships among these quantities. For example, "current" problems are similar at a quantitative level because they contain an additive relationship between the rate of the vehicle (steamer, canoe, etc.) and the rate of the medium in which it travels (current, tide, etc.). While other aspects of the quantitative structure for a pair of current problems can be dissimilar, such a shared substructure may contribute to subjects' estimates of problem similarity. As in the results of Hinsley *et al.* (1977), similarity judgments at the level of "river current" versus "DRT" problems may appear an educational failure: problem solvers demonstrate content-specific categorizations when the instructional goal was to facilitate their learning of mathematical forms. Another interpretation is that quantitative substructures are learned through instruction and problem-solving experience and thus form part of the underlying competence in this domain. Since particular substructures are correlated with problem types, the resulting categorizations appear overly content-specific. However, there may be a functional or pragmatic basis for learning these problem classes: despite dissimilarity of overall mathematical structure, shared quantitative substructures require similar partial solution strategies.

2.4.2 Situational structure

The quantitative network formalism does not attempt to account for the problem structures that subjects actually generate during problem solving, although some constraints are placed on combining quantitative types into relational triads. In this section, I examine another level of analysis — the situational structure of a story problem — as a source of additional constraint when people construct a solution-enabling representation of an algebra story problem. This view of the situational structure of an algebra story problem is not synonymous with what other researchers have called "surface content." Although surface materials like trains, buses, or letters are important problem constituents, and may be particularly so for novice problem solvers, I will not focus on these materials.

Instead, I present a language for describing the situational structure of "compound" algebra story problems involving related linear functions, and use the language in a detailed examination of problems involving motion or work[8] (see Table 1.1). As with the quantitative network formalism, this language for describing the situational structure of problems can play several roles: as an analysis of problem structure, as a hypothetical cognitive representation, or as an educational medium. Here I develop a relational language for describing problems, argue for its utility in generating key problem-solving inferences, and then use the language to present a viewpoint on the

[8]Motion and work are frequently used as the setting for story problems in algebra texts, comprising 20% of an extensive sampling by Mayer (1981).

space of possible algebra story problems that is complementary to problem classes based on quantitative structure. In later chapters, I also use the language to help interpret various activities observed in exploratory studies of algebra story problem solving and then to consider the theoretical and educational implications of these findings for mathematical competence.

A relational language of situational contexts. Basic terms of the relational language are presented first, followed by an example of their use shown in Figure 2.3. Compound motion and work problems are assembled around related events — e.g., traveling in opposite directions, working together, riding a bus and walking, or filling envelopes and checking them. In each event, an agent engages in activity that produces some output (distance or work) over a period of time. Hence, output and time are the basic *dimensions* that organize story activities. These activities start and stop with particular times, locations, or work products that can be modeled as *places* along the appropriate dimension. Places that bound an activity define particular *segments* of output or time, and these segments can be placed in relation to each other within a common dimension.[9] Rates of motion or work give a systematic correspondence between dimensions of output and time, and using rates in the solution of a quantitative problem requires a strategy for integrating these dimensions. Arranging output and time dimensions orthogonally gives a rectilinear framework in which rate is a two-dimensional entity. Rate entities are modeled as *inclines* that associate particular output and time segments. Relational descriptions involving typed dimensions, places, segments, and inclines provide a language for expressing the *situational context* of an individual problem.

The situational context of problem MOD (from Table 1) is shown in Figure 2.3. Parts (a) and (b) of the figure show place and segment representations for output (distance) and time, while part (c) of the figure shows an orthogonal integration of these dimensions with time on the vertical axis. In part (a) of the figure, trains traveling in opposite directions from the same station provide two spatial segments (Distance 1 and Distance 2) sharing a place of origin (S) but with unknown places for destinations. These segments are collinear and oriented in opposite directions. Since trains leave from the same place of origin, these distance segments are also adjacent and can be arranged within the horizontal dimension shown in part (c) of the figure. In part (b) of the figure, trains leave at the same time (t0) and are separated by 880 kilometers at some later time, providing time segments (Time 1 and Time 2) that are congruent (i.e., coinciding at all points) when arranged within the vertical dimension. This assumes collinearity and the same directional orientation for all time segments. Representing rates of travel as two–dimensional inclines, part (c) of the figure puts particular instances of output and time in correspondence (e.g., 60 versus 100 kilometers in the first hour of travel). Inclines can either represent a concrete situation, as shown here, or an invariant relation between output and time dimensions. Treating

[9]Segment relations within a dimension are similar to Allen's (1983, 1984) relational descriptions of temporal intervals.

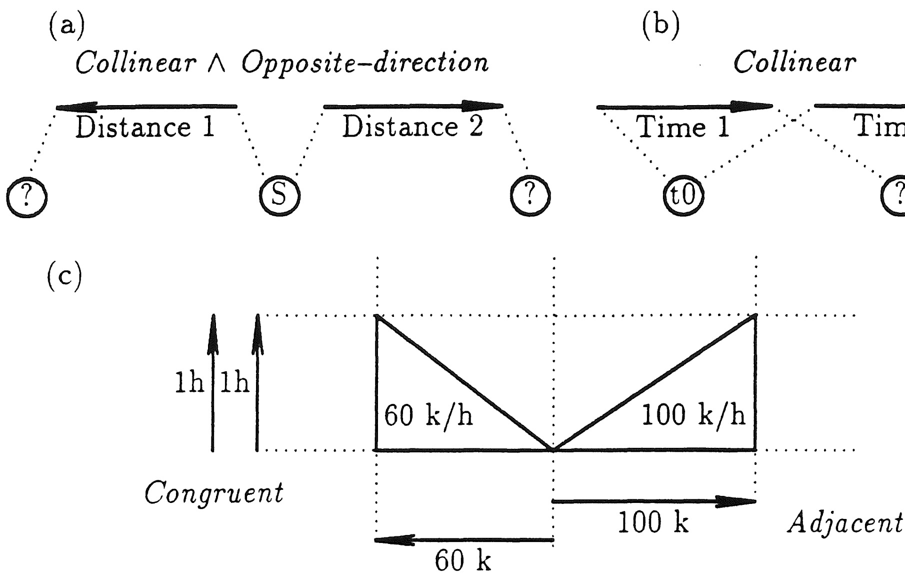


Figure 2.3: A situational context for motion in opposite directions: (a) and (b) show places and segments for output and time, while (c) shows inclines for rates when these dimensions are arranged orthogonally.

rate as an invariant relation approaches the algebraic sense of rate as a linear function. Each interpretation enables different problem–solving activities, discussed below.

Problem–solving inferences based on situational contexts. Before using this relational language to describe a space of situational contexts for algebra story problems, I briefly consider its utility as a representation for problem solving. First, I describe how a representation of situational context like that shown in Figure 2.3 could be constructed; second, I consider how this relational description might be useful for problem comprehension and solution. Both questions bear on the role of this situational language as a representational hypothesis and as an instructional medium.

On the issue of how these representations might be constructed (either spontaneously or as an educational exercise), I will propose a series of *constructive inferences* that operate on a case frame representation[10] of the events described in the text of a compound algebra story problem. These inferences build a situational model of the problem by assembling a relational description of a particular situational context. Assuming the case frames contain roles that specify typed places and segments (e.g., the starting location versus the starting time), a problem solver could model these roles as situational places and segments within output and time dimensions. From these initial entities, a series of elaborative inferences identify places and segments implicit in

[10]See Brachman (1979) for a review of related representation schemes and Kintsch and Greeno (1985) for an example of a case frame representation for the text of arithmetic word problems.

the problem statement and relations over segments within each dimension. What results is a relational description of situational context as in Figure 2.3. Constructive inferences that assemble a relational description of situational context are similar to the comprehension strategies that Kintsch and Greeno (1985) use to take propositional encodings of arithmetic word problems into a set-based representation.

On the issue of utility, I suspect that segment relations within situational dimensions support the construction of quantitative representations like the networks of Shalin and Bee (1985). For example, knowing that spatial segments are collinear and adjacent while times are congruent supports two useful problem-solving inferences in problem MOD: constituent distances can be added to yield a total distance, and the rates of each train can be added to give a combined rate. The first inference is a necessary quantitative constraint for solution (i.e., a "conservation assumption" for Paige and Simon, 1966), while the second inference effectively compresses the compound problem into a simpler problem which can be solved without extended algebraic manipulation. These are precisely the inferences about problem structure that were not accounted for in an examination of quantitative structure. For example, the network components shown with dashed lines in Figure 2.2(a) would result if a student decided to add motion or working rates. Hence, in addition to constructive inferences that build a situational context, there are also *constraint-generating inferences* that take descriptions of situational structure into quantitative relations. Each inference about a quantitative constraint, supported by relevant situational relations, gives a substructural component in a larger set of constraints that may enable a solution.

It is also possible to use dimensions, places, segments, and inclines directly in a solution attempt by treating these representational entities as a model of the problem situation. I will develop a prescriptive account of *model-based reasoning* as a problem-solving tactic here. Following chapters introduce operational categories for interpreting this tactic within the structure of written protocols and give an empirical account of its use and consequence in algebra story problem solving.

Placed within a single dimension to model time or output, segments provide an explicit spatial representation that enables a variety of problem-solving operations like "copying," "stacking," "comparing," or "decomposing" their one-dimensional extent. Similarly, using inclines as models of rate enables operations like "joining" or "scaling" their two-dimensional extent. Joining, shown in part (a) of Figure 2.4, places copies of the concrete incline along the diagonal in an iterative fashion. Scaling, in part (b) of the figure, treats the incline as an invariant relation by estimating the extent of a segment in one dimension and then projecting that value through the incline to generate an associated extent in the other dimension. Each operation is based on a different interpretation of rate as a relation across dimensions, and each coordinates operations on associated segments within single dimensions.

Both join and scale operations enable problem solving by model-based reasoning without requiring algebraic representation. Figure 2.5 shows hypothetical solution attempts using join and scale operations on the opposite direction motion problem

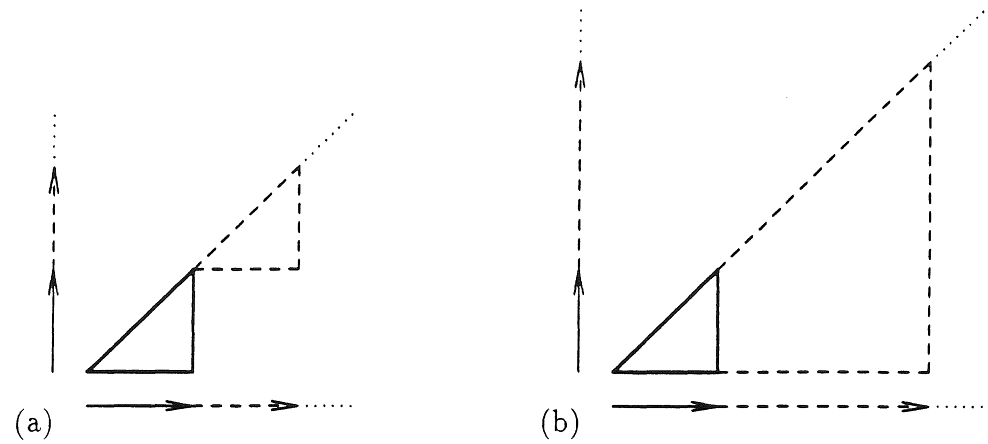


Figure 2.4: Operations based on different interpretations of two–dimensional inclines: (a) shows a concrete situation successively "joined" to give an iterative simulation of states within the problem model; (b) shows an invariant relation "scaled" to give a heuristic estimate of a final state in the model.

(MOD). Treating inclines as concrete entities in part (a) of the figure, the join operator enables an iterative simulation over five successive one hour increments in the time dimension. These correspond to intermediate states in a two–dimensional model of the problem, successively constructed and tested against the given constraint of being 880 kilometers apart after a common interval of time. Treating inclines as invariant relations in part (b) of the figure, the scale operator enables a heuristic estimate of the problem's final state by choosing five hours as the time at which the trains will be 880 kilometers apart and projecting this choice of a common time through each incline to find associated distance segments. In both solution attempts, spatial relations within the two–dimensional model support and organize relatively simple quantitative operations like addition, multiplication, and value comparison. Thus, even without utilizing the metric qualities that such a model might afford (e.g., testing whether adjacent distance segments precisely "fill" the composed 880 kilometer segment), model–based reasoning can lead to a solution without explicitly constructing an algebraic representation of the problem.

While entities and operations in model–based reasoning can support solution attempts directly, they also provide a vocabulary of problem–solving activities that could be used to construct an algebraic representation. For example, introducing a variable, t, as a label on the unknown common time in part (b) of Figure 2.5, we can use the scale operator to project that variable into expressions for labels on each distance segment in the horizontal dimension. Since these segments are adjacent and must fill the given combined distance of 880 kilometers, addition of label expressions in the horizontal dimension gives an algebraic expression for the combined distance, $100t + 60t = 880$. Thus, model–based reasoning operations can also participate in *constraint–generating inferences* described earlier.

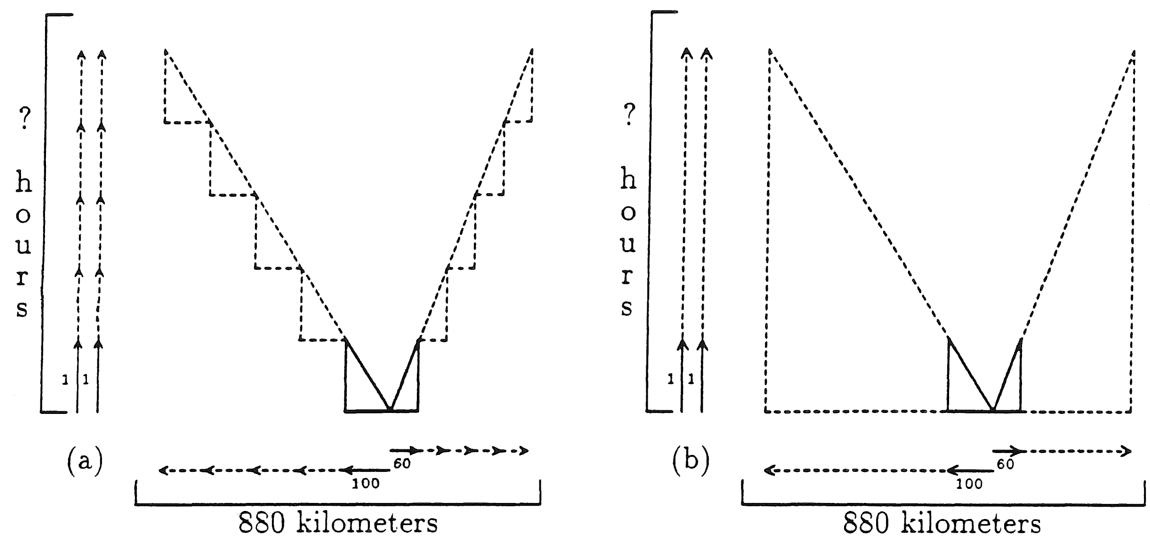


Figure 2.5: Solution attempts using model–based reasoning on problem MOD: (a) "joins" successive concrete inclines in an iterative simulation; (b) "scales" inclines as an invariant multiplicative relation in a heuristic estimation.

In general, inferences in model–based reasoning correspond to relatively opaque operations in the algebraic formalism (e.g., distribution of a product). Their spatial character and granularity may provide an accessible problem–solving medium for subjects who are newcomers to the algebraic formalism. In addition, the results of these operations could justify more abstract activities in an algebraic or quantitative network representation, allowing problem solvers to verify quantitative constraints or results about which they are uncertain. Evidence for these hypothetical roles of model–based reasoning, even in competent problem solvers, is presented in the chapters that follow.

Situational contexts as problem classes. Beyond their role as a representational hypothesis or an instructional medium, situational contexts provide a viewpoint on the space of possible compound algebra story problems that is complementary to the problem classes provided by quantitative structure. Even if we restrict analysis to compound motion problems in which movement must be collinear and directed, a variety of situational contexts are possible. Taking two collinear distance segments we can select from a set of spatial relationships (e.g., congruent or adjacent) and combine this selection with directional orientation (e.g., same or opposite) to yield a distinct spatial situation. Also selecting a relation between time segments (e.g., congruent or adjacent), we can combine segment relations for distance and time dimensions to yield a particular situational context for a compound motion problem. For example, problem MOD has adjacent distance segments oriented in opposite directions and has congruent time segments, yielding the situational context used in Figure 2.5.

A similar approach is possible with compound work problems. Work outputs can also be modeled as collinear segments, although their directional orientation is less

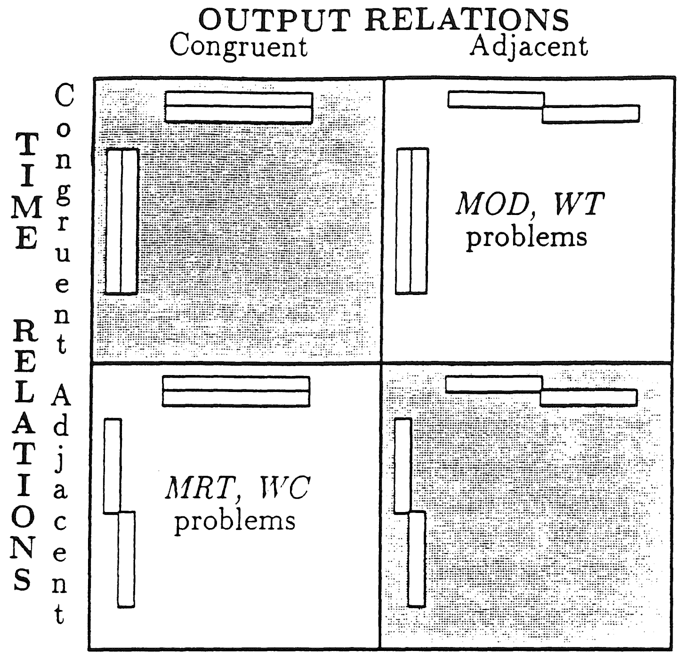


Figure 2.6: A matrix of situational contexts for pairs of isomorphic motion and work problems.

directly interpretable. In the present analysis, I exclude a sense of direction for work outputs. Working "together" can be modeled as adjacent output segments and "competitive" work as congruent output segments. For example, the work together (WT) problem has adjacent output segments that add to yield a single job and congruent time segments that, in concert with additive output, allow addition of working rates. This corresponds directly with the situational context of problem MOD, without directional orientation of output segments. The competitive work problem (WC) can be modeled in a similar fashion. Since Randy and Jo each work on the same set of boxes, we choose congruent segments to model the same output. Adjacent time segments are associated with the completion of each output, leading to a direct situational correspondence with the round trip problem (MRT).

Figure 2.6 shows a matrix of situational contexts formed by crossing segment relations from output and time dimensions. Compound motion and work problems in each cell have a common situational structure (e.g., problems MOD and WT in the upper right cell), and off–diagonal cells contain pairs of problems that reverse segment relations for time and output. For example, reversing adjacent distances and congruent times in problem MOD produces problem MRT, provided that opposite directions are retained in both problems. Problem structures in diagonal cells of the figure (shaded) are not used in this study but also provide the basis for particular algebra story problems. For example, the lower right cell of Figure 2.6 contains what Mayer (1981) called "speed change" problems. This constructive approach to situational contexts can be extended to larger relational vocabularies for output and time (e.g., including overlap, disjoint, etc.), yielding a sizable space of situational contexts that provide the dimensional basis for algebra "stories" about motion and work.

These examples show that a language of dimensions, places, segments, and inclines can be used to model compound motion and work problems. I have also examined the coverage of this language over different classes of algebra story problems, like those included in Mayer's exhaustive taxonomy (1981). Useful models of situational context can be constructed for most of these classes, including current, mixture, simple interest, cost, and coin problems. Some extensions of the language appear necessary to model relational constraints involving additive and multiplicative comparisons (e.g., "12 more than" or "twice as fast as"). In general, however, models of situational context are possible for any problem in which related linear functions can sensibly be shown within two dimensions. Although arbitrarily complex quantitative relations can be graphed in a Cartesian plane, the provision that their dimensions be "sensible" restricts our modeling language to situations where one–dimensional relations like adjacent and two–dimensional operators like "joining" or "scaling" have meaning. Thus, dimensional models of situational context may be applicable beyond textbook algebra story problems and include everyday situations involving related linear functions.

Comparison of situational and quantitative structure. Isomorphism within cells and reversed structure across cells of the matrix in Figure 2.6 partition the space of compound algebra story problems in a way that is complementary to the problem classes described in the preceding section on quantitative structure. In fact, the problems paired in each cell also have an isomorphic quantitative structure, and problems from off–diagonal cells reverse quantitative relations. For example, an additive triad over distance extensives in problem MOD contrasts with a shared extensive for distance in problem MRT. I will argue that this complementarity arises precisely because the quantitative substructures serve as a mathematical abstraction for describing situational contexts. In turn, a relational language of situational contexts provides an abstraction for describing (or modeling) events within particular problems. Thus, choosing segment relations for output and time gives rise to an organized space of situational contexts for compound motion and work problems, each with a corresponding quantitative structure.

While quantitative and situational viewpoints on algebra story problems are complementary, they are not identical. The quantitative network formalism models conceptual entities of time, output, and rate as abstractions that preserve quantitative type (e.g., extensives versus intensives) and value, either as a number or an algebraic expression. In contrast, situational segments and inclines model these same entities as individuals that preserve semantic type (e.g., time versus output), dimensional order (i.e., segments versus inclines), quantitative value, a physical sense of extent (i.e., the length of a segment or the slope of an incline), and local "spatial" relations between individual instances of extent (e.g., the 60 and 100 kilometer segments after the first hour of travel are adjacent). Preserving physical extent and relations of locality may allow problem solvers to utilize spatial knowledge when identifying or verifying quantitative constraints. For example, when a total distance can be decomposed into component distances which exactly fit within the total, there is a direct physical justification for their addition. "Joining" or "scaling" inclines using a two–dimensional model of rate promises a similar physical justification for operations on intensive quantities. Whether students actually use such a vocabulary for justification is an open question that is addressed in later chapters. My hypothesis is that shared aspects of situational structure, in addition to quantitative structure, contribute to subjects' judgments of similarity between an arbitrary pair of algebra story problems.

Quantitative and situational structure are not the only materials in the domain of algebra story problems that are important for problem solving, learning, and teaching. Neither can we tacitly assume that these structures, as described above, are actually held by subjects during problem solving. However, these structural abstractions may help to understand what subjects actually do when confronted with a problem to be solved, and to hypothesize what must be learned for competent problem solving to be achieved. Knowledge sources that guide the generation of quantitative representations, and the manner in which they are manifested during problem solving, comprise an important part of competent performance. By grounding quantitative structure in conceptual understanding, these knowledge sources may allow a problem solver to effectively assemble and validate representational structures and operators in the algebraic formalism.

2.5 Discussion

This chapter draws across several disciplinary boundaries to frame concerns for a descriptively adequate theory of applied quantitative inference. The history and cultural status of applied problems as instructional materials is reviewed first to show that these problems are quite "real" for very different communities. Next, studies of applied mathematical problem solving from different perspectives are reviewed. On the one hand, traditional cognitive science approaches to these problems have refined the actual phenomenology of problem solving into a collection of loosely related theoretical components that do not provide a coherent or empirically compelling description of human activity. Most of the large collection of studies reviewed in this chapter provide interesting glimpses of quantitative inference, but their starting assumptions about problem–solving and the methodological consequences of these assumptions give a necessarily fragmentary view of the whole activity. On the other hand, studies of quantitative inference in cognitive anthropology and cultural psychology move between classroom, laboratory, and work settings to describe problems and their solutions as integrated activity. These studies also provide an interesting view of both "school" problems and "after–school" or everyday problems.

In light of these literatures, the structure of algebra story problems is examined at both quantitative and situational levels, with the explicit purpose of leaving open to empirical study the fine–grained episodic structure of solution attempts evident in the written protocol example shown in Chapter 1. The account of problem structure that results is prescriptive at several levels: as an analysis of "tasks" for protocol studies of algebra story problem solving, as a collection of entities for hypotheses about mental or concrete representation, and as an illustrative artifact with some potential for instruction. These prescriptive aspects are carried through the remainder of the dissertation, influencing the kinds of problems presented to study participants and the analysis made of their responses. However, they do not delimit either the interpretive vocabulary for observed problem–solving activity or the kinds of explanations that are possible for these observations — i.e., interactions between quantitative and situational aspects of "problem structure" is not proposed as a closed system or a simulation of human activity. This concluding section attempts to find meaningful points of intersection within the diverse collection of reviewed studies. The general conclusion is that much remains to be understood about even seemingly mundane cases of applied quantitative inference.

2.5.1 Towards an ecological theory of applied quantitative inference

Finding a meaningful intersection between traditional cognitive science studies of human inference and studies of "everyday" cognition is an explicit theme of this dissertation. As this theme is developed over the remaining chapters, applied quantitative inference is reconsidered as constructive activity that both responds to and acts upon the immediate material setting. In this sense of interaction, problem solving can be studied as an ecological system that opens up the received view of "representation" in cognitive science to include correspondences between aspects of mental, material, and interpersonal experience. This theme resurfaces throughout the remaining chapters. Several points of commonality and difference remain implicit in the preceding reviews, and they are worth bringing forward here. To simplify the discussion, I will use the term "cognitive" to describe traditional cognitive science accounts of quantitative inference and the term "ecological" to describe studies of quantitative inference that move into the wider settings of "everyday" life. Neither term is offered as a satisfactory description of these diverse research programs.

A primary difference between these perspectives is the status of "problems" and their "solutions" in human activity. The cognitive approach generally takes applied mathematics problems, as they appear in textbooks, as representatives of an idealized class of tasks. From this vantage, the characteristics of existing textbook problems provide an objective definition of a "problem" (e.g., Mayer, 1981) that can then be manipulated to assess the relative influence of different characteristic features. Likewise, the correct solution of these problems is best described within the standard ontology and notation for expressing and manipulating quantitative relationships, in this case algebraic expressions. In contrast, the ecological approach starts with human activities in particular settings and asks what recurrent disruptions in these settings require changes in activity. Thus, the need to make formal calculations while cooking (de la Roche, 1986) presents a "snag" in ongoing activity that can be framed as a problem. The changes in activity around this problem comprise its solution, and both may have relatively little to do with conventional mathematics instruction. I focus on the "task" of solving algebra story problems based on the observation that these problems are recurrent disruptions in the instructional and assessment experiences of many people. This is not to insist that these problems are "real" substitutes for after-school life.

A closely related question is the issue of ascribing "competence" to individuals on the basis of their solutions to problems. In cognitive studies, competence is often operationalized precisely around the actual solution offered for a representative problem (e.g., an unknown value or collection of algebraic equations). Alternately, competence can be factored across an analytic decomposition of problem-solving into a set of related processes, as in studies of problem categorization or comparison (e.g., Reed, 1987). Ecological approaches instead hold competence accountable to the local demands of the setting in which the problem and its solution are enacted. As with Scribner's (1984) dairy workers, the issue may not be simply whether or not one has gotten the correct quantity, but whether the activities required to find that quantity allow the person to remain engaged in an ongoing line of work. Thus, high school students may be prodigious calculators, but their relative inflexibility in comparison with icebox workers makes them "novices" in the dairy setting. That theoretical ascriptions about competence have consequences can be seen in the discontinuity between relatively principled quantitative reasoning among preschoolers and the progressive inversion between manipulative skills and conceptual understanding in later school mathematics (Resnick, 1987a). The studies reported in the following chapters contrast people with different levels of competence as conventionally ascribed in school settings, but analyses of their activities when solving algebra story problems bring forward several surprising commonalities.

Perhaps the deepest schism between cognitive and ecological approaches is over the issue of what constitutes "knowledge," its location, and its relation to activity. The cognitive perspective proposes mental structures as the primary basis for coherent individual activity. The boundary between an individual's knowledge and the material or social setting is permeable through processes that encode perceived aspects of relevant external inputs. Thus, Paige and Simon's (1966) intriguing observations about the material structure of diagrams for making quantitative inferences are attributed to "functionally equivalent" mental representations of the diagrams. In contrast, the ecological perspective admits more to the material and social setting when analyzing the relation between knowledge and activity. For example, the "knowledge" involved in ship navigation appears to be distributed across people and "cognitive artifacts" like the nomogram or "3-minute rule," both of which are required to manage a competent navigational performance (Norman and Hutchins, 1988). The role of material structure in "models" of functional relationships is treated in a similar fashion in the following representational conventions were adopted or how these fit together to give a coherent understanding of any individual problem. Third, these introspective efforts gave little insight into problem–solving attempts where routine classification broke down and the textbook problem effectively came "alive" as a genuine problem.

On the basis of pilot studies collected with undergraduate and graduate students, the study of competent problem solvers described in this chapter was undertaken. An analysis of written solution attempts on representative algebra story problems taken with undergraduates in nontechnical majors suggested that a wide variety of problem–solving activities were commonplace, including various forms of nonalgebraic representation (i.e., diagrams, tables, and written natural language arguments). Surprisingly, similar activities were observed in a small group of graduate–level computer science students, suggesting that a detailed account of competence in the task domain of algebra story problems might diverge widely from existing accounts in the empirical literature. The study of written protocols described in this chapter provides one line of evidence on the episodic structure of solution activities observed in a group of competent mathematical problem solvers.

3.2 Method

The primary goal of this study is to characterize the activities of "competent" problem solvers on representative algebra story problems. When compared with the activities of beginning algebra students, the contrast should give a rough image of the terrain over which a learner must travel to become a skilled problem solver. Participants were chosen who had clearly mastered the algebra curriculum up to existing institutional standards, but who were not recent recipients of algebra–based instruction. Thus, findings describe a primary target of traditional instruction in algebra: a problem solver who has mastered the tools of the algebraic formalism, has practiced these skills during instruction, and should be able to apply these skills in novel settings. The study involves minimal experimental intervention, so interpretation and analysis of problem–solving protocols are primarily descriptive.

3.2.1 Participants

Participants in this study were 85 undergraduate computer science majors in their junior and senior years. They were enrolled in an introductory course in artificial intelligence and participated in the study as part of their classroom activities. They could be viewed as "experts" in algebra story problem solving, because they must have successfully completed courses in algebra during secondary schooling. In addition, prerequisites to the artificial intelligence course included three university–level courses in calculus and completion or concurrent enrollment in courses covering discrete mathematics. As a result, the level of mathematical sophistication in this sample of problem solvers should be high — i.e., they are institutionally certified problem solvers. Still, as shown in later analyses, the solutions offered by many members of this sample do not resemble the smooth execution of a practiced skill.

3.2.2 Materials

Participants were asked to solve the four algebra story problems shown in Table 1.1. Problems MOD, MRT and WT were taken directly from Mayer's (1981) sample of algebra story problems, with minor alterations in their number set and phrasing. These alterations were intended to free participants from unwieldy calculations during problem solving and to make wording between selected pairs of problems more similar. Problem WC was constructed to be isomorphic to the MRT problem at the level of quantitative structure.

These problems were selected for two reasons. First, with the possible exception of WC, they are typical of problems found in secondary school texts. From an exhaustive set of 1,097 algebra story problems drawn from 10 texts, Mayer found that problems like MOD, MRT, and WT accounted for 7.8% of all observed problems. Second, different pairings of these problems present participants with opportunities for positive or negative transfer across contiguously presented problems.

Aside from their use as representative problem–solving tasks, algebra story problems often serve as materials for studies of analogical transfer.[1] Given a target problem to solve, participants exhibit positive transfer when they can use the solution method from a previously encountered source problem to help solve the target problem. Alternately, participants exhibit negative transfer when they access and use the solution from an inappropriately related source problem.

Specifically, problem pairings MOD–WT and MRT–WC are isomorphic in their quantitative structure (see Figure 2.2) and have similar situational contexts. In the MOD–WT pair (see Figure 2.6), output segments are adjacent, being collinear and sharing a starting point, while time segments are congruent, overlapping completely by sharing both starting and ending times. In the MRT–WC pair, outputs are congruent while times are adjacent and of different value. Should subjects recognize this similarity, they may exhibit some form of positive transfer. Alternately, problem pairings MOD–MRT and WT–WC are similar at a more superficial level, sharing types of surface materials (e.g., distance traveled or parts of a job completed) while having quite dissimilar quantitative and situational structures. In fact, relations over output and time dimensions are exactly reversed, as described in the preceding section on quantitative structure. In the MOD–MRT pair for example, outputs in MOD are

[1]See Hall (1986, 1989a) for comparative reviews of analogical inference in problem solving and learning.

manipulative errors, and the status of the episode in the overall solution attempt. The last category covers relative correctness and the reason for transition to a new episode. With the exception of conceptual content, each category is further differentiated into alternative subcategories, as shown in Table 3.1. In some cases only one subcategory is selected as best describing the more general category (e.g., simulation as a type of model-based reasoning under tactical content); in other cases, each subcategory can occur within a single episode (e.g., various kinds of conceptual and manipulative errors).

The remainder of this section takes up each of these interpretive categories in detail, showing representative written protocols as examples of their use in scoring the episodic structure of participants' solution attempts. For example, student M20 in Figure 3.1 goes through three error-free episodes, each with a specific purpose, tactic, content, and transition. In the protocols shown in figures as illustrations of various categories, episodes are separated by dashed lines, and their sequence is shown with circled numbers. Several protocol excerpts are presented directly in the text without accompanying figures.

3.3.1 Strategic purpose

The strategic purpose of an episode describes its relation to the goal of finding a solution. Judgments of a problem solver's "purpose" are clearly a matter of our own interpretation, although scoring criteria are presented that make these judgments operational across individual ratings. In this regard, our scoring distinguishes among three abstract problem-solving modes.

Comprehension. The participant is not directly seeking a final solution, but is constructing a representation of the problem by incorporating various constraints. In Episode 1 of Figure 3.1, the participant finds a way to express working rates by considering their outputs after one hour.

Solution attempt. The participant is attempting a series of operations that work directly toward a solution (Figure 3.1, Episode 2).

Verification. The participant has already produced a solution to the problem and is now seeking confirmatory evidence for it, for instance by rederiving the solution with another tactic or by inserting the answer in some intermediate equations (Figure 3.1, Episode 3).

3.3.2 Tactical content

The tactical content of an episode is the method a participant uses to achieve some strategic purpose. Operational criteria refer primarily to the protocol material for the current episode, but in a few cases information contained directly in the protocol was

Table 3.1: Categories for interpreting the content of problem–solving episodes.

I. Strategic purpose	III. Conceptual content
Comprehension	
Solution attempt	IV. Errors
Verification	Conceptual errors
	Errors of commission
II. Tactical content	Errors of omission
Annotation	Manipulation errors
Problem elements	Algebraic errors
Retrieval of formulas	Variable errors
Diagram	Arithmetic errors
Algebra	
Model–based reasoning	V. Status of episode in solution attempt
Simulation	Consistency
Heuristic	Before
Ratio	During
Whole/part	After
Part/whole, part/part	Transition
Proportion	Subgoal
Scaling	Wrong
Unit	Impasse
Procedure	Lost
	Final solution
	Found solution wrong

Mary can do a job in 5 hours and Jane can do the job in 4 hours. If they work together, how long will it take to do the job?

Mary does 1/5 job in 1 hr
Jane " 1/4 job in 1 hr (1)

$\frac{1}{5}x + \frac{1}{4}x = 1$

$x\left(\frac{1}{5} + \frac{1}{4}\right) = 1$

$x\left(\frac{4}{20} + \frac{5}{20}\right) = 1$ (2)

$x\left(\frac{9}{20}\right) = 1$

$\boxed{x = \frac{20}{9}}$

DOUBLE CHECK:

$\frac{1}{5}\left(\frac{20}{9}\right) + \left(\frac{20}{9}\right)\frac{1}{4} = 1$ (3)

$\frac{4}{9} + \frac{5}{9} = 1$

okay.

Figure 3.1: Protocol of student M20 on the WT problem.

insufficient to make an operational category judgment. In these cases, surrounding episodes and post hoc written explanations supplied by the participant were used to assist scoring.

Annotation. These episodes usually occur early in the protocol when participants are collecting information about the problem. Three cases are covered.

- *Problem elements.* The participant is recording elements of the problem text (e.g., $V_A = 60km/hr$, Figure 3.2, Episode 2).
- *Retrieval of formulas.* The participant is remembering and writing down memorized formulas that seem relevant (e.g., $v = \frac{d}{t}$, Figure 3.2, Episode 4).
- *Diagram.* The participant draws a picture of the problem situation (e.g., Figure 3.2, Episode 1).

Algebra. An episode is algebraic if it makes use of one or more equations placing constraints on the value of one or more variables. However, simple assignments are not treated as equations. Thus neither $100 + 60 = 160$ nor $d = 880$ is considered an equation, whereas $d = 100 \times t$ is considered an equation. As shown unusually clearly in the protocol of Figure 3.3, the tactical approach of the typical algebraist is to express constraints as a system of one or more equations (or proportions) and to solve for the appropriate unknown. There are also cases of participants trying equations in a generate–and–test fashion until, as one explained, an equation "looks good."

Model–based reasoning. This category is scored when a participant "executes" a model of the problem situation along the dimension defined by an unknown quantity such as time, distance, or work. The two subcategories of model–based reasoning relate to constructive problem–solving inferences described in the preceding chapter in the section on situational structure.

- *Simulation.*[2] The participant selects a base unit for the chosen dimension and "runs" the model for each successive unit increment as illustrated in Episode 3 of Figure 3.2. Consistent with earlier descriptions of situational structure, a simulation episode could be interpreted as an iterative "joining" of concrete individual inclines. Simulation can also be partial (just one or two increments) in that it is not used to reach a solution, but to examine relations between quantities and to enable some other solution method. In both Episode 1 of Figure 3.1 and Episode 5 of Figure 3.7, a simulation for one hour establishes the quantitative combination of entities from distinct events.
- *Heuristic.* The base quantity "jumps" by variable increments whose magnitude is determined at each point by estimations of closeness to the solution. A heuristic

[2]This sense of *simulation* differs from computational studies of common–sense reasoning. For example, de Kleer's (1977, 1979) *envisionment* uses quantitative calculation to resolve qualitative ambiguity, whereas the current sense of simulation uses a qualitative model to help disambiguate quantitative constraints.

Two trains leave the same station at the same time. They travel in opposite directions. One train travels 60 km/h and the other 100 km/h. In how many hours will they be 880 km apart?

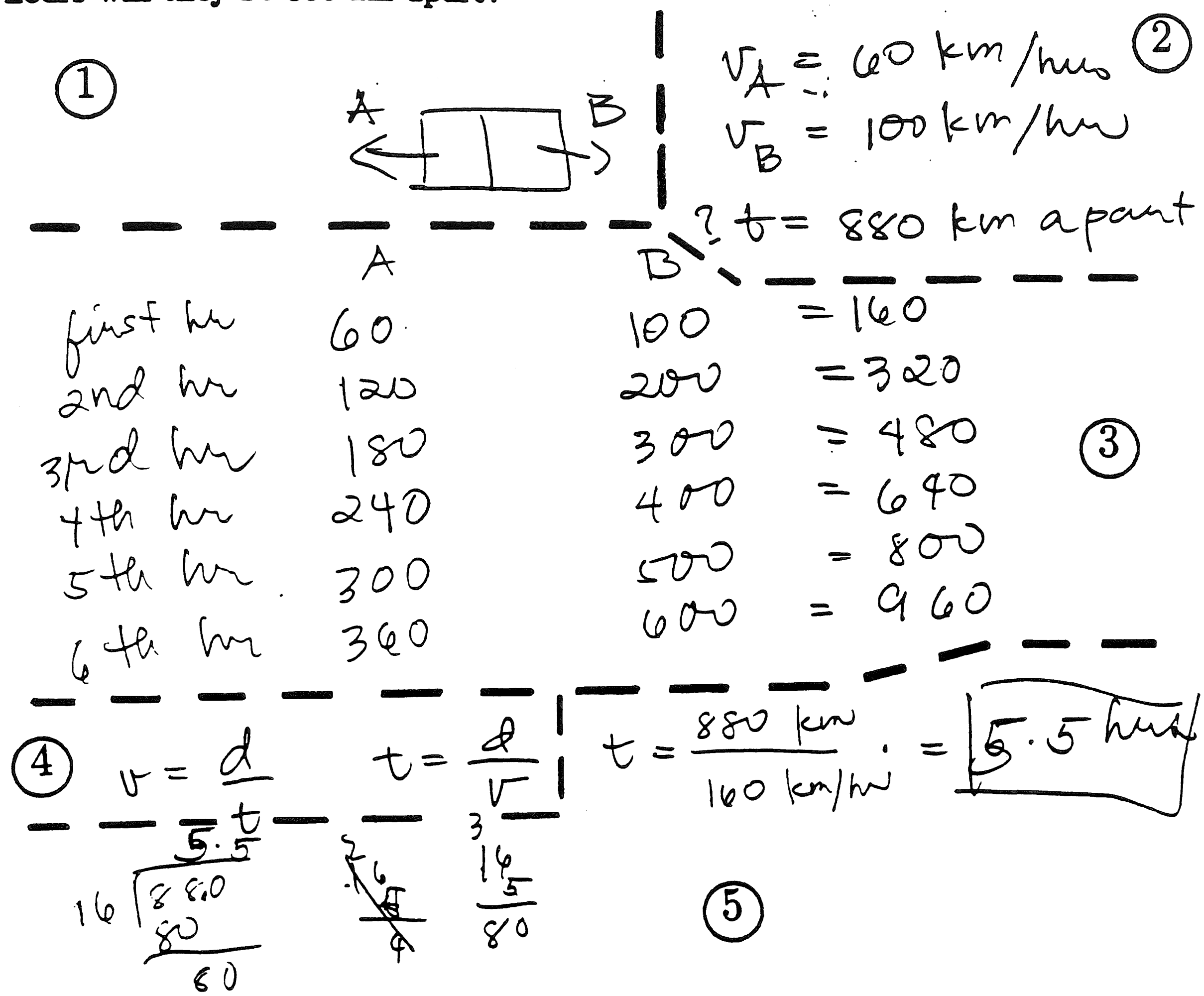


Figure 3.2: Protocol of student W06 on the MOD problem.

George rode out of town on the bus at an average speed of 24 miles per hour and walked back at an average speed of 3 miles per hour. How far did he go if he was gone for six hours?

bus distance = (24 miles/hr)(x hours)

walking distance = (3 miles/hr)(6 − x hours)

bus distance = walking distance

(24 miles/hr)(X hours) = (3 miles/hr)(6 − X hours)

$$24x = 18 - 3x$$

$$27x = 18$$

① $$X = \frac{18}{27} = \frac{2}{3} \text{ hours}$$

bus distance = (24 miles/hr)($\frac{2}{3}$ hours)

bus distance = 16 miles = walking distance

One way = 16 miles

Round Trip = 32 miles

Figure 3.3: Protocol of student M39 on the MRT problem.

Two trains leave the same station at the same time. They travel in opposite directions. One train travels 60 km/h and the other 100 km/h. In how many hours will they be 880 km apart?

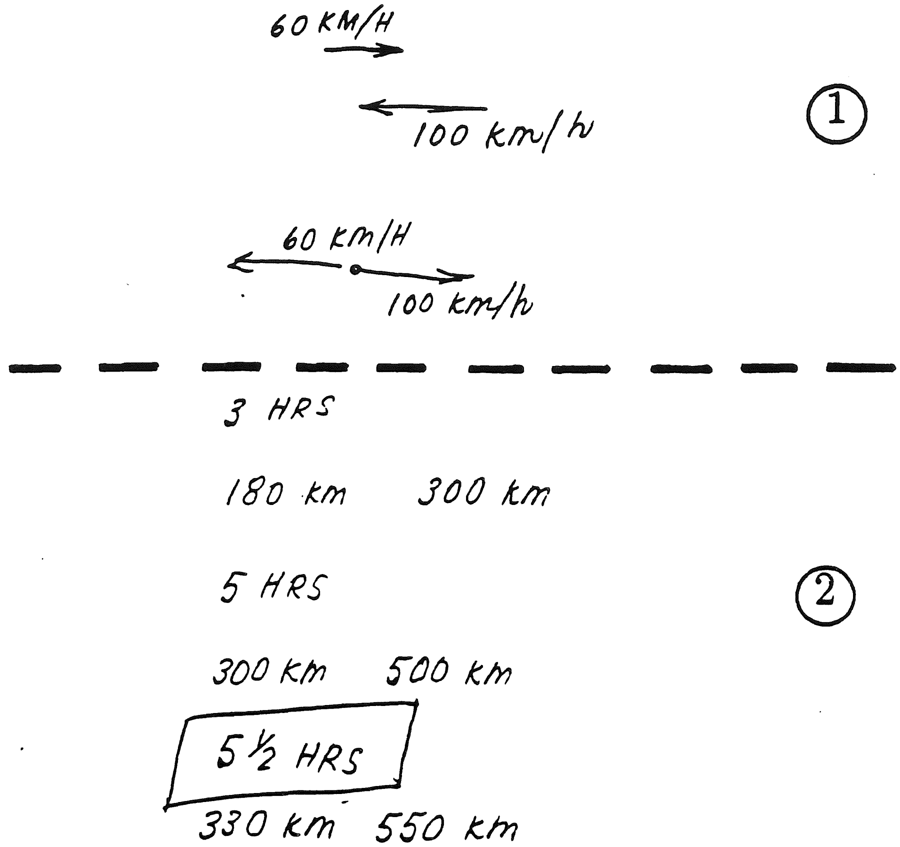


Figure 3.4: Protocol of student M03 on the MOD problem.

model–based reasoning episode could be interpreted as "scaling" inclines that represent invariant relations, as described earlier. The progression of this generate–and–test approach can be monotonic, as in Episode 2 of Figure 3.4, or it can follow some form of interpolation search. After each generation of a base value, the state of the problem situation being modeled is reconstructed and evaluated.

Ratio. This subcategory covers tactics in which relations of proportionality between quantities are used, sometimes providing clever "shortcuts" to a solution. These tactics clearly utilize a representation of quantity (e.g., intensive quantities, as described earlier), but the manner in which related quantities are integrated may depend on constructive inferences within the situational context (e.g., composing segments or inclines).

Mary can do a job in 5 hours and Jane can do the job in 4 hours. If they work together, how long will it take to do the job?

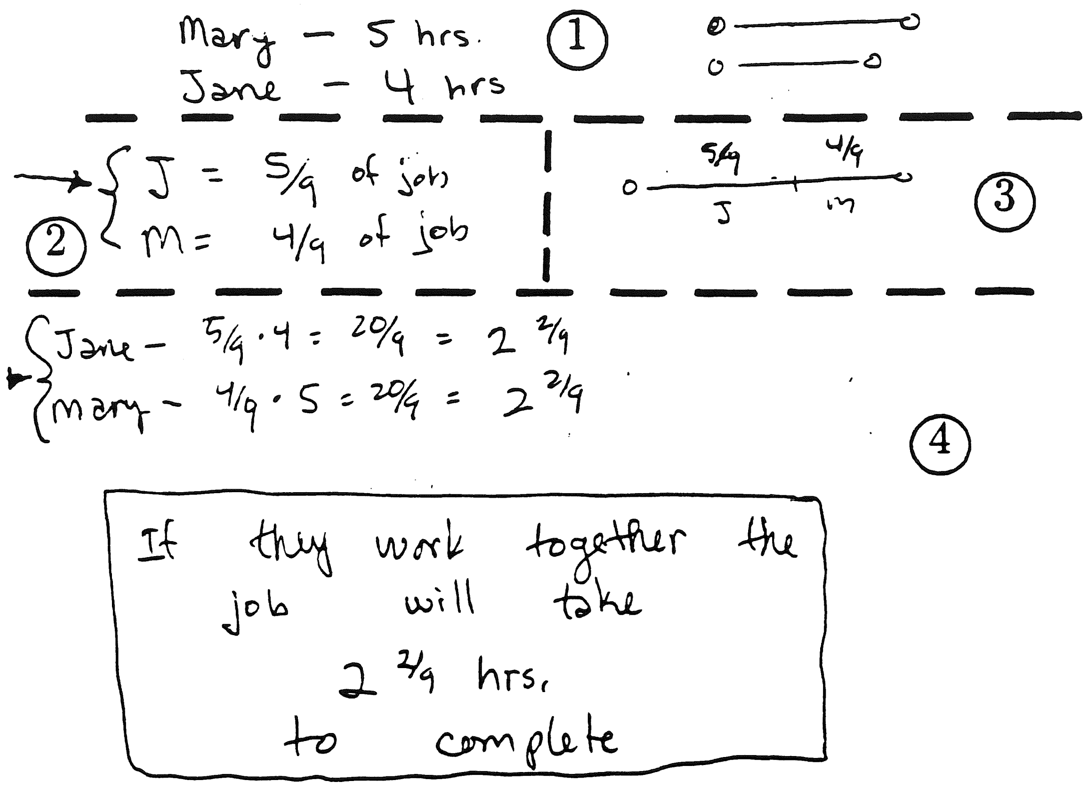


Figure 3.5: Protocol of student M32 on the WT problem.

- *Whole/part.* The participant views a part as fitting some number of times into a whole quantity, as in Episode 6 of Figure 3.7 where the student divides a given distance apart by an incorrect rate of separation.
- *Part/whole and part/part.* These two types of ratios compare portions of entities. Use of the part/whole ratio is illustrated in Episodes 2 to 4 of Figure 3.5, where the participant considers parts of the total job. A version of the part/part ratio appears in Episode 2 of Figure 3.6, involving the respective rates of bus and foot travel.
- *Proportion.* Nonalgebraic proportions cover reasoning of the type exhibited by student m05 on the WT problem: "they've done $\frac{9}{10}$ [of a job] in 2 hrs, so $\frac{2}{9}$ hr more would do for [the job] left to be done."
- *Scaling.* The participant solves a related version of the problem or reaches an unexpected answer and simply scales the answer to fit the quantities given in the problem. This may relate to our earlier description of scaling rates as invariant two–dimensional inclines. In Episodes 3 and 4 of Figure 3.6, for example, the participant solved an easier problem by heuristic model–based reasoning and then scaled her answer to "fit" the MRT problem.

Unit. In a few cases, a participant reasons purely in terms of units of measurement given in the problem. For instance, on the work competitive problem (WC), student M44 examines alternative rate forms with the following manipulations:
"$\frac{box}{min(utes)} \cdot min = box \qquad \frac{min}{box} \cdot box = min.$"

Procedure. This subcategory is scored when there is evidence that a participant is executing some stored sequence of actions or operations other than routine algebraic or arithmetic manipulation. For example, on the work together problem (WT), student M21 appears to use a simple averaging tactic for combining quantities, writing "$total = \frac{1}{2}(5+4) = \frac{9}{2} = 4\frac{1}{2} hrs.$"

3.3.3 Conceptual content

The conceptual content of an episode reflects the participant's conceptualization of the problem situation and the resulting set of constraints between problem entities. There is a subtle but crucial distinction between situational understanding and the quantitative constraints that are implied by it, as suggested in previous sections. The scoring of conceptual content simply includes constraints that the participant clearly recognizes and uses in the episode. For instance, participant M39 in Figure 3.3 appears to understand all of the necessary constraints in problem MRT: equal distances, additive composition of times, and the distance–rate–time relation.

In order to track correct or incorrect structural inferences across episodes, a normative account of the constraints contained in a problem must be imposed. Naturally, this account of a problem's structure may be at odds with quite reasonable

George rode out of town on the bus at an average speed of 24 miles per hour and walked back at an average speed of 3 miles per hour. How far did he go if he was gone for six hours?

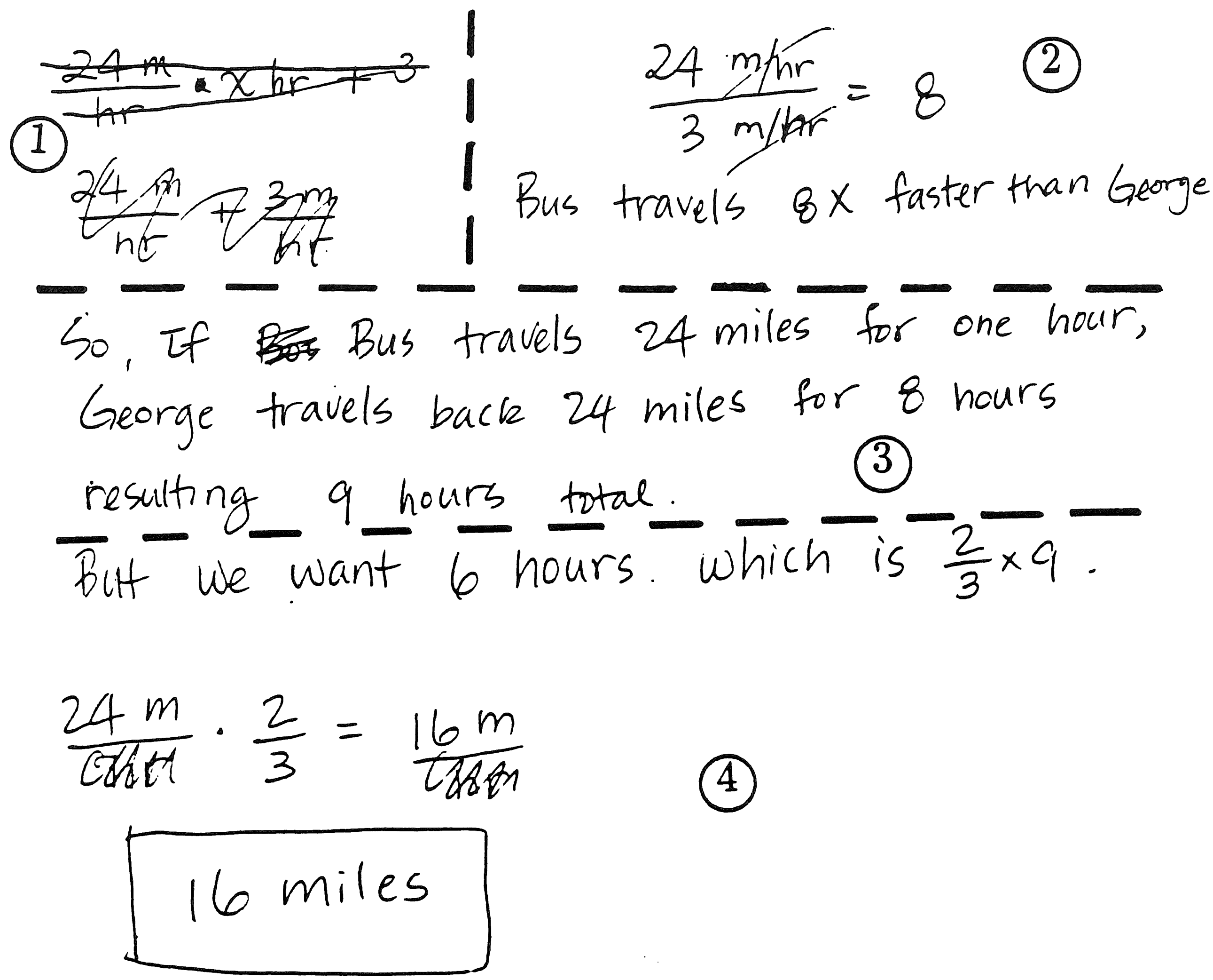


Figure 3.6: Protocol of student W17 on the MRT problem.

conceptualizations that participants construct, and the normative interpretation used here attempts to remain open to coherent alternative viewpoints on problem structure. Returning to the earlier analysis of problem structure, the components of a problem can be divided into one and two dimensional entities, with specific relations possible between these entities. Comparing the structural relations between these components that are evident in a participant's written protocol with a normative view of problem structure, correct structural inferences or conceptual errors are identified in each of the following categories:

- relations over output or time entities (e.g., distances are the same),
- the form in which the rate is expressed (e.g., as output over time),
- relations between rates (e.g., addition or ratio), and
- complex expressions relating multiple entities.

Correct and incorrect relations between one dimensional entities (i.e., output or time) vary over problems as described earlier (e.g., congruent times have equal values in problem MOD). Rate form as a relation over one dimensional entities and relations between rates present more difficult interpretations, since status as correct or incorrect must be determined by considering the way in which the participant uses rates to organize one dimensional entities. For example, in problem WC student M13 forms a rate as time–per–box (5 and 2 minutes per box), composes these rates by adding (7 minutes per box), and then uses the composite rate as a divisor of the total time to find that eight boxes are filled and checked in 56 minutes. This is a reasonable and effective conception of problem structure, but the rate form reverses a normative view of working as box–per–time and leads to a very different algebraic construction (e.g., $1/5\ T_f = 1/2\ T_c$).

The complex expression category covers relational constraints over several entities that could not be unambiguously broken into more specific constraint categories during protocol scoring. For example, on problem MOD student W01 writes "x/60 + x/100 = x/880," a complex expression that cannot lead to a correct value for an unknown time (assuming x is a variable for a common time), but it is difficult to tell whether 60 and 100 are being treated as rates or partial distances in some kind of algebraic proportion. The policy taken in this case is to treat the entire expression as a conceptual error, categorized as a complex expression.

After coding these categories as absent or containing a particular relational constraint for conceptual content, errors of omission, and errors of commission on each problem–solving episode, we can track the introduction of correct and incorrect structural inferences can be tracked across successive episodes within each participant's solution attempt on each problem. Introducing operational accounts of the origin of correct or incorrect constraints enables finding which tactics are responsible for inferences about problem structure and how incorrect structural inferences might be "repaired" during subsequent episodes.

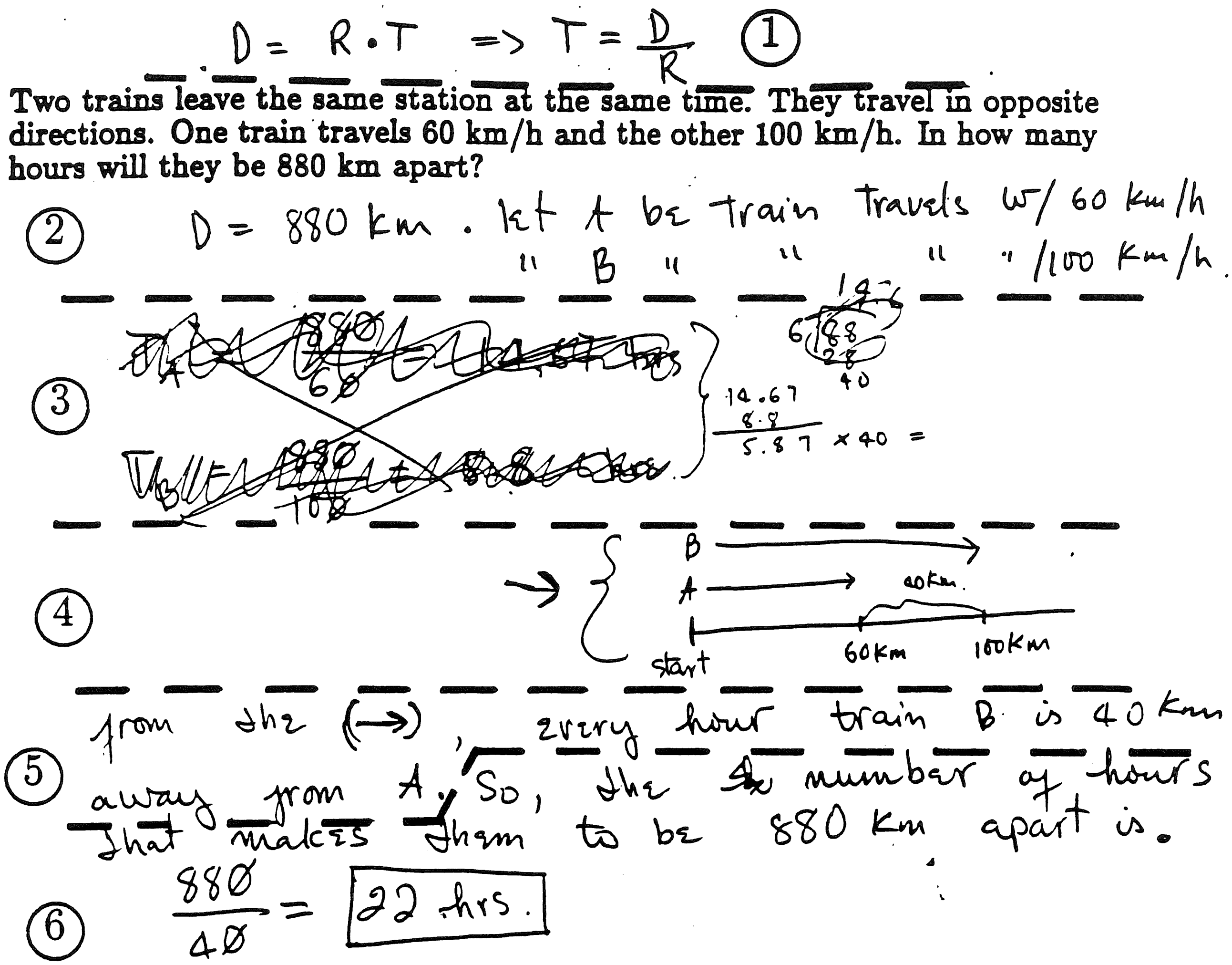


Figure 3.7: Protocol of student M19 on the MOD problem.

3.3.4 Errors

Within each problem–solving episode, two broad classes of errors are considered. **Conceptual errors.** These are scored when a participant either includes a constraint that is inappropriate for the problem or excludes a constraint that is a critical requirement for the current episode.

- *Errors of commission.* These errors are incorrect constraints that the participant introduces during an episode, either by incorrectly representing the situational context of the problem or by making erroneous quantitative inferences. For example, in Episodes 4 to 6 of Figure 3.7, the participant subtracts distances thinking the trains are going in the same direction.
- *Errors of omission.* These errors are overlooked constraints. To be scored as an error of omission, an overlooked constraint has to be critical to the solution method applied by the participant. This usually means that two entities are explicitly used, whereas the relation between them is ignored. In Episode 3 of Figure 3.8, the participant overlooks that working times represented as x and y are equal.

Manipulation errors. Because written protocols usually display algebraic or arithmetic manipulations clearly, scoring identifies manipulative errors of three types.

- *Algebraic errors.* For example, on the MOD problem, student W39 wrote "$880 = \frac{160}{T}$" followed by "$T = \frac{880}{160}$."
- *Variable errors.* Two types of errors related to the concept of variable are scored. In switch errors, the meaning of a variable changes in the course of problem solving. In label errors, participants use variables as labels for quantities. For instance, in the round trip problem (MRT), student M10 writes the equation "$1B + 8W = 6hrs$" and explains that "for every 1 hour on the bus, it takes 8 hours to get back."
- *Arithmetic errors.* For example, on the opposite direction motion problem (MOD), student M20 wrote "$\frac{880}{160} = \frac{11}{4}$." After detecting this arithmetic error in a verification episode, they recovered by using the ratio scaling tactic mentioned earlier.

3.3.5 Status of episode within solution attempt

Categories listed so far deal with internal characteristics of an episode. The final categories of the scoring scheme, consistency and transition, describe the relation of an individual episode to the overall problem–solving effort.

Consistency. This category assesses the correctness of an episode in the context of the surrounding problem–solving sequence and is scored correct or incorrect for three facets.

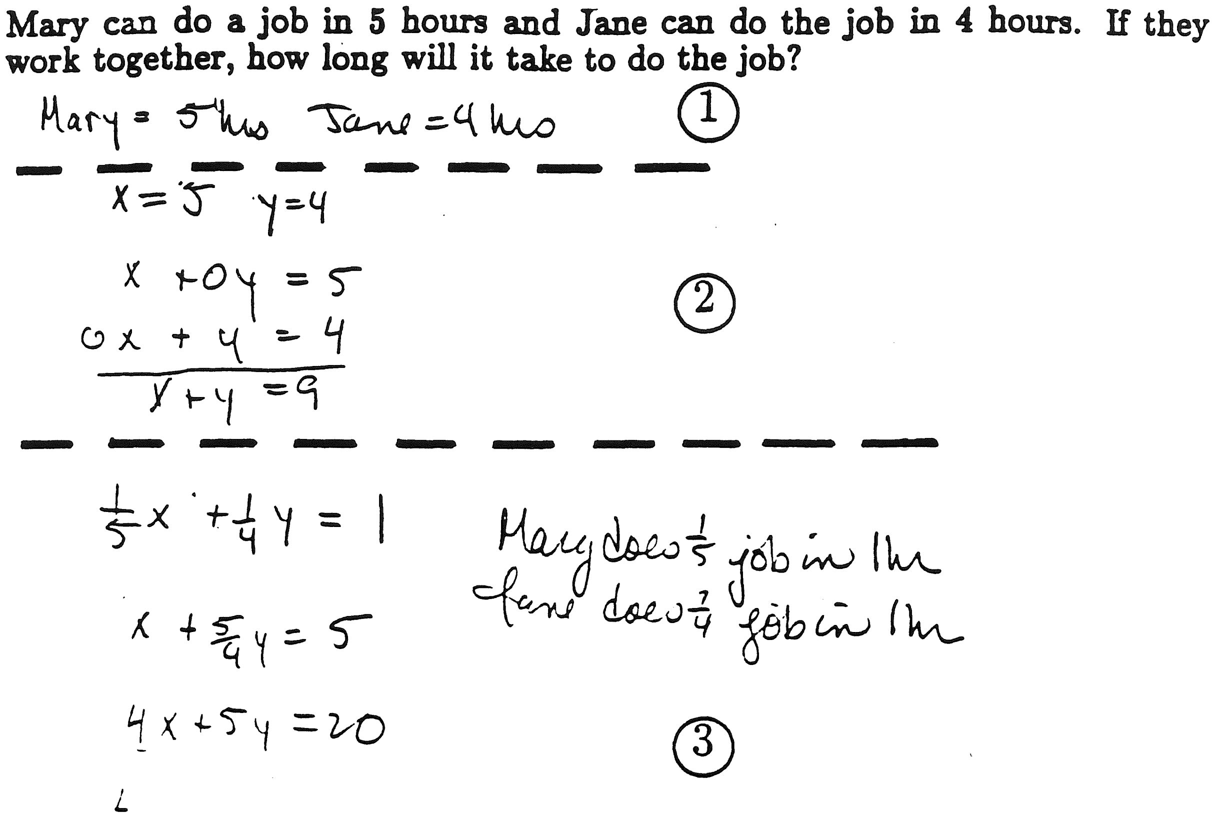


Figure 3.8: Protocol of student W23 on the WT problem.

- *Before.* This subcategory reflects the correctness of the context inherited by the episode. For example, errors may be generated in former episodes and passed into the current episode, as with the conceptual error of commission (same direction) passed between Episodes 4 and 5 of Figure 3.7.
- *During.* This scores the correctness of the current episode with respect to the inherited context. An episode producing an incorrect result can be internally correct if it is consistent with an incorrect context. For example, Episodes 5 and 6 of Figure 3.7 are internally consistent with the conceptual error of commission introduced in Episode 4.
- *After.* This subcategory assesses the absolute correctness of the outcome of the current episode. If a solution is presented, the scoring reflects its correctness; otherwise, scoring assesses whether or not the participant is on a possible right track in their solution attempt.

Transition. The intent of scoring transition is to determine the reason a participant passes from one episode to the next. Unlike consistency, which reflects the scorers' judgment of correctness, this aspect attempts to capture the participant's point of view.

- *Subgoal.* The participant accomplishes an intermediate goal, bringing the episode to an end (Figure 3.1, Episodes 1 and 3). Information identified when achieving a subgoal (e.g., changing the form of a working rate) is generally carried into subsequent episodes.
- *Wrong.* The participant decides that she is on the wrong track and abandons the current approach, usually by marking through the work (Figure 3.7, Episode 3). This transition is often the result of an explicit verification episode.
- *Impasse.* The participant reaches a point where she cannot continue with the current method. A good example of impasse is shown in Episode 3 of Figure 3.2, where the student correctly applies simulation by hourly increments, overshoots the noninteger solution, and then switches to an algebraic tactic after adding rates.
- *Lost.* The participant reaches a point where she cannot determine how to proceed, as in Episode 2 of Figure 3.8.
- *Final solution.* The participant reaches a result and presents it as a solution to the problem.
- *Found solution wrong.* The participant realizes or believes that the solution presented is incorrect.

This presentation of a framework for interpreting written protocols gives an overly linear picture of its use in scoring solution attempts. In fact, categorizing the episodic structure of a written protocol within this framework was usually done quickly (from 5 to 20 minutes per protocol) and with little subsequent disagreement among the scorers. By design, each category was rated with at least 75% agreement over four scorers; most categories approached unanimous agreement.

In addition to determining whether or not a participant has managed to find a correct solution to an algebra story problem, this framework for interpreting problem–solving episodes allows us to describe the internal structure of their solution attempt. Analysis of episodic structure supports more fine–grained explorations of the strategic and tactical course of problem solving. In the quantitative results section which follows, composite analytic categories are formed by identifying episodic patterns among the atomic category judgments already described. These analyses can speak of participants reaching a "final episode" with some particular tactic and content or can examine a series of contiguous episodes during which model–based reasoning is used.

3.4 Quantitative Analysis of Problem–Solving Episodes

Developed as a working hypothesis in Chapter 2, competent problem solving proceeds as an elaborative, interdependent exploration of two distinct problem spaces: (a) the situational context of a story problem and (b) the quantitative constraints given explicitly or implicitly in the problem statement. Results presented in this section provide evidence for this interdependency at a global level of problem–solving activity and at a more detailed level of episodic content. The analysis distinguishes between participants' ***problem–solving attempts*** and the *episodic structure* of those attempts. A problem–solving attempt, includes all the activities evident in the written protocol, which may include several distinct episodes. Episodic structure describes the alternation of problem–solving episodes of various types, and the constraints or errors that are contained within and across those episodes.

The tactical content, strategic purpose, transitional status, and errors present in participants' solution attempts are analyzed first. These analyses pool episodes within solution attempts to show the prevalence of different interpretive categories, and so they provide only a coarse image of competent problem solving. The second level of analysis looks within individual solution attempts and examines two episodic patterns in detail. An analysis of the episode during which a final solution is offered provides a finer image of problem–solving outcome, describing relations between solution outcomes and other interpretive categories within the episode. This level of analysis also identifies individual episodes of model–based reasoning to permit a closer examination of problem–solving activity outside the traditional algebraic formalism. By considering the content of surrounding problem–solving episodes, participants' reasons for using model–based reasoning can be examined, and the effectiveness of this tactic for making correct problem–solving inferences or recovering from existing errors can be assessed. Third, we pool episodes across participants to look at the origin of inferences about problem structure, asking which problem–solving tactics are "responsible" for the introduction of correct structural inferences versus conceptual errors of omission or commission. These analyses also examine problem–solving tactics for those episodes in which prior conceptual errors are "repaired," either by being eliminated or replaced with a correct structural constraint.

3.4.1 Problem–solving attempts

Because interpretive categories advance hypotheses about problem–solving processes, their frequency of occurrence among participants is presented. Table 3.2 shows the percentage of participants having one or more episodes in which various rated categories were observed. Percentages are shown separately for each problem (MOD, MRT, WT, WC) but are collapsed over groups (M, W) since none of these contrasts were statistically reliable. Most findings are as expected, while several are surprising.

Tactical content of scored episodes. Although most participants use algebra in their solution attempts (63.5% to 85.9% across problems), reasoning outside the algebraic formalism is surprisingly common.

Table 3.2: Percentage of subjects with a scored category during their solution attempts.

	Problem			
Category	MOD	MRT	WT	WC
Tactical content				
Algebra	82.4	85.9	71.8	63.5
Model	30.6	22.4	35.3	47.1
Ratio	17.6	14.1	15.3	42.4
Procedure	0.0	1.2	21.2	0.0
Units	3.5	1.2	1.2	1.2
Notations	7.1	15.3	21.2	29.4
Diagram	69.4	36.5	8.2	9.4
Strategic purpose				
Comprehension	84.7	64.7	57.6	60.0
Solution attempt	100.0	100.0	100.0	100.0
Verification	28.2	20.0	7.1	20.0
Episode transitions				
Solution	97.6	75.3	85.9	97.6
Impasse	9.4	10.6	7.1	4.7
Lost	4.7	21.2	15.3	3.5
Wrong	16.5	38.8	25.9	16.5
Errors				
Omission	7.1	21.2	23.5	11.8
Commission	17.6	49.4	42.4	14.1
Arithmetic	9.4	4.7	3.5	2.4
Algebra	5.9	8.2	8.2	0.0
Variable	1.2	5.9	14.1	2.4

- Looking within individual problems, at least one model–based episode is used by 22.4% to 47.1% of participants, depending on the problem. A separate analysis pooling across problems shows that 72.9% of participants have one or more episodes of model–based reasoning in their written protocols. These episodes are explored more fully later.
- Use of ratios is the next most prevalent nonalgebraic tactic (14.1% to 42.4% across problems) and may depend on a variety of factors: (a) the complexity of the constraints presented by a problem's quantitative structure, (b) the accessibility of situational justifications for those constraints, and (c) the manner in which the constraints are presented in the problem text.
- Few solution attempts contain episodes using a "procedure" or reasoning with "units." Most participants using a procedure on problem WT chose to take an average over working rates, a strategy that violated the situational meaning of "working together" in that problem and generally led to an incorrect solution.
- Annotations, in the form of diagrams or notations about problem elements, were either scarce or common, depending on the situational and surface content of the story problem. Motion problems (MOD, MRT) showed few notations for problem elements (7.1%, 15.3%) but more frequent diagrams (69.4%, 36.5%), whereas work problems showed frequent notations (21.2%, 29.4%) but fewer diagrams (8.2%, 9.4%). Although it is likely that the spatial content of motion problems makes them more suitable for diagrams, some participants are able to construct effective diagrams for work problems (e.g., see Figure 3.5, Episode 3).

Strategic purpose of scored episodes. Most participants show explicit attempts at comprehension in their written protocols (57.6% to 84.7% across problems), typically through diagrams, notations about problem elements or model–based reasoning. Although all participants make some attempt to solve the problem, only a minority give evidence of attempting to verify the results of their work (7.1% to 28.2% across problems).

Transitions out of scored episodes. Most participants find and explicitly present a solution (either correct or incorrect) as part of their problem–solving attempt, although problems MRT and WT appear more difficult than their quantitative isomorphs in this regard (WC and MOD). A more detailed analysis of solution outcomes follows shortly. Likewise, the three transitions without solution (i.e., impasse, lost, or wrong) are most common in the more difficult problems (MRT and WT).

Errors in scored episodes. Conceptual errors of omission and commission increase for the more difficult problems (MRT and WT), and appear much more frequently than manipulative errors (arithmetic, algebraic, or variable errors) on all problems.

Several interesting patterns emerge in these findings. First, participants' written protocols are not composed solely of material generated while performing algebraic transformations. Instead, many use various forms of model–based reasoning, conducted within their understanding of the context posed by a story problem text. Second, although most participants do present a solution in some form, their efforts do not appear as a smooth progression toward a quantitative solution. Rather, their problem–solving efforts are often interrupted by varied conceptual difficulties that must be repaired before a solution is found. Third, manipulation errors within algebraic and arithmetic formalisms *do* occur, but these are overshadowed by conceptual errors of omission or commission as a primary source of problem–solving difficulty. Consistent with the treatment of problem structure in Chapter 2, these findings are interpreted to mean that participants form an understanding of the problem at the level of its situational context and then use this understanding to introduce quantitative constraints. As a result, *many of the activities present in an episodic analysis of algebra story problem solving fall outside the traditional algebraic formalism.*

3.4.2 Final episodes: outcome, tactical content, and errors

Examination of the written protocols clearly shows that participants undertake a variety of problem–solving activities when attempting to solve these problems, particularly when they encounter difficulties in reaching a solution. However, the previous findings speak only to the *presence* of various conditions in participant's problem–solving efforts. As scored, participants average approximately 2.5 episodes per problem–solving effort, with some protocols presenting evidence for as many as 10 distinct episodes. The following analyses look within individual protocols for more finely detailed episodic structure.

Within an individual's efforts on any given problem, a final episode is extracted for a first level of detailed analysis. This episode need not be their last effort in a solution attempt, but it is final in one of three senses: it is the last episode during which they present a solution that is correct, the last episode during which they present a solution that is incorrect, or the last episode of a problem–solving effort in which no solution is presented. Incorrect means the participant presents an incorrect final solution without detecting any errors. The no solution category includes participants who present an incorrect solution but realize they have done so during a subsequent attempt at verification, without being able to recover. Thus, the final episode may be correct, incorrect, or present no solution.

Performance outcomes and problem order. Table 3.3 shows the final outcomes for each problem, broken down to show anticipated effects of problem ordering. On some target problem, participants are expected to perform better if they have just seen a relevant source problem than if they have seen no prior problem or have just seen an irrelevant source problem. For example, if participants recognize and use an analogy between adjacent problems, group W should perform better than group M on problem MOD (shown as M < W in the table), since those in group W are exposed to a relevant problem (WT) just before seeing problem MOD. Problem WT is a relevant source, as described in the Methods section, since it is isomorphic in quantitative

Table 3.3: Final Episodes: Percentage Correct by Subject Groupings.

	Problem by Group Contrast†							
	MOD		MRT		WT		WC	
Outcome	M < W		M > W		M > W		M < W	
Correct	89.1	92.3	47.8	56.4	58.7	64.1	93.5	89.7
Incorrect	6.5	7.7	19.6	15.4	28.3	20.5	6.5	5.1
No-solution	4.3	0.0	32.6	28.2	13.0	15.4	0.0	5.1

†Group contrasts (M vs. W) anticipate effects of problem order.
M sees MOD, WT, WC, MRT; W sees WT, MOD, MRT, WC.

structure to problem MOD. If positive analogical transfer occurs, participants in group M should be at a relative disadvantage, having seen no prior problem.

Similarly, on problem MRT group M should perform better than group W (shown as M > W in the table), since those in group M are exposed to a relevant source problem (WC, a quantitative isomorph) immediately prior to seeing problem MRT. In contrast, participants in group W will have just seen an irrelevant source problem (WT, with reversed quantitative structure) and may be at a greater disadvantage on the MRT problem.

None of the group contrasts shown in Table 3.3 are statistically significant, and several show trends that are opposite what would be expected if participants were transferring material between adjacent problems. However, this analysis does not consider outcomes on prior source problems, which might facilitate or inhibit participants' use of problem comparisons. Table 3.4 and Table 3.5 also show percentage correct but take into account whether participants were correct, had no solution, or were incorrect on preceding problems. Thus, participants with an incorrect solution or no solution on a relevant prior source, and hence little of value to transfer, can be separated from those with a correct solution on a relevant prior source.

In these more detailed analyses, trends for problems MOD and WT are in the direction expected if positive transfer plays a facilitating role, but contrasts on problems MRT and WC show no similar effects. Even when these analyses were repeated using the presence of any conceptual error in a participant's solution attempt for each problem, group contrasts were still mixed and statistically unreliable.

Thus, the problem–ordering manipulation introduced to provide opportunities for positive and negative transfer had little effect on performance at the level of solution correctness. This finding is considered in more detail in the Discussion. Clearly, problems MRT and WT were most difficult, with percentages of participants reaching a

Table 3.4: Percentage correct on a target problem without any prior source problem versus with a relevant prior source problem.

Target	No source	Relevant source outcomes[†] C	N	I	Total
MOD (n)	(46)	(25)	(6)	(8)	(39)
C	89.1	92.0	100.0	87.5	92.3
N	4.3	0.0	0.0	0.0	0.0
I	6.5	8.0	0.0	12.5	7.7
WT (n)	(39)	(41)	(2)	(3)	(46)
C	64.1	65.9	0.0	0.0	58.7
N	15.4	12.2	50.0	0.0	13.0
I	20.5	22.0	50.0	100.0	28.3

[†]Outcomes: C = correct, N = no solution, and I = incorrect.

correct solution (51.8% and 61.2%, respectively) falling well below those reaching correct solutions on problems MOD and WC (90.6% and 91.8%, respectively).

Relations between solution outcome and tactical content. Table 3.6 shows tactical content and error categories for final problem–solving episodes. For tactical content, a participant receives a single category score, so cell frequencies sum to give appropriate column totals. A few protocols contain insufficient information to score tactical content in the final episode. For errors, a participant may achieve a correct solution in the final episode but still demonstrate several types of errors. As a result, cell entries for errors do not always add up to coincide with column totals.

The prevalence of tactical content and error categories in the final episode is generally consistent with findings for overall solution attempts. Looking within these attempts, the analysis can focus more closely on relations between tactic and outcome.

- Even within the final episode, not all solutions (correct or incorrect) are found using algebra. Excluding those with no solution or with contents that were not scorable, between 22.0% and 44.0% of participants (across problems) used other tactics to find their final solution.
- Use of ratios is the most prevalent form of nonalgebraic reasoning in final episodes, with the exception of an incorrect averaging procedure on problem WT. Model–based reasoning is the next most prevalent tactic.
- Algebra, model–based reasoning, and ratio tactics are comparably effective in the final episode. Averaging across problems, algebra is slightly more successful (79.5% correct) and slightly less error–prone (8.8% incorrect) than either of the nonalgebraic tactics in participants' final problem–solving episode.

Table 3.5: Percentage correct on a target problem with a relevant prior source problem versus an irrelevant prior source problem.

	Irrelevant source outcomes†				Relevant source outcomes			
Target	C	N	I	Total	C	N	I	Total
MRT (n)	(36)	(0)	(3)	(39)	(43)	(0)	(3)	(46)
C	58.3	0.0	33.3	56.4	48.8	0.0	33.3	47.8
N	27.8	0.0	33.3	28.2	34.9	0.0	0.0	32.6
I	13.9	0.0	33.3	15.4	16.3	0.0	66.7	19.6
WC (n)	(27)	(6)	(13)	(46)	(22)	(11)	(6)	(39)
C	96.3	83.3	92.3	93.5	86.4	90.9	100.0	89.7
N	0.0	0.0	0.0	0.0	4.5	9.1	0.0	5.1
I	3.7	16.7	7.7	6.5	9.1	0.0	0.0	5.1

†Outcomes: C = correct, N = no solution, and I = incorrect.

Table 3.6: Final episodes: tactical content and errors by outcome.

	Problem by Outcome†											
	MOD			MRT			WT			WC		
Category	C	I	N	C	I	N	C	I	N	C	I	N
n	77	6	2	44	15	26	52	21	12	78	5	2
Tactic												
Algebra	58	6	0	36	8	20	43	5	7	44	2	1
Model	3	0	0	4	2	6	2	1	2	12	1	0
Ratio	13	0	2	4	3	0	5	3	2	22	1	1
Procedure	0	0	0	0	0	0	1	11	1	0	0	0
Units	2	0	0	0	0	0	0	0	0	0	0	0
Not scored	1	0	0	0	2	0	1	1	0	0	1	0
Errors												
Conceptual	1	6	0	0	14	16	1	27	10	1	4	0
Manipulative	7	2	0	1	5	2	4	7	1	2	1	0

†Outcomes: C = correct; I = incorrect; N = no solution.

Thus, even within the final episode where a solution might be found, a normative account of problem solving consisting of successive algebraic transformations would be disconfirmed by these data. Instead, participants find solutions through a variety of reasoning strategies that, in some cases, involve relatively little formal algebra. The episodic structure of model-based reasoning tactics is examined more closely later.

Relations between solution outcome and errors. Errors observed during final episodes are also interesting, although more difficult to interpret because individual students can have multiple errors. The error categories are shown in the lower panel of Table 3.6 distinguish between conceptual errors, which arise through omission or commission of specific quantitative constraints, and manipulative errors, which arise through improper use of arithmetic, algebraic operations, or variables.

- With the exception of problem MOD, conceptual errors are more prevalent than manipulation errors. This is particularly true of the more difficult problems (MRT and WT). Participants who achieve a correct solution have fewer conceptual errors than those with an incorrect solution or no solution (1:6, 0:30, 1:37 and 1:4 across problems MOD, MRT, WT, and WC, respectively). In the few cases where a solution is found despite conceptual errors, offsetting manipulative errors fortuitously "correct" these conceptual errors.
- Although manipulative errors are found among participants not reaching a correct solution, they are also observed among those giving a correct solution. These errors are "repaired" within the final episode to allow for a correct solution. Among participants reaching an incorrect solution, the number with manipulative errors could not account for more than about a third of these failures (2/6, 5/14, 7/27, and 1/4 across problems). Thus, approximately two thirds of the incorrect solutions must be based on conceptual errors.

One interpretation of these results is that manipulative errors are less frequent and *more recoverable* than conceptual errors. That is, participants who make an error during a problem-solving episode are more likely to recover from that error if it stems from arithmetic or algebraic manipulation than if it is a result of misunderstanding the structure of the problem. Because errors may persist across episodes, this conclusion cannot be unambiguously supported. Nonetheless, the most serious errors among this group of relatively competent problem solvers are conceptual rather than manipulative.

3.4.3 Episodic structure of model–based reasoning

One of the most intriguing findings in these data are participants' use of model-based reasoning. In these episodes, they depart from the algebraic formalism and reason directly within the situational context presented by the story problem. This section examines the functional role that model-based reasoning plays within the overall solution effort. The analyses attempt to determine under what circumstances this form

Table 3.7: Errors and transitional status of a previous episode compared with the purpose of a model-based reasoning episode.

	Problem by Purpose†											
	MOD (n = 26)			MRT (n = 19)			WT (n = 30)			WC (n = 40)		
Previous episode	C	S	V	C	S	V	C	S	V	C	S	V
None	7	1	0	1	4	0	17	4	0	10	2	0
No errors												
On track	3	9	0	0	6	0	2	2	1	10	11	2
Abandon	1	0	0	1	1	0	0	2	0	0	0	0
Errors												
On track	1	0	0	0	0	0	0	0	0	1	1	0
Abandon	2	2	0	1	5	0	0	2	0	1	2	0

†C = comprehension; S = solution attempt; V = verification.

of reasoning occurs, what purpose it serves within a particular solution attempt, and what outcomes are likely when problem solvers reason in this fashion.

As with the analysis of final episodes, specific episodes within participants' solution attempts are identified where model-based reasoning occurs. The analysis also extracts the preceding problem-solving episode in the hopes of identifying enabling conditions for model-based reasoning. Because some participants' only use of model-based reasoning occurs during their first scored episode, they will have no preceding episode.

Precursors to model-based reasoning. A first task for describing the role of model-based reasoning in participants' solution attempts is to determine their reasons for using this method. These are evident in the contrast between correctness and transition out of an immediately preceding episode and the purpose (as rated) for using model-based reasoning.

Table 3.7 shows the number of participants using model-based reasoning for some purpose (scored as comprehension, solution attempt, or verification) subsequent to various conditions in the preceding episode. An individual may have (a) no preceding episode, (b) a preceding episode without errors, or (c) a preceding episode with one or more scored errors (i.e., an error of commission, omission, or manipulation from which the individual does not recover in that episode).

- From 26.3% (5 of 19 on MRT) to 70.0% (21 of 30 on WT) of model-based reasoning episodes occur as the first episode in a solution attempt. Of these initial model-based episodes, the majority (except for problem MRT) are undertaken for the apparent purpose of comprehending some aspect of the presented problem. The remaining initial episodes are scored as solution attempts.
- For participants having a preceding episode, their transition out of this episode is scored as achieving a subgoal, finding a solution, reaching an impasse, or deciding they are wrong.

Of the model-based reasoning episodes following an error-free episode, there are two essentially different scenarios. In the first, a participant's preceding episode ends with achieving a subgoal or finding a solution. This solution attempt can be considered "on track." In the second scenario, participants "abandon" the preceding episode after reaching an impasse (also after getting lost, as described earlier) or deciding that their efforts are wrong. They are technically on track because their preceding episodes are error-free, but they encounter sufficient difficulty that they abandon a previous line of reasoning in favor of model-based reasoning.

- Almost all participants who are on track in a preceding episode either attempt a solution or continue attempts at comprehension during the model-based reasoning episode.
- Only a few participants are on track and undertake model-based reasoning for the purpose of verification. On problem WC, these verification episodes follow finding a solution; the single verification attempt on problem WT comes from a student who verified a recalled formula using a simplification of the original problem.
- Participants abandon (i.e., decide they are lost, at an impasse, or wrong) a prior, error-free episode infrequently and subsequently use model-based reasoning for comprehension or to attempt a solution.

Model-based reasoning episodes following an episode with errors are less frequent than the case already discussed, but fall into similar categories. Relatively few participants have preceding errors, are unaware of those errors, and proceed as if on track (i.e., achieve a subgoal or find a solution). Participants who are aware of their preceding error nearly always decide that they are wrong and abandon the preceding episode. Among those who abandon a preceding episode with errors, subsequent model-based reasoning is used either for comprehension or as an attempt to find a solution.

Although based on a subset of all participants studied, these findings support an interpretation in which model-based reasoning plays four basic roles in problem solving: (a) as a *preparatory comprehension* strategy when the model-based episode is either the first problem-solving activity attempted or follows other comprehension episodes, (b) as a *solution* strategy when participants feel they are on track, (c) as an *evidence gathering* strategy when a solution has been found previously (this is infrequent), and

Table 3.8: Errors before and during model–based reasoning.

	Problem by Errors During Model Episode							
	MOD (n = 26)		MRT (n = 19)		WT (n = 30)		WC (n = 40)	
Before	Errors	None	Errors	None	Errors	None	Errors	None
No episode	0	8	2	3	1	20	0	12
Errors	1	4	2	4	0	2	1	4
No errors	0	13	1	7	0	7	0	23

(d) as a *recovery* strategy when participants suspect that their comprehension or solution efforts may be "off track." These interpretations are consistent with the earlier argument that reasoning within the situational context of a problem supports the generation of quantitative constraints, can be used directly as a solution method, or can be used to verify that these constraints are appropriate.

Effectiveness of model–based reasoning. To assess the effectiveness of model–based reasoning, the occurrence of any errors within successive episodes is examined. Table 3.8 shows the relationship between errors during a preceding episode (when there is one) and errors within the model–based reasoning episode.

- When model–based reasoning is the participant's first evident activity, as indicated by "No episode" in Table 3.8, errors are not often encountered within that episode. The two errors shown for problem MRT are misconceptualizations in which participants assumed that round–trip times were equal. The error in problem WT comes from a student who assumed that Mary and Jane did equal amounts of work.
- When a previous episode contains errors, the subsequent model–based episode is usually error–free. Thus, existing errors may be "repaired" during model–based reasoning. Across problems, from 66.7% to 100% (problems MRT and WT, respectively) of participants showed no conceptual or manipulative errors in the following model–based reasoning episode.
- Following an error–free episode, only one student introduced a new error with model–based reasoning by omitting the constraint that distances are equal on problem MRT.

Although these findings are not conclusive, they are again consistent with the four hypothetical roles for model–based reasoning described in the analysis of final episodes. First, preparatory comprehension promotes an error–free conceptualization of the problem situation, enabling participants to assemble correctly the quantitative

Table 3.9: Percentage of all tactical episodes introducing any correct structural inference.

	Tactic					
Problem	Algebra	Model	Annot.	Ratio	Proc.	Unit
MOD	68.6	83.9	33.3	15.8	0.0	25.0
(n)	(105)	(31)	(81)	(19)	(0)	(4)
MRT	66.9	76.2	54.1	43.8	100.0	100.0
(n)	(118)	(21)	(61)	(16)	(1)	(1)
WT	57.1	90.0	22.6	38.9	25.0	0.0
(n)	(84)	(33)	(31)	(18)	(20)	(1)
WC	77.6	54.9	5.9	17.0	0.0	0.0
(n)	(76)	(51)	(34)	(47)	(0)	(1)

structure of the problem during later reasoning episodes. Second, participants also attempt to find solutions directly through model–based reasoning, generally without introducing errors. Third, after encountering an error during previous problem–solving activities, participants may be able to recover through the use of model–based reasoning. Fourth, model–based reasoning can play a confirmatory role when participants have identified important problem constraints or a possible solution.

3.4.4 Tactical Course of Structural Inferences

The preceding analyses of model–based reasoning as a problem–solving tactic looked within participants' solution attempts to find what relations held between this tactic and conceptual errors, both before and during an episode of model–based reasoning. In this section, a similar form of analysis tracks the introduction of correct inferences about problem structure and conceptual errors of omission or commission. The analysis also takes a closer look at the tactical course of "repairing" conceptual errors during a solution attempt. These analyses test hypotheses about the role of model tactics in generating and evaluating quantitative constraints.

Tactical origin of correct structural inferences. Table 3.9 pools over participants' solution attempts to show which tactical episodes are most likely to introduce correct structural inferences about each problem. A structural constraint is "introduced" on its first scored occurrence in a participants' solution attempt. For example, on problem MOD a student whose first use of the constraint that times are the same comes during a model–based reasoning episode would contribute one case to the number of structural inferences attributed to model–based reasoning. By analyzing the percentage of all episodes of a particular tactic that introduce a structural inference, this analysis adjusts for the number of opportunities for structural inferences afforded by more or less frequent tactical categories. Naturally, tactics that occur infrequently give poorer estimates of the origin of structural inferences.

- Excluding tactical episodes that occur very infrequently (i.e., Unit and Procedure, except on problem WT), model–based reasoning is more likely than any other tactic to introduce correct structural inferences on problems MOD, MRT, and WT. By comparison with algebraic tactics on these problems, an episode of model–based reasoning is from 1.2 to 1.6 times more likely to introduce a correct problem constraint (problems MOD and WT, respectively).
- On problem WC, this trend is reversed, with algebra tactics 1.4 times more likely to originate a correct structural inference. Even on this problem, however, over half of the problem–solving episodes scored as using model–based tactics introduce a correct structural inference.
- The constraints introduced using a procedure tactic on problem WT correctly add work output during an otherwise incorrect attempt to find an average time.
- Pooling over problems and excluding procedure and unit tactics, model tactics are most likely to introduce correct constraints (73.5%), followed by algebra (67.4%), annotation (33.3%), and ratio tactics (25.0%). These differences are statistically reliable (Tactics × Correct constraints, $\chi^2(3) = 118.0$, $p < .001$).

Rather than playing an "auxiliary" role in competent problem solving (Paige and Simon, 1966), using a dimensional model to reason about problem structure appears to play a central, generative role. This supports two important hypotheses: (a) situational structure as a normative domain ontology may enable constructive and constraint–generating inferences, and (b) model–based reasoning as a problem–solving tactic is important for preparatory comprehension.

Tactical origin of incorrect structural inferences. Just as some tactics are more likely to introduce correct structural inferences about a problem, this tracking analysis can show which tactics are most likely to introduce conceptual errors. Table 3.10 shows an identical analysis of the tactical origin of conceptual errors, again pooling over participants' solution attempts. Excluding infrequent tactical categories (i.e., Unit and Procedure, except on problem WT), a complementary image of the relative risks of model–based and algebraic tactics is found.

- On all four problems, algebraic tactics are more likely to introduce conceptual errors than are model–based tactics — from 1.6 times more likely on problem MOD to 9.8 times more likely on problem WC.
- Ratio tactics are comparable to or more error prone than model–based tactics, except on problem MOD where most participants constructed a ratio of the total distance to the distance travelled after the first hour without introducing any conceptual errors.

Table 3.10: Percentage of all tactical episodes introducing any incorrect structural inference.

	Tactic					
Problem	Algebra	Model	Annot.	Ratio	Proc.	Unit
MOD	15.2	9.7	6.2	5.3	0.0	25.0
(n)	(105)	(31)	(81)	(19)	(0)	(4)
MRT	42.4	23.8	8.2	25.0	100.0	0.0
(n)	(118)	(21)	(61)	(16)	(1)	(1)
WT	27.4	3.0	0.0	44.4	85.0	0.0
(n)	(84)	(33)	(31)	(18)	(20)	(1)
WC	19.7	2.0	2.9	6.4	0.0	0.0
(n)	(76)	(51)	(34)	(47)	(0)	(1)

- On problem WT, ratio and procedure tactics introduce more conceptual errors than any other tactic. For ratio tactics on this problem, most errors originate when participants introduce a conceptual error of commission over working times (e.g., adding times).
- For procedure tactics on problem WT, most conceptual errors originate when participants attempt to find an average working time.
- Again pooling over problems and excluding procedure and unit tactics, algebra tactics are most likely to introduce incorrect constraints (27.2%), followed by ratio (16.0%), model (7.4%), and annotation tactics (5.3%). These differences are also statistically reliable (Tactics × Conceptual errors, $\chi^2(3) = 56.9$, $p < .001$).

As with the origin of correct structural inferences, these findings largely support hypotheses about quantitative inferences based on situational structure and model–based reasoning as a problem–solving tactic. Furthermore, the relatively error–free status of model–based episodes admits the possibility that this tactic might be used to recover from conceptual errors introduced during algebraic or other tactical episodes. This hypothesis is examined next.

Detection and repair of conceptual errors. Since participants' solution attempts are broken into successive episodes of coherent problem–solving activity, it is possible for conceptual errors introduced early in a solution attempt to be "repaired" during later episodes. For example, student M42 subtracts distances in an algebraic episode on problem MOD (an error of commission), but then decides to add distances during a subsequent model–based episode. This represents one operational case of

Table 3.11: Percentage of tactical episodes retracting or replacing an earlier conceptual error.

	Tactic					
Problem	Algebra	Model	Annot.	Ratio	Proc.	Unit
MOD	9.5	19.4	2.5	0.0	0.0	25.0
(n)	(105)	(31)	(81)	(19)	(0)	(4)
MRT	22.0	23.8	9.8	31.2	0.0	0.0
(n)	(118)	(21)	(61)	(16)	(1)	(1)
WT	15.5	6.1	3.2	11.1	25.0	0.0
(n)	(84)	(33)	(31)	(18)	(20)	(1)
WC	10.5	7.8	0.0	8.5	0.0	0.0
(n)	(76)	(51)	(34)	(47)	(0)	(1)

repair, where a conceptual error is replaced by a correct constraint. The other case of repair occurs when participants remove an earlier conceptual error but do not explicitly replace it with a correct structural inference.

Table 3.11 shows the relative frequency with which prior conceptual errors are removed or replaced with correct structural constraints during various tactical episodes.

- On problems MOD and MRT, model–based tactics are comparable to or better than algebraic tactics for accomplishing repairs, while on problems WT and WC, algebraic tactics surpass model–based reasoning for repairing prior conceptual errors.
- On problem MRT, constructing a ratio is more likely than any other tactic to accomplish repairs, and on closer inspection most of these cases show either specific attention to an earlier conceptual error (e.g., establishing a ratio over times after omitting their addition) or a complete reconceptualization of problem structure.
- The large percentage of "repairs" using a procedure on problem WT are an artifact of displacing earlier conceptual errors with an equally problematic averaging tactic.
- Pooling over problems and excluding procedure and unit tactics, algebra tactics are most likely to repair conceptual errors (14.9%), followed by model (12.5%), ratio (11.0%), and annotation tactics (4.3%). These differences are statistically reliable (Tactics × Repairs, $\chi^2(3) = 15.0$, $p < .002$).

Combined with estimates of the relative likelyhood of introducing correct and incorrect structural inferences, the analysis of "repairs" provides support for the hypothesis that model-based reasoning can be used to repair conceptual difficulties introduced during previous problem-solving episodes. Tracking the introduction, removal, and repair of conceptual errors within solution attempts shows that model-based reasoning has no particular advantage over other tactics in repairing conceptual errors. However, since these errors are most likely to occur within episodes using other tactics (i.e., algebra, ratio, or procedure), reasoning about a dimensional model of problem structure plays an important role in competent problem solving by introducing correct structural inferences, seldom introducing conceptual errors, and sometimes repairing earlier conceptual errors.

3.5 Discussion

Interpreted as a series of problem-solving episodes, the written protocols provide an opportunity to look within individual solution attempts for strategic and tactical activity. Analyses reported in this Chapter look across a relatively large sample of mathematically sophisticated students in an effort to describe "typical" problem-solving behaviors. This section briefly summarizes the major quantitative findings and compares these results with other research on mathematical problem solving. A discussion of the implications of these findings for conceptions of mathematical knowledge and instruction appears in the concluding chapter (Chapter 6). This study yields several unexpected findings and leads to a set of hypotheses about constructive mathematical competence that provide a framework for more detailed protocol studies of algebra story problem solving (Chapter 4) and motivate a functional account of how people construct and use dimensional models of problem structure as a major component of competent problem solving (Chapter 5).

3.5.1 Summary of quantitative findings

As part of an effort to explore the episodic structure of algebra story problem solving, the previous section presented four levels of quantitative analysis: (a) the prevalence of different interpretive categories in participants' overall *solution attempts*, (b) relations between outcomes, tactical content, and errors in their *final episodes* of problem solving, (c) the role and effectiveness of *model-based reasoning* episodes within the wider problem-solving context, and (d) the tactical course of *structural inferences* across participants' solution attempts. Each successive level of analysis tightened the focus on findings at coarser levels.

A global view of solution attempts revealed significant nonalgebraic reasoning as a prevalent and somewhat unexpected activity in competent problem solving. Most prevalent among these tactics was model–based reasoning. Among observed errors, conceptual omissions or commissions were more frequent than manipulative errors within arithmetic or algebraic formalisms. An examination of final episodes, the "bottom line" in a very lean view of these problems, corroborated this global image of significant nonalgebraic reasoning on nonroutine problems. Looking more closely at errors, analyses showed that manipulative errors were both less frequent and less damaging than conceptual errors, because problem solvers were more likely to recover from errors of manipulation within the final episode.

After examining the episodic structure of model–based reasoning, four roles for this tactic were proposed: (a) as preparatory comprehension, (b) as a solution method, (c) as evidence–gathering for a candidate solution, or (d) as a recovery method for errors generated earlier in the solution attempt. These hypothetical roles were supported by an analysis of the tactical course of structural inferences, both correct and incorrect, across participants' solution attempts. Model–based tactics competed favorably with algebraic and other tactics for introducing correct structural inferences, while at the same time introduced far fewer conceptual errors than algebraic tactics. In addition, model–based reasoning was effective in repairing conceptual errors introduced in preceding episodes and was generally comparable to other problem–solving tactics in this role.

These quantitative analyses of problem solving episodes corroborate the earlier description (Chapter 2) of an interplay between the quantitative and situational structure of algebra story problems. The following sections take up the implications of these findings for normative accounts of competent algebraic problem solving, the prevalence and consequences of problem comparisons for transfer across solution attempts, and the position of model–based reasoning in a theoretical account of mathematical understanding and the construction of mathematical representations.

3.5.2 Competent problem solving

The study reported in this chapter explores competent algebra story problem solving. The term *competent* contrasts problem–solving activity observed among advanced undergraduates with images of problem–solving "expertise" portrayed in the literature. For example, Hinsley *et al.* (1977) and Mayer *et al.* (1984) reported that experienced problem solvers use problem–solving schemata to categorize problems by type and then solve these problems using recalled quantitative constraints. Although this account corresponds with some of the protocols, many participants in this study appear to *construct* solutions to algebra story problems. Rather than resembling a smooth execution of a highly practiced skill or the application of a familiar schema, these constructions often proceed with difficulty and include reasoning activities only partly connected to algebraic or arithmetic formalisms.

As noted earlier, participants in this study should be considered competent mathematical problem solvers. Nonetheless, judging from the episodic activities observed, these algebra story problems were not routine tasks. On problems MRT and WT, for example, many failed to reach a correct solution, and those who did succeed often experienced considerable difficulty. Analyses of errors during solution attempts show that conceptual errors of omission and commission are both more prevalent and more damaging than manipulative errors in algebra or arithmetic. These results support an account of algebra story problem solving in which problem comprehension and solution are complementary processes. Integrating situational and quantitative materials in a problem is a central aspect of competence. These intermediary structures provide a representational bridge between the text of a problem and a quantitative solution. Reasoning about the situational context of a problem can serve as a justification for assembling quantitative constraints that may eventually lead to a correct solution. Empirical findings show that a substantial portion of problem-solving activity is devoted to reaching an understanding of the problem that is sufficient for applying routine calculations.

Despite their mathematical backgrounds, perhaps advanced undergraduates have yet to achieve competent algebra story problem solving, well beyond the curricular setting designed to teach it. Alternately, they may have been "experts" during and shortly after algebra instruction, but with the passage of time they have lost the facile performance demonstrated by Hinsley *et al.* (1977). As a practical matter, neither position provides an entirely satisfactory explanation, because the issue remains how to characterize ostensibly competent problem solving in a clientele for whom the algebra curriculum is designed. Recent studies of mathematical problem solving in "practice" (Carraher, Schliemann, and Carraher, 1988; de la Rocha, 1986) have presented similar images of competent quantitative reasoning: problem solvers organize their quantitative knowledge around the demands of the situational context presented by the task, often using the problem situation (or knowledge of it) to assemble or verify quantitative constraints. As a theoretical matter, then, the question of what it means for people to construct and understand a mathematical representation remains open. If an image of competent problem solving in this domain is to inform teaching efforts — that is, if it is to have some predictive capacity as described in the introduction of this dissertation — then activities like these are a legitimate topic of study.

3.5.3 Transfer effects

The problem materials used in this study were chosen as quantitative isomorphs and then presented in a sequential order that might influence participants' use of analogical comparisons between adjacent problems. Other studies of analogical transfer with algebra story problems have produced mixed results, but show that both positive and negative transfer sometimes occur. Positive transfer has been more likely when participants are alerted to the experimental manipulation (Reed, 1987; Reed *et al.*, 1985) or are high in mathematical achievement (Novick, 1987). Transfer effects related to higher achievement have been attributed to participant's improved attention to aspects of quantitative structure (Novick, 1987; Silver, 1979) and better memory for previous solution methods (Silver, 1981). Negative transfer in participants with lower achievement (Novick, 1987) has been attributed to a reliance on inappropriate problem features and an inability to reject misleading analogical sources. Finally, Dellarosa (1985) experimentally manipulated participants' use of analogical and schematic problem comparisons to produce improvements in their categorization and solution of related problems.

This study did not alert participants to the comparability of problems or encourage them to look back over their prior solutions as they worked through the problems. Their backgrounds ensure high mathematical achievement, and entrance requirements for academic majors in computer science and engineering preselect for high quantitative abilities. There is no performance–level evidence of positive or negative transfer within the problem–solving session, despite a manipulation of problem structure and presentation order to elicit these effects. At the aggregate level, participants appear to take the "school math" task at face value: each problem, presented individually on a blank sheet of paper, is treated as a self–contained exercise, rather like what a student might face during examinations in an algebra course. In many ways, this is an entirely reasonable reading[3] of the study setting by participants, since relevant similarity between test items would be redundant in a conventional school assessment.

Despite finding no reliable evidence of transfer between solution attempts, a closer inspection of individual protocols and explanatory remarks shows that some participants do exhibit negative transfer. In some of these cases, transferred material directly violates the quantitative and situational structure of the target problem. For example, student W08 incorrectly attempted to add working rates on problem WC, first writing $1/5 * boxes + 1/2 * boxes = 56$, followed by $7/10 * boxes = 56$. In explanatory remarks, they stated: "The mistake I made was that I assumed it was like problem 1 where they work together." In the preceding solution to WT, this student had written "Together = 1/5 + 1/4 in one hour = 9/20" and then correctly divided one job by the combined rate. Adding working rates in problem WT is justified because Mary and Jane work together at the same time. However, situational and quantitative relations are exactly reversed in problem WC (see Figures 2.2(b) and 2.6 in Chapter 2). Because times are added together (adjacent) and work is performed on the same boxes (congruent), the addition of working rates (i.e., output over time) cannot be similarly justified.

In other cases, participants recognize an appropriate source problem but then fail to transfer information at the correct level of abstraction. For example, on problem MOD student W01 correctly attempted to add motion rates, but used an algebraic expression of the form: $1/60 + 1/100 = x/880$. On the previous (WT) problem, the student managed a correct solution using an expression of the form, $1/5 + 1/4 = 1/x$, and

[3]Pea (1987) makes a similar argument for "reading situations" when reviewing studies that compare quantitative reasoning in practical versus school settings.

remarked that this "is a formula used to find a total of time they work together." Although the addition of rates can be justified in both problems, it appears that the rate form in the retrieved formula was reversed (i.e., time over output) when used in a solution attempt on the MOD problem. Thus in a situation where participants should benefit by transfer of a solution approach, their failure to justify transferred material actually produces a negative effect.

It may be that the problem–solving context (completing a test booklet in a proctored examination setting) as well as not alerting participants to the comparability of problems prevented them from recognizing and elaborating effective analogical comparisons between problems. In verbal interviews reported in the next chapter, participants are sometimes prompted to make problem comparisons. Analyses of their protocols show that spontaneous attempts at analogical inferences between algebra story problems are common. These comparisons are sometimes lengthy and can introduce misconceptions, but they also frequently lead to fruitful explorations of problem structure, both quantitative and situational. In addition, comparisons need not encompass the entire problem structure but can make effective use of relevant substructural similarities. These alternative findings are largely consistent with other verbal protocol studies of learning from worked examples (Chi, Bassok, Lewis, Reimann, and Glaser, 1989; Pirolli and Anderson, 1985; Singley, 1986) and suggest that analogical comparison may be a common problem–solving and learning strategy in settings where people have some control over their work.

3.5.4 Model–based reasoning

The findings of this chapter are not the only documented evidence for model–based reasoning during mathematical problem solving. Several psychological studies have found similar evidence, although interpretations of this activity vary. Paige and Simon (1966), comparing human protocols with Bobrow's (1968) computational model of translating algebra story problems into equations, found that participants with varied mathematical backgrounds used "auxiliary representations" of the physical setting of a problem. These representations allowed some people to detect impossible problems before constructing algebraic expressions or to assemble relevant quantitative constraints (e.g., additivity in part–whole relations). Studying the prevalence of Polya's (1945) heuristics for mathematical problem solving, Kilpatrick (1967) reported that 60% of an above–average group of eighth graders used "successive approximation" while attempting to solve word problems. These trial–and–error approaches were often successful and were sometimes combined effectively with more deductive solution strategies. Silver (1979) found similar successful approximation strategies in students who had yet to receive formal algebraic training.

Studying geometry problems, Schoenfeld (1985) found that students used a trial–and–error approach to generate hypotheses about geometric relations and then evaluated these hypotheses by physical construction. He argued that these exploratory episodes of "naive empiricism" were usually poorly organized and often interfered with forms of deductive verification that students had been trained to use. Finally, Kintsch and Greeno (1985) described a process model of solving arithmetic word problems in which quantitative strategies were triggered by information contained in a "situation model" of the problem. The situation model was constructed during text comprehension and contained a set-based representation of typed quantities and their interrelationships (e.g., part-whole). Kintsch (1986, 1988) has since shown that the construction of a situation model is important for recall, inference, and learning from text.

Looking over this evidence, it is clear that studies of mathematical problem solving consistently encounter activities similar to model-based reasoning: people construct some form of situation model, make inferences within the model to help comprehend and sometimes to solve a quantitative problem, and use the model in a supportive role to assemble or to verify quantitative constraints. Beyond model-based reasoning in mathematical problem solving, similar evidence is available across a wide range of cognitive activities. For example, Johnson-Laird (1983; Johnson-Laird and Bara, 1984) argued for a model-driven account of syllogistic reasoning that underlies common-sense inference. Given a pair of premises such as *All the artists are beekeepers* and *All the beekeepers are chemists*, Johnson-Laird's participants appeared to build successively more elaborate models of the situation described by the premises when searching for valid inferences. The validity of each inference, rather than being logically deduced by sound rules of inference, was evaluated with respect to these concrete models of the premises. Errors occurred when participants were unable to build sufficient models of the premises and thus overlooked or failed to eliminate various inferences. Relatively concrete forms of reasoning outside traditional (i.e., schooled) formalisms have also been observed for decision making under uncertainty (Tversky and Kahneman, 1974), various forms of statistical reasoning (Nisbett, Fong, Lehman, and Cheng, 1987), and explanations of physical processes (Clement, 1983; McCloskey, 1983).

In general, these studies raise questions about the relationship between what students bring to an educational setting — that is, their preconceptions about a subject matter — and materials that the curriculum explicitly presents. In the domain of mathematical problem solving, students' preconceptions and associated activities are often pushed to the background of legitimate practice and inquiry. At best they are auxiliary to quantitative reasoning; at worst they interfere with preferred problem-solving activities and produce "lost opportunities, unfocused work, and wasted effort" (Schoenfeld, 1985, p. 308). In their stead, the manipulation of symbolic representations of quantity, apart from the situations and activities that give rise to these quantities, is held in the foreground. Findings for model-based reasoning in this chapter, in concert with other studies reviewed briefly, suggest that this foreground/background conception of quantitative problem solving may need to be reconsidered.

In this sample of advanced undergraduates, a routine problem is one in which the use of familiar algebraic operations will provide a *precise* value for an unknown entity. This is the power of the algebraic formalism: it is perfectly general, sound, and often simple to apply. However, *quantitative precision is of little value when a participant is uncertain about the problem's structure.* Findings on overall episodic activity, the frequency and consequence of conceptual versus manipulative errors during those episodes, and the role of model–based reasoning show that routine activities within the algebraic formalism make up only a portion of competent problem solving. For many competent problem solvers, algebra story problems are not routine exercises. Instead, much of their problem–solving activity is devoted to assembling a sensible set of constraints on a desired quantity, an effort that uncovers the problem's structure. When algebraic constraints are unclear, participants sometimes attempt solutions using model–based reasoning (e.g., Figure 3.2), a tactic that approximates a *certain* value for an unknown entity. The value is certain when quantitative constraints that determine its derivation are embedded in a model of problem structure that is familiar to the participant.

The strategic significance of model tactics is consistent with various explanations. On one hand, enacting a set of physical constraints may provide otherwise skilled problem solvers with an efficient means of estimating quantitative solutions. Under this interpretation, the model–based episode shown in Figure 3.2 results simply from the student's preference for repeated additions over a more complicated division. Wilkening (1981) made a similar argument when interpreting results of a developmental study on the relationship between velocity, time, and distance. In contrast, the preceding analyses suggest that episodes of model–based reasoning serve as problem–solving strategies in their own right and are used when more "formal" activities (e.g., algebraic substitution) are unreachable given the current problem representation. Under this interpretation, the student in Figure 3.2 undertakes model–based reasoning because her representation of the problem cannot justify a division of the total distance by a combined rate. Enacting motion and time constraints over successive hours of travel makes the quantitative structure of the problem more certain and supports a conceptualization of quantitative constraints in which the total distance can be divided by a combined rate to give a precise account of the elapsed time. Further constraints are introduced by establishing that the correct quantitative solution must fall between the fifth and sixth hours of travel.

Chapter 4
Representations Observed: Learning and Teaching

4.1 A wider view of problems, problem solvers, and solutions

The preceding study of written problem–solving protocols presents an image of competence for the domain of algebra story problem solving. Findings of significant episodic activity outside the algebraic formalism are surprising in contrast with existing accounts of competent problem solving in the empirical literature and intriguing for a wider account of nonroutine mathematical reasoning. This chapter presents a more detailed analysis of think aloud protocols taken during problem–solving interviews. These analyses expand the restricted observational window on problem–solving activity that written protocols afford, both to confirm earlier findings and to obtain a more detailed account of several hypotheses about the construction and evaluation of quantitative inferences. In addition, these verbal protocols extend the analysis to a wider range of participants, including newcomers to the algebraic formalism at one extreme and algebra teachers at the other extreme.

The chapter is organized as a conventional cognitive study, first describing the methods used to collect verbal reports of problem solving and analytic conventions for using these data. Results present a series of comparisons between algebra students, their teachers, and advanced undergraduates from the previous chapter. Comparisons range across problem structures and across tactical variations in problem–solving activity to explore a set of outstanding questions that open the chapter. In addition, these comparisons provide new explanatory material for understanding competent quantitative inference and its acquisition.

4.1.1 Overview of major findings in written protocols

In the written protocol study, competent algebraists produced varied materials during their solution attempts on representative algebra story problems. In particular, the analysis focused on three major interpretive findings:

1. Solution attempts were interpreted as having an episodic structure, with individual episodes identified on the basis of coherence in strategy, tactic, and content.
2. Episodes using nonalgebraic tactics were quite common, including: diagrammatic annotations, inferences drawn over a dimensional model of problem structure, and various uses of ratios without algebraic notation.
3. Model-based reasoning tactics played several different roles within the episodic structure of solution attempts: (a) generating inferences about structural constraints including algebraic notation, (b) directly determining unknown values without introducing algebraic notation, and (c) evaluating or repairing inferences about problem structure.

Each of these findings was sensible but a little curious, given that the initial purpose of the study was to gain an empirical image of algebraic competence. Why is it that people, well past mastery of school algebra, place so much of their activities outside of the traditional algebraic notation?

4.1.2 Outstanding questions

While the interpretive analysis of written protocols allows us to look within the episodic structure of problem–solving attempts, these analyses are inherently conservative because the means of observation are so narrow. As mentioned earlier, we have no timing information, must rely on the written order in which various tactics are used, and cannot query participants about what they take various notations to mean or about their reasons for making transitional decisions. This interpretively conservative window onto problem–solving activities raises a number of additional questions that call for more detailed verbal protocols collected with more diverse participants. These questions concern:

- The time course and pattern of episodic sequences.
- Participants' reasons for transitioning between episodes, including their accounts of problem–solving impasses.
- The internal structure and content of model–based reasoning episodes.
- Interactions between algebra and model tactics, particularly as they give rise to inferences about problem structure.

- The prevalence and outcome of problem comparisons, particularly given a change in problem setting from a written examination format to a problem–solving interview.
- A contrast between competent problem solvers and newcomers to the algebraic formalism.

Each of these issues is important for understanding the cognitive activities involved in competent mathematical reasoning. If significant aspects of what we have called model–based reasoning persist at the strategic and tactical levels of competent problem solving, do algebra students and their teachers exhibit similar forms of quantitative inference? In their problem–solving activities, do they use model–based reasoning to construct, support, and evaluate different kinds of quantitative representation? This chapter combines qualitative and quantitative analyses of materials collected during relatively open–ended verbal interviews with algebra students and algebra teachers. Results of these analyses confirm and extend exploratory findings from Chapter 3, as well as comparing different problem solvers sampled at various locations along an idealized continuum of problem–solving competence.

4.2 Method

In Chapter 3, an analysis of each solution attempt as a sequence of episodes allowed various quantitative summaries of the relation between tactics, representational content, and outcomes. In the analyses of verbal protocols that follow, contrasts between problem–solving tactics are primarily qualitative, informed by earlier findings but directed towards an account of how various tactical approaches lead to structural inferences and transitional outcomes in different problems.

There are methodological problems with studying any spontaneous human activity, particularly where the circumstances of the activity are part of the research question. The study reported in this chapter replicates and extends earlier findings by (a) looking for evidence of model tactics among different problem solvers in less restrictive settings, and (b) collecting verbal report data about transitions between tactical episodes.

There is a simple logic to this design that follows a "constant comparative method" developed in sociological approaches to grounded theory (Glaser and Strauss, 1967; Strauss, 1986). In grounded theory, comparative analyses of qualitative data iteratively confirm and extend theoretical categories by moving their interpretation strategically across different settings. Each of the comparisons listed above is intended to carry interpretive categories constructed in the the written protocol study into different positions along relevant sampling dimensions: problem solvers, settings, and detailed personal accounts of problem–solving activity.

The materials of these comparative analyses are careful interpretive readings of observed problem–solving activity in terms of theoretical categories carried into the analysis or generated during its undertaking. For example, the previously monolithic category of model–based reasoning is further refined by the observation that quite different notational systems are employed during these tactical episodes. Qualitative and quantitative analyses in this chapter provide one explanatory account of participants' problem–solving activity, presented in enough detail that a sceptical reader could construct a different explanatory account.

Like conventional reports of verbal protocol studies, I use transcript material as an illustration of theoretical categories and contingencies. Certainly this presentation is selective, driven by an attempt to sample different settings with an interpretive category in hand (e.g., diagrammatic notations for state simulation). Unlike conventional uses of verbal protocols, however, I also treat transcript material as a resource with continuing value beyond my own reading of it. Thus, I include both verbal and written transcript materials freely within the body of the chapter, in addition to reporting analytic summaries of problem–solving interviews.

4.2.1 Participants

Participants in these problem–solving interviews were selected to give boundaries on either "side" of the competent problem solvers from the preceding chapter, along a continuum of algebraic expertise. At the lower end, two students were selected, one in an introductory algebra class at a local junior high school and the other in an intermediate algebra class at a local junior college. At the upper end of the continuum, algebra teachers from similar settings were selected, one an instructor in math and science at the high school level and the other a mathematics instructor at a local junior college. These four respondents are not presented as a statistically representative sampling from various levels of mathematical schooling and achievement. Instead, they provide a theoretically relevant contrast for questions about tactical activity and personal significance in competent mathematical problem solving. Participants were paid $3.00 per hour for the problem–solving interviews.

Celeste. Celeste is an eighth grade student taking a beginning algebra course. She describes her performance in the class as "average" and relates that algebra is not her strongest subject area. The problem solving interview is conducted in the setting where Celeste usually does her algebra homework: she sits on her bed, writing on a notebook held in her lap, with finished problems placed around the work surface (her bed) in small piles. While working on a new problem, she periodically rearranges the finished work, pulling selected sheets to the surface and leaning over them to inspect her earlier activities.

Celeste's introductory algebra course is organized around 110 "problem sets," and students progress en mass through the sets during the year. She is currently on lesson 88, which includes a worksheet with several word problems like:

> The creature wolfed down 36% of the food. If there were 75 pounds of food to start, how many pounds were left?

According to Celeste, her teacher solves example problems at the board, while students read the problem and suggest equations to write down. When the students are silent, he writes in equations of his own and tells them why those equations fit the story. She describes the teacher's explanations as using "another language" that she and her classmates would later translate for each other, depending on who was having difficulty. Asked if his explanations contain words like "expression," "term," "variable," or "polynomial," she laughs and reports that she and her friends talked to each other by saying "times by this thing" and so on.

Karen. Karen is a 24 year old undergraduate in a local junior college. She is currently enrolled in an intermediate algebra course, and completed an introductory course the previous semester. These courses satisfy background mathematics requirements for further undergraduate study in a four year undergraduate program. Her study plans are still open regarding a major concentration, and she has not yet made applications to other colleges or universities. While Karen is also new to algebra, she has received more instruction than Celeste and has more concrete contingencies surrounding her school performance, since her matriculation involves transferring algebra courses for undergraduate credits.

Karen works at her dining table, where she usually does her algebra homework. On seeing that we would be doing story problems, she commented that she finds story problems very difficult and that her algebra instructor spends lots of time on these problems in class. Karen takes algebra as a personal challenge, mentioning that by treating the algebraic notation as a foreign language to be learned (she likened it to Chinese) she is able to lessen her apprehension about these classes. She has designed a system in which a roommate gives her M-and-M candies for each equation she solves, placing herself on a continuous reinforcement schedule for what she finds an arduous task. Piecing her comments on algebra together, she finds the subject difficult, feels that her lack of confidence interferes with performance, and has turned algebra into a personal project that extends beyond the classroom.

Paul. Paul is a 40 year old science and math teacher, who works in a local hospital with an adolescent CARE Unit for children with substance abuse problems. He teaches mathematics and science in a classroom setting, depending upon the clientele in any given week, but his teaching regularly includes introductory algebra. Paul specializes in teaching approaches with "learning disabled" children and, during discussions of his current teaching activities, shows a creative facility for making quantitative problems concretely accessible to students. He creates applied problem–solving projects for his students, such as estimating the number of beans in a jar or the number of hairs on a student's head, that encourage students to develop physical approaches when using a given formula (e.g., the volume of a cylinder). These projects are enormously popular with otherwise very difficult students.

The interviews were conducted out of the hospital setting where Paul teaches, both in office settings where Paul sometimes prepares classroom materials. Between problem–solving attempts, he describes two objectives in his approach to teaching mathematics: (a) to make the problems real and accessible for his students and (b) to encourage practical approaches to finding quantitative solutions, using algebraic manipulation of formulas as supporting tools. According to Paul, this approach reflects his own discovery of *what* works with students, rather than a teaching approach he learned as part of his certification. Though his work in the interviews and his reflective statements about mathematics contain material about teaching students and about how he learned to solve story problems, Paul engages the interview tasks as a problem solver rather than as a teacher.

Richard. Richard is a mathematics instructor in a community college and was recommended by an adult (the mother of another subject, not reported in the following analyses) who had taken an algebra course with him. He is well known for making difficult concepts accessible for students who have had few positive experiences in the mathematics curriculum and are returning to school for some form of professional certification. His reputation is both for putting apprehensive students at ease and for giving many of them their first experience of having understood a mathematical idea. Richard also writes a popular math puzzle column for a local paper. While Paul describes mathematics as a relatively minor part of his work as a teacher, Richard describes his work as teaching mathematics and his interests in the subject extend well beyond the classroom.

The interview was in an empty classroom where Richard teaches. He works at the board, alternating between explaining his teaching style when the problem is routine and a genuine verbal report of problem solving when the problem is not routine. The classroom and board provide an interesting mix of public and private settings, on the one hand making his work a performance for an audience trying to understand what he is doing (myself or his students), but on the other hand making it possible for him to be wrong while embedded in surroundings that emphasize his role as an authority. As a result, the difference between routine and nonroutine problems takes on a special significance, since the uncertain outcome of nonroutine problems introduces risk into Richard's public performance. Nonroutine problems, presented by an academic observer, are "live" in a way that classroom demonstrations often are not.

4.2.2 Materials

Problems used in this and the preceding chapter are drawn from a space of compound algebra story problems about motion and work. Problems can be generated by choosing different time and output relations to be depicted in the story text. For example, in problem MOD of the preceding chapter, distances travelled are collinear and adjacent, since the trains share a starting place and travel away from each. Times in this problem are the same, since the trains start together and reach 880 kilometers apart at some

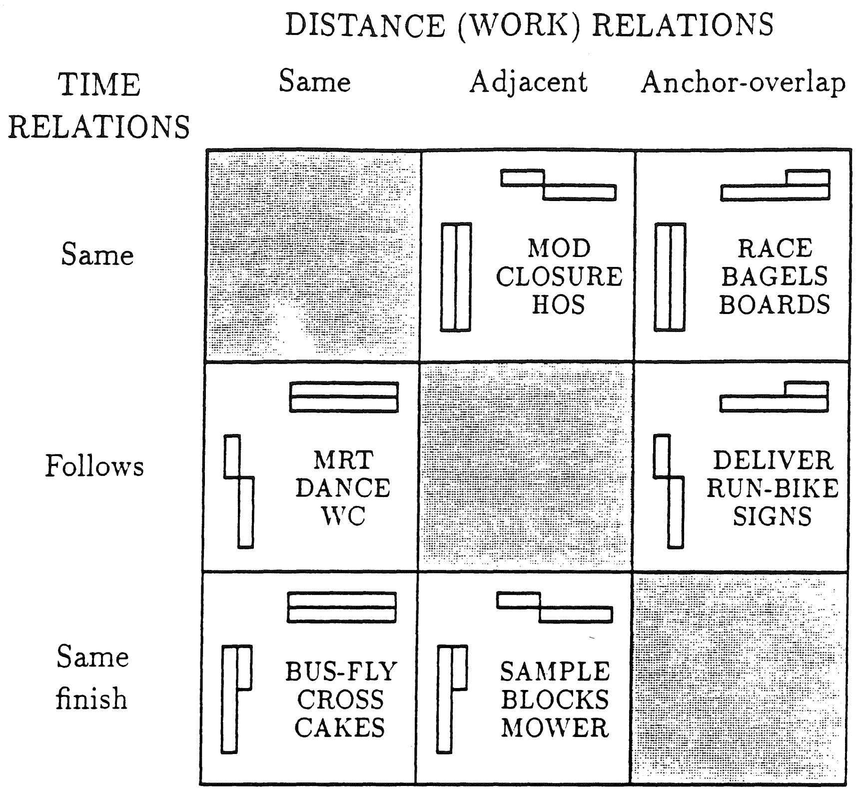


Figure 4.1: Algebra story problems constructed within a problem matrix that crosses selected output and time relations.

later time. We can generate other problems in this relational space by choosing different output and/or time relations, as in problem MRT where round trip distances are the same but travel times are sequential.

The problems used in this study were generated from the relational structures found by crossing same, adjacent, and anchor–overlap relations between distances or work outputs with same, sequential, and same finish relations between times. Excluding problem structures with identical relations in distance (or output) and time, this gives the six off–diagonal problem classes shown as a matrix in Figure 4.1. In each class, three problems were constructed, either directional variations of motion (e.g., separation or closure at the same time) or a work problem (e.g., working together at the same time). The problem labels shown in Figure 4.1 are used to identify problems within the chapter, individual problems are reproduced as needed in the text, and the full set of problems are shown in Appendix A.

Since participants were interviewed for different lengths of time (described below), each worked on a different subset of the problems shown in Figure 4.1. From the resulting transcripts, solution attempts on a set of "reference problems" (MOD, HOS, MRT, and WC) provide a basis for comparison with advanced undergraduates, while solution attempts on a larger set of "comparison problems" (including RACE and BAGELS) provide a basis for comparing students and teachers. Celeste was the only participant who attempted to solve problems from each structural class.

4.2.3 Procedure

Participants were interviewed individually, asked to work through a common series of problems, and encouraged to "think out loud" as they worked. As possible, problem solving interviews were conducted in the settings where participants usually worked on algebra problems, either at home or in the classroom. Interview sessions lasted from one and one–half to two hours, and participants were interviewed either once (one teacher), twice (one teacher and one student), or four times (one student). Approximately fifteen hours of audiotaped interviews were transcribed for further analysis.

Settings. The work settings in which verbal reports were collected differed from the classroom examination format of the written protocol study. During the interviews, participants were allowed access to earlier work, and at times the interviewer encouraged them to look back over earlier solution attempts, as described in a moment. There were no time constraints placed on the solution of individual problems, and several of the problems elicited lengthy solution attempts. In contrast, undergraduate students in the written protocol study worked through a bound problem booklet, were interrupted and told to move to the next problem at 8 minute intervals, and were never encouraged to consult earlier problems. Finally, the interpersonal character of the verbal protocol sessions, attended by a generally benevolent observer, differed dramatically from the anonymous examination format of the written protocol study.

These differences provide a meaningful contrast between problem–solving as a form of assessment in the classroom and problem–solving as a form of work, carried out at home (on beds, desks, and kitchen tables) or at a blackboard at the front of a classroom. The presence of an interviewer changes the character of this mathematical work, which is usually private and can be "cleaned up" later if need be, but participants acclimate quickly to this change, even taking the interview as an occasion to relieve themselves of their misgivings about formal algebra instruction.

Interview format. Problem–solving interviews were structured as "revised clinical interviews" (Ginsburg, Kossan, Schwartz, and Swanson, 1983) by treating the work surface and participants' written notations as concrete problem–solving supports. Thus, participants were asked to "talk aloud" while working on problems; they were also able to organize their work setting as they liked; and the materials they generated were taken up as part of the interview process.

Several forms of intervention by the interviewer result in protocols that differ from traditional studies of verbalization during problem solving (Ericcson and Simon, 1984). As is customary, the interviewer prompts the problem solver to "talk" or "tell me what you're thinking" after prolonged silence. However, the interviewer also asks for clarification when what the problem solver means by some verbal or written material is puzzling. Beyond solving presented problems, participants were also explicitly asked to recall earlier problems (generally at the start of the second interview) and were periodically prompted to compare solved problems with earlier problems. Thus, participants were indirectly alerted to the possible utility of problem comparisons during the interview, although explicit requests for problem comparison were generally post hoc.

More direct interventions were used when problem solvers reached an impasse or appeared lost. For example, a respondent might state "I really don't know" or enter a prolonged period of indecision over alternative conceptions of problem constraints. In these cases, the interviewer identified possible misconceptions and suggested strategic or tactical alternatives. Direct interventions were least frequent with Richard (an algebra teacher) but were undertaken with equal frequency across the other participants. Sometimes these interventions allowed a problem solver to continue what would likely have been an unsuccessful or abandoned solution attempt. Finally, the interviewer presented short tutorial sessions on relationships between various problem–solving representations and problem structure to both groups. These interventions were sometimes didactic, but generally led to collaborative representational constructions by interviewer and participant. Prompts for explanation, suggestions about alternative strategies, and tutorial interactions result in interviews that sometimes resemble a tutorial dialogue more than a traditional think–aloud protocol study. Thus, in addition to relatively uninterrupted problem–solving sessions, the interviews include significant instructional content and discussion of participants' problem solving activities.

4.2.4 Conventions for presenting transcript material

Any qualitative analysis of ongoing human activity must structure the observational record, risking unwarranted generalizations as ambiguous or seemingly unrelated observations are deleted from the interpretation. In some respects, these deletions are a routine part of scientific work (Star, 1983) that are required if the complex scope of observations will be given any coherent interpretation. This chapter attempts to balance these tensions by liberally including temporally contiguous excerpts from interview transcripts, together with written protocols taken during the interview. Thus, analytic distinctions are not only exemplified in the observational materials presented, but that material is presented in enough detail so that a reader can assess whether analytic categories are responsible to the ongoing activities of problem solving. One test for whether this balance has been achieved is whether or not a reader could construct alternative interpretations of the problem–solving setting and subjects' activities.

Several notational conventions are used in presenting verbal transcripts: ellipses (...) indicate silent periods of three or more seconds; parenthetical elements note descriptive or timing information made during transcription from audiotape; capitalized words show EMPHATIC tone; and bracketed elements [...] indicate omitted material. Utterances are numbered within interviews, both for reference in the text and to provide relative timing information. Written material shown in figures has been redrawn for clarity in reproduction, with care taken to preserve its relative position and local topology.

4.2.5 Analytic framework

The analysis of verbal and written protocol material in this chapter starts with problem MOD, the first problem seen by each participant. This problem provides a common reference point for the empirical findings in Chapter 3, and analyses presented here follow the same interpretive framework by dividing solution attempts into episodes of coherent problem–solving activity. A persistent methodological problem with protocol materials is that they are difficult to analyze in ways that allow one to use transcripts as more than exemplary materials (Anderson, 1987; Seifert and Norman, 1987). As a way of managing verbal and written interview material, the interpretive framework of problem–solving episodes acts as a relatively neutral organizational overlay that can later serve as an analytic index for qualitative and quantitative comparisons. This is similar to Newell and Simon's (1972) use of problem–behavior graphs, but the interpretive categories do not depend on any particular theory of problem solving.

Thus, we can look for replications of the episodic patterns examined in the study of competent problem solvers (e.g., model–based recovery from conceptual errors) within a much richer sample of problem–solving activities that correlates verbal reports with written notations. In addition, we can compare beginning and career algebraists with advanced undergraduates to determine if significant activity outside the algebraic formalism is used by these participants as well. I call algebra teachers "career algebraists" because they are neither professional mathematicians nor institutionally trained to be competent mathematical problem solvers in the sense of computer science or engineering undergraduates. Instead, they are trained to introduce newcomers to algebra in a way that may lead to institutional certification, so in a very particular sense they have made algebra (and other areas of mathematics) their career.

A detailed analysis of solution attempts on problem MOD not only relates to hypotheses advanced in the earlier study, but these analyses also generate several new theoretical categories that help to explain problem–solving tactics, content, and outcomes. These can be broadly previewed as (a) the transparency of written notations for supporting inferences about problem structure, (b) the role of these notations in maintaining control over quantitative manipulation, and (c) the personal epistemological position of different participants towards various written notations in algebraic problem solving. These are taken up with supporting protocol material in the next chapter, which sets out an explanatory account of competent quantitative inference.

4.3 Moving in opposite directions

(MOD) Two trains leave the same station at the same time. They travel in opposite directions. One train travels 60 km/h and the other 100 km/h. In how many hours will they be 880 km apart?

As mentioned above, problem MOD provides a reference point for findings and hypotheses presented in the preceding chapter. Each problem–solving interview starts with this problem, before any significant interaction between participant and interviewer. These solution attempts provide a baseline against which to examine the effect of various instructional interventions, either planned or spontaneously undertaken in response to the course of events within an individual interview. For example, with the beginning student (Celeste), there is evidence for significant learning events over the course of several interviews, and these can be understood against the tactics, episodic transitions, and written materials of her initial problem–solving activities.

Qualitative analyses of solution attempts on this problem are shown separately for each participant. Each follows the interpretive framework developed earlier, treating the solution attempt as a sequence of tactical episodes with content, errors, and outcomes that either finish the solution attempt or transition to another episode. Embedded in these analyses are references to the outstanding questions that opened this chapter, and these are collected in the discussion section that ends this chapter.

4.3.1 Celeste: Model–based inference as the "weird way out."

Celeste starts the interview with the first of many model–based reasoning episodes, using iterative simulation as a "weird way out" of the quantitative dilemma posed by uniform rates related under given constraints (S–3, below). Her written protocol for problem MOD is shown in Figure 4.2. Celeste's reflective comments about model–based iterative simulation are a revealing summary of the tension between mathematical activities that are familiar for her but illegitimate by contrast with the precision of prescribed ways of making sense of a school math problem. Her tactic is certain, in that she can determine an unknown value within her understanding of the problem's structure. However, as a public report, the tactic is a "guess" and she knows "there's some other way to do it." Although her use of a model of the problem situation is not yet well integrated with prescribed quantitative activities, she is generally able to find solutions to these problems without introducing algebraic notation.

S–1 They travel in opposite directions... OK, if there's one train... they're going different directions... so say this one is going 60 and this one's going 100 (gestures some distance apart).

I–1 So they are yeah far apart, on either end?

S–2 Yeah, is this per hour?

I–2 Yes, those are kilometers per hour.

S–3 So, I try to... I guess I take the weird way out. I would, I guess... add like 100 to that (indicates 100 kph) because that'll make two hours and add 60 to that (indicates 60 kph) to make two hours for that. I know there's some other way to do it.

I–3 So you'd see what happened in two hours? Why don't you write this down – don't worry about making mistakes.

S–4 I'd probably go 120 2 hours... would be 120 and 200 (writes "2 hrs 120 200" in a row)... and then times by 2 again... and see what that got and that would be 240 and 400. So far this is 640 and we need... more... 880. So... this would be 3 hours and then (writes "3 hrs 2 2" in a row)... I'd say that if you did it again, it'd be too much cause you'd have 800 and 480 and that's too much (nothing written). So I'd take like the half way out.

I–4 Half way of what? Or half way between what?

S–5 Divide this by two (indicates 240) and this by two (indicates 400)... but I don't know where that would get me because you're just going back to where you started.

Celeste sets up a model–based tactic by gesturally enacting the first hour of travel (S–1). Her narrative account of rate as an invariant capacity of trains (i.e., "going 60") is converted into a concrete product for the first hour of travel. The first arithmetic operation comes after gestural enactment, when she constructs the second state by adding concrete rate values to find a cumulative distance apart (S–3). This is an interesting transition from a model of train travel as a horizontal movement of her hands apart, to a narrated calculation that accumulates distances travelled after 2 hrs. Both sets of activities yield problem states in which trains are some distance apart after a common interval of time, yet two hours into the train trip Celeste's paper is unmarked.

At I–3, the interviewer suggests writing down her work, and Celeste begins working on her own paper (shown in Figure 4.2). Celeste's initial calculations become a table that starts with distances at 2 hrs and then stacks columnar multiplications to generate values for successive states (S–4). Common times are recorded at the left margin of a row, while distances for each train are organized under columns. She records distances for the second hour, but then reports "times by 2 again" as a repetition of the calculations taking her between successive states. Celeste uses this doubling operation to find that the trains are 640 km apart after the third hour,[1] a distance that should

[1]Figure 4.2 shows "4 rs" as a row label, reflecting a later repair by Celeste, who originally wrote "3 hrs" directly under "2 hrs" in a row with common multiplicative factors (i.e., "3 hrs 2 2").

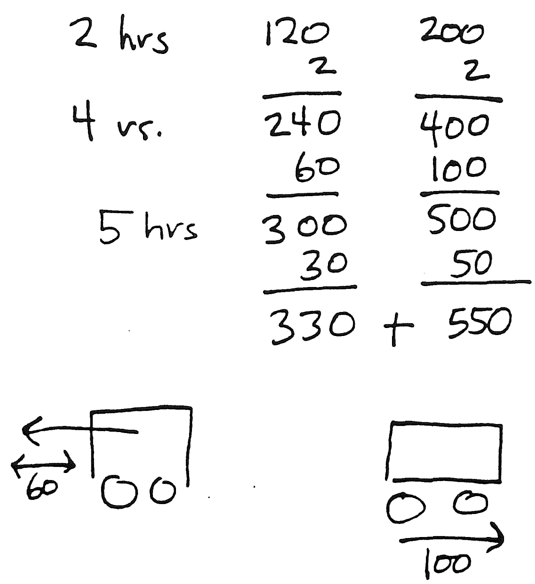


Figure 4.2: Celeste's solution to problem MOD.

actually occur after the fourth hour of travel. Considering the next (her fourth) hour, she narrates doubling prior distances to get 800 and 480 and remarks "that's too much." Attempting to recover, Celeste selects "the half way out" and halves excessive distances, only to find that she returns to the prior state and is "going back to where you started" (S–5).

An impasse in model–based iterative simulation. Interpreting these materials as a problem–solving episode, Celeste reaches an impasse in model–based iterative simulation used as a solution attempt. She has correctly inferred that travel times are the same (and equal valued), that adjacent distances can be added together to find the total distance between trains, and that the sum of these cumulative distances will eventually equal 880 km. Her difficulty, which can be interpreted as an error of commission for rate as a relation between time and distance, comes from incrementing hours travelled by 1 hr units at successive states while doubling the distances trains travel after each hour.

What does Celeste's conceptual error tell us about her understanding of rate as a linear function? From a formal view of the problem, is something wrong with Celeste's understanding of train separation as an instance of the mathematical concept of related linear functions? To preserve linearity under composition of these functions, any transformation in one dimension (i.e., multiplying distances by two) must have an equivalent transformation in the other dimension (i.e., multiplying times by two), a principle violated by Celeste's doubling transition between successive states. She may not yet correctly integrate related dimensions in her concrete conception of a linear rate, perhaps even omitting rate as a relational constraint across dimensions. Although she almost certainly is not thinking of train travel in this way, a literal reading of her calculations might even suggest that she is simulating successive states in a model of nonlinear acceleration (e.g., $D = R \times 2^{T-1}$).

Alternately, Celeste's difficulties may not lie solely within the mathematical concepts implicit in the problem's structure. The contested row in her table of state calculations is intended to be about the third hour of travel, perhaps independent of formal principles of rate as a linear function. Celeste is clearly using some relation across dimensions, because calculations are subordinate to and driven by her intention to construct successive states in the model. Moving from gestures that enact the model's state at one hour to a written record of the second hour, Celeste replaces a correct constructive activity (adding successive concrete rate values) with an incorrect inference about the quantitative relation between successive states (doubling distances from the prior state). Arithmetic operations that are originally embedded in a correct conception of problem structure implicitly diverge from what Celeste understands the calculations to be about.

The structure of the table as an external record of simulated states may partly explain her use of doubling to construct successive states. Noticing that train distances double in value from the first to second hours, Celeste constructs the third hour of travel by doubling distances at the second hour. Written as a table, prior calculations act as an external template for subsequent state constructions and allow Celeste to flexibly extend the written record of her simulation down the page. However, the two dimensional structure of rows and columns in her table provides little explicit feedback about quantitative operations or their results. State construction clearly motivates Celeste's calculations and provides meaningful points of reference for derived quantities (e.g., Celeste never adds times and distances), but her control over these calculations relies on a notation that does little more than record prior results. Interpreting calculation and its written record as tools for organizing simulated states, Celeste has made an unfortunate choice of tools in this solution attempt.

State–driven repair and solution within the model. Treating Celeste's complaint (S–5) as a genuine impasse, the interviewer recommends that she reconsider the third hour of travel and prompts her to explain the values she finds for time elapsed and distance travelled in this state (not shown). After considering how far trains travel in a single hour, Celeste detects the doubling error and replaces 3 hrs with 4 hrs as a row label in her table of state calculations. She also manages to synchronize time and distance in the iterative simulation, and when asked to recount the first hour of travel (I–13, below), she spontaneously resumes state simulation at the fifth hour (S–15), constructs another state at 5 and 1/2 hours, and finds the solution. Celeste's change from columnar multiplication at 4 hours to columnar addition at 5 hours shows that she has repaired her earlier error of doubling successive state distances.

I–13 Just to start it off, show me what would happen in 1 hour.

S–13 In 1 hour it'd be 120... no, no, wait... it'd be 60.

I–14 One train would go 60...

S–14 And the other 100. And then in another hour... yeh.

I–15 It would be another 60 and another 100. And that's how you get 120 and 200.

S–15 Ok, ok, so you'd go plus 60 and that'd be 500 and 300 and that's 800.

I–16 How many hours is that?

S–16 5. Yeh, and then... wouldn't it have to be like, 5 1/2 hours or something? Because if you wanted to do this (5 hour row) by half or something... these two by half?

I–17 So in a half an hour, how far would the 60 mph train go?

S–17 That'd be 30 and... um 50. Hey! Then I got it! Because you add 330 and 550 and you... and that's 880. Yeh.

A retrospective account of problem structure. Celeste's table organizes states as successive rows with like dimensional quantities arranged under columns. Paradoxically, the table shows very little about relations between states that would help to discover her conceptual error, but at the same time the table provides a natural structure for controlling the simulation by extending new rows down the page. Since the table plays such a mixed role in her performance, the interviewer asks Celeste to draw a picture of what is going on in the problem (I–21). She draws a literal scene directly below her tabular record of states in Figure 4.2.

I–21 Could you draw a picture of this thing?

S–21 Yeh, so here's the train... with little wheels and stuff.. and he's going this way... 60 here... and this guy is going 100 and you wanna find out... If you switched the problem around and you say, how far will it go in 5 1/2 hours, then you could like go backwards.

I–22 Backwards, you mean you could solve it backwards?

S–22 Yeh, you could do it... You know that this is 60 and you say that in one hour... and in two hours it'd be 120. I guess it would be easier if you said that... in 5 hours because you know how much it is in one hour. So if you just go 1 and 1 and 2 and 2 and 1/2, then that's five.

I–23 Do you often solve problems in this way? (she nods) You do?

S–23 Yeh, sometimes I try to say... It used to be hard for me to picture that in my head... a train actually doing that... it'd just confuse me because it doesn't seem like numbers and trains go together. But if they all have something in common then that's a lot easier for me to do.

I–24 So, what do they have in common here?

S–24 They're going in opposite directions for one thing. And one is going faster than the other, but they both have kilometer per hour. So you know that in some time both of them are going to be this far apart from each other. So they're gonna be like... just like moving your arms out.. and they're gonna be a certain distance apart from each other and that's gotta be something that has to do... that can be evenly into that.

I–25 Into what?

S–25 Into 880, both of them because they have to be some distance apart. This one goes 60 kilometers per hour and this one goes 100 kilometers per hour, and then eventually they're gonna get there somehow. And since you know the hour and the distance that they need to be, you can check your answer and everything.

Although Celeste's diagram shows a very literal account of train travel without significant rendering of state information or boundary constraints on train travel, she nevertheless gives an insightful verbal summary of problem structure.

- At S–21 Celeste changes the format of the problem by giving time (5 1/2 hrs) and working "backwards" through the rate to find how far apart the trains would end up. At S–22 she appears to scale the rate to 5 hrs, saying it "would be easier" than an iterative simulation.
- At S–23 Celeste remarks on the difficulty of learning to think about situations and quantities as having "something in common." In this sense, algebra story problems are a disruption in more prevalent school–math activities like performing calculations or recalling formulas.
- At S–24 what makes "numbers and trains go together" is a justification for various quantitative relations in the problem, expressed both verbally and gesturally. These provide a retrospective justification for various quantitative inferences that Celeste made during the solution attempt.

With a solution in hand, Celeste gives clear evidence of understanding the multiplicative relation between distance, time, and rate by exchanging given and unknown values in the problem statement. Her doubling error during iterative construction of states in a model of train separation, initially hidden in a two dimensional table, cannot be explained by a lack of knowledge about this multiplicative relation or by an understanding of train travel as being accelerated. Instead, her doubling error originates at the transition between gestural and written state simulation, leads to an impasse between halving and doubling as symmetrical calculations, and is detected and repaired during a prompted examination of travel during a single state.

The anatomy of model–based state simulation. Looking back to outstanding issues that motivate these verbal interviews, the first solution attempt collected during this study both corroborates earlier descriptions of the strategic role of model tactics and presents new evidence for the internal structure and content of these episodes. The origin of Celeste's solution attempt is a state construction that transforms rate into a concrete product. The narrative, gestural, or written activities producing subsequent states literally have a structure in space and in time (i.e., moving hands apart, writing a next row, making another calculation). These activities partly correspond to the dimensional structure of quantities in the problem by serially accumulating both distances and times, although this correspondence is critically dependent upon the notation that Celeste chooses — a tabular arrangement of written calculations.

Relations between problem, activity, and setting — events in the "story" presented by the problem statement, quantities given in the statement, activities in the spatio–temporal setting of the interview, and changes to that setting as written notations are introduced — are all integral to a complex human performance of "model–based reasoning." After limited intervention, the tactic is successful without algebraic notation, and this reflects a personal "way out" of a dilemma that Celeste experiences at several levels: managing residual quantities, finding a precise solution to this problem, and making public a "weird way" of dealing with school algebra. The qualitative analysis shows that (a) model tactics do occur among newcomers to algebra and that (b) these tactics have a characteristic structure, content, and meaning for their users.

4.3.2 Karen: Working between diagrams and recalled formulas

In contrast with Celeste's relatively exclusive use of model tactics, Karen's solution attempt on problem MOD moves between diverse notations and tactics in an attempt to construct a suitable algebraic expression. At S–2 (below), Karen's reading of the problem text is interleaved with construction of a diagram, shown at the top of her written protocol in Figure 4.3. Opposite–directed segments emanate from a common origin ("we've got the station"), drawn as a vertical line segment. These segments are labelled with quantities that are narrated as rates ("he's going 60") and letters that differentiate trains.

Model–based evaluation of an algebraic conjecture. Having constructed the diagram, Karen reports that her annotation serves primarily as an occasion for recalling algebraic formulas that might be suitable for this problem (i.e., "God, I don't remember formulas or anything for this."). At S–3, however, she recalls the three term multiplicative relation between time and distance and considers trying to find the time required for a single train to travel the given distance (i.e., $880/60 = T$). Before writing or manipulating an expression, Karen rejects this algebraic conjecture by comparison with her earlier diagram, which she updates with boundaries drawn as large parens to designate the given distance apart (i.e., "Oh, but they have to be 880 apart, like that"). This interpretation of Karen's activities is corroborated by her retrospective comments at S–4 (i.e., "... that was wrong because they both have to be this much apart, 880 apart.").

S–2 Two trains leave the station. So we've got the station. They travel in opposite directions. Say train A goes that direction (draws right directed segment) and he's going 60 (labels 60 km). B goes that direction... opposite, yeh (draws left directed segment). That's train B (labels 100 km). How many hours... will they be 880... God, I don't remember formulas or anything for this.

I–2 You think there's a formula that would help?

S–3 Yeh, I'm trying to think of that. I don't remember it, for setting that up. 60 kilometers in an hour... 880 kilometers... Oh, but they both have to be 880 apart, like that (draws large parens and labels 880). Oh, Rogers! (both laugh)

I–3 One formula that might help is distance equals rate times time. Did you already know that?

S–4 Yeh, I'm trying to figure out how to plug that into... what's stumping me is that... first I was thinking to take the 880 and figure out how long it would take train A to go that distance. But that was wrong because they both have to be this much apart (indicates large parens), 880 apart.

I–4 Ok.

Interpreted as a series of problem–solving episodes, this excerpt shows an interesting interaction between annotation, algebra, and model tactics. Karen's initial annotation uses a diagram to integrate event boundaries with given quantities from the problem statement, but she does not introduce significant structural constraints. A conceptual error of omission (the faster train's trip) is introduced when she pursues a recalled formula relating distance, rate, and time. This algebraic approach is rejected when Karen returns to her earlier diagram and extends its structure to render the final state in train separation, converting a relatively weak annotation into a diagrammatic model that shows the part/whole composition of distances as an important structural constraint.

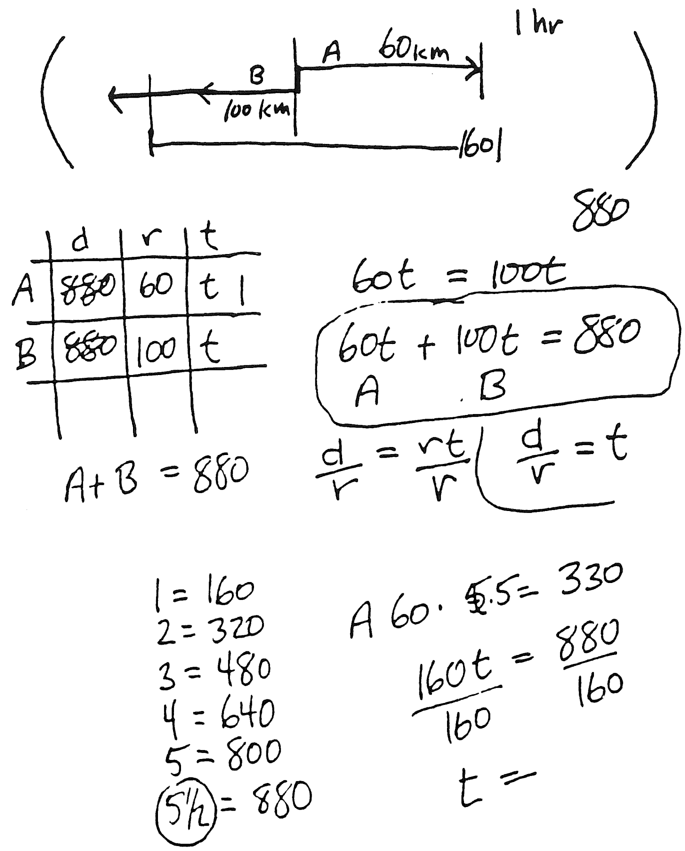


Figure 4.3: Karen's solution to problem MOD.

Doing the dirt: filling holes and making expressions. Karen has yet to assemble an effective set of constraints on the unknown time, but she has recovered from a conceptual error and constructed or recalled two important constraints: individual distances as parts of a given total and rate as a relation between time and distance. She is "stumped" (S–4) by being unable to recall a formula that integrates composite distances or to "plug" the problem into the simple multiplicative relation over distance, rate, and time that she has recalled. At S–5 she transitions to another written notation, a "dirt table" organizing quantitative roles (middle left in Figure 4.3). As conventionally presented in algebra texts, the table provides an organized array into which given and unknown quantities can be placed. Within the interpretive framework of the preceding chapter, this excerpt could be interpreted as the beginning of an algebraic episode, since Karen is constructing and instantiating equations around the role table. Alternately, her recall and use of the role table might be interpreted as an instance of a new tactical category, using a specialized notation for constructing algebraic expressions.

S–5 Ok, well let me set up my little dirt table... I call it dirt... R T (draws a 2 by 3 table, labels d, r, t for columns). So we've got train A and train B (labels A and B for rows). Ok, the rate is... for train A is 60. Is that right? Yeh, rate... so I'm trying to figure out the time. ... The distance... is 880 for both of them. Well, yeh, its A plus B equals 880. That's right... well, yeh. (writes 880 in each cell under d) That seems wrong to put 880 in there, though.

I–5 Where would you put it?

S–6 In the distance box for either of those. This to me seems wrong. Ok, let me see... One thing is A plus B is *going* to equal 880 (writes A + B = 880 below table). So...

I–6 What you thinking?

S–7 Well, this is screwing me up. It may be right, but I'm feeling that its wrong to plug that distance in there. So, maybe there are 2 equations... so we can put a t here (writes t in time cells for A and B). 60 t equals 100 (writes 60t = 100t)... no it doesn't equal 880, it equals part of 880. Oh, oh, its coming to me! Oh, good. 60 t equals... no that's not right. That rate... that rate PLUS that rate... 60 t... plus 100t equals 880 (writes 60t + 100t = 880). Is that right? I'm getting away from that formula, though. Rate times time... (long pause)

The "dirt" table directs Karen's attention to typed quantities for each train, but as a notational device leaves the relations between table quantities largely implicit by giving no evidence for whether or how various table cells might be quantitatively combined. After correctly inserting rates in rows A and B, she narrates a choice between treating distances as being equal ("880 for both of them") or as an additive composition ("its A plus B equals 880"). Karen incorrectly inserts 880 into distance cells for each train (S–5), producing a table that is at odds with her earlier diagram. Puzzled over this inconsistency, Karen correctly inserts a common variable in each time cell of the table, but she then generates an incorrect algebraic expression equating train distances (60t = 100t). At S–7, she notes that each train travels a part of the combined distance, a structural constraint used in her earlier diagram to reject an incorrect quantitative conjecture (S–4) and generates a correct algebraic expression (60t + 100t = 880). However, Karen worries that this sum of products is more complex than the three term relation she has recalled between distance and time ("I'm getting away from that formula, though.").

Reminded of using multiple equations in her algebra course, Karen abandons the correct algebraic expression and begins several unsuccessful attempts to find two equations using only the three term relation between distance, rate, and time. After several minutes of algebraic manipulation (not shown), she reports that these activities lead back to an expression she previously rejected, 60t = 100t. The interviewer intervenes by asking about the algebraic expression she abandoned earlier, 60t + 100t = 880.

I–19 So what else do you need other than this equation? (indicates 60t + 100t = 880)

S–20 (long pause) Oh! I need... Ok, I'm thinking I need something to go in here other than the t, to be able to times it by that to get the... Well? Ahh...

I–20 What do you know about their times?

S–21 Well, let's see. He travels 60 kilometers per hour... (long pause) Oh! Well time is 1 (writes 1 in time cell for train A)... No.

I–21 So in one hour...

S–22 In one hour, he goes 60 kilometers, and in one hour he goes 100 kilometers. So then that would just be 60 plus 100, and that's not 880 (laughs). Wait a minute, 60 times 100, no that's not it either. I'm confused.

I–22 In one hour, what would happen?

S–23 Well, yeh. He'd have gone 60 and he'd have gone... that far in one hour (extends segment for B and draws a composite brace below A and B).

Karen again fails to recognize that the complex sum of products can be manipulated to determine a value for the unknown time. Asked to clarify the relation between times, she narrates the rate ("60 kilometers per hour") and shifts from thinking about this quantity as an invariant property of trains to a concrete description of activity, as evident in her statement "Well time is 1" (S–21). Karen also records this value in the upper time cell of her table, signalling a dramatic change in her view of quantities in this problem.

From equations to states. Asked to explain this inference (I–21), Karen searches for an arithmetic operation that relates 60, 100, and 880 and then concludes that the trains will be 160 km apart after 1 hr. This state construction leads her to return to her earlier diagram, redrawing the segment for the faster train (B) so it is longer than the slower train and recording the sum of partial distances with a brace directly underneath the original directed segments (top of Figure 4.3). In contrast with much of her preceding algebraic activity, Karen again has a set of familiar constraints (literally) in hand: times are the same, distances are added together, and this sum of distances must eventually equal 880 km. When the interviewer restates the required unknown (not shown), Karen realizes that there is a "weird way" out of her algebraic impasse, much as in Celeste's solution to this problem.

S–26 Allright, so in one hour, we've got 160 kilometers (writes 1 = 160). Well, I could just keep doubling it (laughs). That seems like a weird way to do it. Or not doubling, but...

I–26 What would happen if you did that?

S–27 Well, then in 2 hours, right (writes 2 = 320)... Is that how it would work? Can I use this (calculator)?

I–27 Sure.

S–28 Is that right, though, to do it times 2. So in 2 hours they've gone 320. In 3 hours... but I'm wondering if I'm keeping the same... 160 times 3 hours... 480.

I–28 What are you worried about, with that?

S–29 Oh, no. I'm just wondering if I'm plugging the right thing in. If I keep...

I–29 If after each hour you should plug in 160?

S–30 Yeh, well its 160, and then... yeh, 160 keeps... right? Yeh. 4 times... is 640... I'll do this til I hit 880! (laughs) If I can't figure out the damn formula... Ok, so 160 times 5... I just realized what I could do. Ok, so after 5 hours they go 800 miles... kilometers. And you've got 80 more to worry about. Now what? Um...

[... exchange about efficient use of calculator ...]

I–30 So you've got 80 miles to go...

S–31 Ok, and he can go... ok, so... yeh, which is half of 160, so its just another half an hour. No, wait... That doesn't make sense. What am I thinking? Yeh, well in a half an hour... Well, I'm thinking half of 160, which is what they do in an hour... they go together in 1 hour. Um... in half an hour they do 80 kilometers, 80 and 80 is 160. Ok, so... 5 hours plus a half would give me 880.

I–31 Ok.

S–32 So in 5 and 1/2 hours, it would take them 5 and 1/2 hours to reach 880 kilometers apart. Now why can't I figure out how to put that into the formula? Or the damn table? That's weird. (writes 5 1/2 = 880)

After settling on "how it would work," Karen iteratively constructs states for successive hours of travel, continuing until she is within 80 kilometers of the given distance apart (S–30). Noticing that this is half of the combined distance per hour, she chooses a half hour to complete the simulation by finding a value for the unknown time. Ironically, despite having just enacted a model of the same complex expression she did not recognize as determining the unknown time (60t + 100t = 880), she does not see how her table of state values relates to either the formula she has recalled (d = r × t) or the role table that she has constructed with some difficulty (S–32).

Abandoning interleaved annotations ("dirt table") and algebraic tactics, Karen succeeds in a solution attempt using model–based iterative simulation. Within structural constraints that travel times are the same and distances are combined to give 880 km, she unpacks the joint multiplicative relation between distances, rates, and time into a series of state calculations, each carefully accountable to her sense of motion in opposite directions. Resulting values are recorded in a table (lower left in Figure 4.3) that organizes common times in the left column, the composite distance apart in the right column, and the functional relation between time and distance as an arithmetic sign for equality (i.e., "="). Under conventional arithmetic interpretation, her series of state expressions (i.e., "1 = 160" through "5 1/2 = 880") would be nonsensical. Thus, table entries are clearly not intended as arithmetic relations between quantities. Instead, Karen reads the final row as a state in a model of train separation: "it would take them 5 and 1/2 hours to reach 880 kilometers apart" (S–32).

Model tactics in action. The summary comments for Celeste are confirmed by a qualitative analysis of Karen's solution attempt: model tactics occur among newcomers, model tactics have a written and active structure that corresponds in important ways with the dimensional structure of related linear functions, and these tactics have a difficult personal significance for newcomers in the face of prescribed school mathematics (e.g., a "weird way to do it" at S–26). Karen's use of model tactics also corroborates hypotheses concerning their role in problem solving:

- *Generating quantitative constraints* — e.g., Karen decides to add concrete rates while updating her diagram to render distances after 1 hour as adjacent segments.
- *Determining unknown values* — e.g., Karen resolves her uncertainty over various algebraic expressions by finding a solution using iterative simulation.
- *Evaluating conjectured constraints* — e.g., Karen rejects an incorrect algebraic relation by comparison with an extended version of her initial diagram.

4.3.3 Paul: Old ways of getting to the new math

On his first problem (MOD), Paul moves quickly from a diagram showing the first hour of train travel into a solution attempt using a whole/part ratio (S–2 and Figure 4.4).

> S–2 Two trains leave... they travel in opposite... one train travels 60, one travels 100... In how many hours will they be... Ok, you start off from one starting spot (marks an X at origin of a line) and then you... this is one hour's, this is 100 (draws two collinear segments, labeled 100 and 60)... so they go 160 kilometers in 1 hour (writes 160 km/ below). How far will they

be in... 160 into 880 (written division), what's that? Divisione... Where's the calculator? So you do the new math...

I–3 (both laugh)

S–3 I don't waste my time with that stuff. New math... 5.5... hours between... to be that distance apart (circles 880 in written division).

Paul's diagram uses collinear, adjacent segments to show the progress of trains after one hour, and these segments are labelled with values that denote scalar quantities for distance travelled. On the basis of this partial simulation, he infers that train distances can be combined and constructs a concrete association between distance apart and a common interval of time ("they go 160 kilometers in 1 hour").

Using this composite scene, Paul first considers "how far will they be in..." some unstated amount of time. Though his diagram does not show any relation between the distance apart after one hour and the given constraint on total distance apart, he immediately transitions into a ratio tactic that partitions the given distance apart (880 km) into 160 km components for each hour of travel. Without explicitly introducing algebraic notation, he constructs a simple arithmetic calculation ("Divisione") that extends from a label on his diagram and yields a precise value for the required time. With this value in hand (S–3), he attempts to describe its place in the diagrammatic model as "hours between," but then explains the value by referring to "that distance apart" and circles 880 in his written notation for division. Ironically, there is nothing in his earlier diagram that can carry the result of his ratio solution.

While Paul transitions smoothly between partial simulation using a collinear segment diagram and a simple ratio calculation, his diagram is interesting in what it leaves out of the problem model: any explicit notation for the temporal dimension, any global constraint on the given distance apart (880 km), and any explicit notation for rate as a diagrammatic convention (e.g., labelling segments with intensive quantities). Asked to explain what his diagram is showing (I–4), Paul recapitulates the inference about combined distance by narrating interdependent constraints between places that bound segments, the quantities that these segments carry as labels, and the quantitative relation (addition) that the configuration of labelled segments implies (S–4). Again, a common travel time for trains is carried implicitly in his verbal report, but his explanation shows how a state, depicted as a spatial scene in the diagram, sanctions an important quantitative inference.

I–4 Ok, tell me what this diagram is showing.

S–4 I just start off with the origin, and then one goes 60 and one goes 100, so its one hour's time. So... they go 160... the distance from here to here is 60, and add to that... its 160. So its that distance. It will be 880 apart...

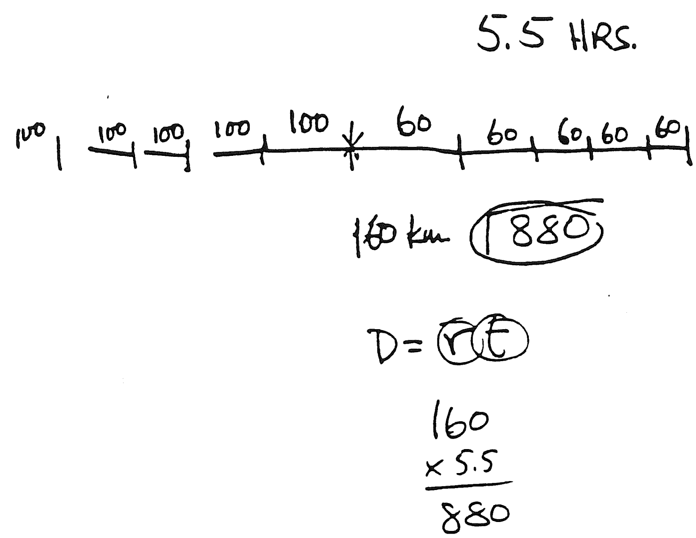


Figure 4.4: Paul's solution to problem MOD.

State simulation as an explanation for arithmetic operations. Given the explicit content of the diagram and Paul's explanation of its use, it would be interesting to see how he might complete the diagram, since he uses a single calculation to find a precise value for the unknown time. Accordingly, the interviewer asks Paul to show what the diagram would look like after 5 and 1/2 hrs of travel (I–5).

> I–5 So after 5 and 1/2 hours of doing that... Show me what it would look like after 5 and 1/2 hours. One of the things I'm interested in is how people use diagrams.
>
> S–5 Ok well this is essentially 1 hour, 2 3 4 5 (draws equal length segments for each)... so 5 and 1/2.
>
> I–6 Ok, they go on out.
>
> S–6 If that's right, I'd have to (checks number of segments)... half an hour.

Paul first reiterates that the existing scene depicts "essentially 1 hour," and then extends the initial diagram by drawing and labelling successive states constructed at 1 hr intervals. Since the interviewer's request is relatively open–ended, Paul's choice to conduct what amounts to an iterative simulation suggests (a) that his conception of

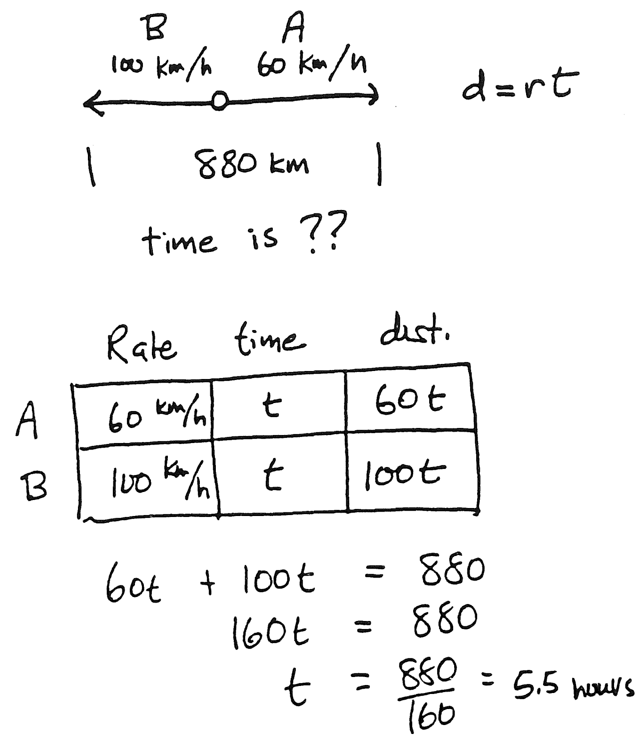


Figure 4.5: Richard's solution to problem MOD.

The diagram mixes potentially incompatible notation schemes, making its proper reading as an integrated collection of constraints problematic. Adjacent, opposite–directed segments terminate at the boundaries of an enclosing brace that is labelled with a distance value (880 km). The spatial configuration of the segments, combined with the brace label, suggests a composite relation between distances. However, labels on the enclosed segments include two additional notations that associate rates with particular trains (e.g., 100 km/hr for train B). Thus, in a single diagram, Richard uses three different notations to label diagram components: one dimensional quantities for distances apart in a retrospective view of train separation, two dimensional quantities for rates associated with train paths, and symbolic identifiers for trains associated with these paths.

To give the spatial configuration of segments and their labels a consistent interpretation, a reader of the diagram must sort through these different notational systems. Richard moves between these alternative interpretive readings smoothly, avoiding potential confusions that might disrupt a student trying to make sense of or repeat his performance. For example, upper case letters used as identifying labels for trains might easily be confused with algebraic variables for the value of distances travelled, perhaps encouraging the modifier/nominal shift reported as a frequent difficulty among newcomers to the algebraic formalism (Sims-Knight and Kaput, 1983).

"Dirt" tables taught. Having identified a collection of structural constraints on given and required quantities, Richard shifts to a "standard technique" for organizing the construction of algebraic expressions: a "box" or "window" divided into rows and columns. This notational device is the instructional version of Karen's "dirt table," used to organize and record inferences about relations between quantitative roles in the problem.

S–4 Otherwise you don't know what you're after. And the standard technique that I think most of us teach around here, is for distance rate time problems we... take a box, sort of a window, and divide it into three columns with as many rows as there are trips that are taking place (draws role table below diagram).

I–4 OK.

S–5 And there are two trips taking place, there's one... I'll call it the 60... well I think I'll call this vehicle A and this will be vehicle B over here (draws and labels rows). I forgot what they were, were they trains?

I–5 Yeh.

S–6 And then the columns are labelled rate, time, and distance (labels them). I'm not doing this now as though you were a student of mine, because I would be going much more slowly.

I–6 Sure, OK.

As in the episodic analysis of Karen's solution attempt, the table can be thought of as an annotational tactic, perhaps a further specialization of the more general tactical category that includes diagrams and notes about various problem elements. The strategic purpose of this episode can also be interpreted as comprehending the problem's structure, ultimately to generate algebraic expressions between given, inferred, and required quantities. By the end of the episode (shown below), Richard succeeds in identifying quantitative relationships that will support a single algebraic expression, and this signals transition into an algebraic episode.

While constructing this role table, Richard narrates general rules for constructing tables on problems that include multiple events, each involving related linear rates of travel (S–4). Rows record quantities for each event, while columns are labelled to identify types of quantities (i.e., "rate, time, and distance" at S–5). This construction appears to be a routine pedagogical activity for Richard, evidenced by his losing track of what types of vehicles were mentioned in the original problem text. Thus, at least as a public performance, Richard's solution resembles traditional accounts of competence in this task domain: retrieving, instantiating, and applying practiced problem–solving schemata (e.g., Hinsley *et al.*, 1977). A final note of interest as Richard sets up the table is his deliberation over whether to use rate values or alphabetic symbols as labels for rows (S–5). As with his labels for diagram components, either choice treats the label's content as a property or modifier by referring to events given in the problem text.

Organizing versus justifying quantitative inferences. At S–7, Richard sets out to fill in empty cells in the table, starting with "the givens" and then considering how to record a common time, implicit in his earlier textual annotation to the diagram. The role table effectively promotes individual quantities and their roles in the event structure of the problem, sorting given or inferred quantities into their respective dimensional types (i.e., rate, time, or distance) and story events (i.e., A and B trains). However, the role table provides no further constraints on what should fill its cells. Much as in Celeste's state table and Karen's role table, the two dimensional spatial array helps to organize structural inferences but does little to generate or sanction these inferences. For example, the multiplicative relation between distance, rate, and time is introduced later (S–9) under the assumption that students will have memorized this formula and be able to recall it as part of a table–building routine.

S–7 The rate is given... let me write in all the givens... the rate is 60 kilometers per hour (fills cell) and for the other, train B is a hundred (fills cell). For the time... well, one thing that we don't know is the time. But what we... this is the part they have so much trouble grasping, is that the time that it takes for each train to reach the point at which the total distance between them is 880 is the same for both trains. I don't know why I'm telling you this, because you know that (laughs).

I–7 No, go ahead.

S–8 If they leave at the same time, which I believe the problem said, and then at some specific time later they're 880 kilometers, then they must have travelled for the same amount of time. If you put two people on the train and started their watches at the same time the train left, then they're watches would say the same thing when you took the measurement between the trains.

I–8 OK.

S–9 So we'll call that time, whatever it is, we'll call it t, but its the same for both, that's why the letter is the same in both boxes (fills cells). And the distance travelled... OH! We need to understand the distance equals rate times time formula, which we would have typically previously studied (writes "d = rt" to right of diagram).

I–9 OK.

S–10 This may be a cop out, probably requires explanation, but the distance would be the rate, 60, times the time, t. In the second box it would be the rate, 100, times the time, t.

I–10 Uh, huh.

As he fills in the role table, Richard carefully justifies his choices for cell contents using model–based reasoning tactics. Consistent with the hypothesis that model tactics can generate or sanction quantitative inferences, Richard switches to a verbal account of places and intervals in a time dimension to justify "the part they (i.e., students) have so much trouble grasping" (S–7). At S–8, he explains the inference about common travel times as an interval bounded by common starting and finishing times, verbally constructing a model of time that is very similar to his earlier segment model for distance. This model of time is further elaborated as a scenario in which travellers on different trains synchronize their watches to measure a common travel time. Neither explanation uses an explicit written notation to show the time relation. Instead, as in his supporting scenario, time is literally carried by movement in an explicitly rendered distance dimension. Thus, his public explanation for filling role table cells appeals to inferences sanctioned outside the role table or algebraic formalism, narrated as a trip through places in a narrative model of train separation.

Reading tables as algebraic structures. Richard offers a final model–based explanation for inferring a part–whole relation that combines distance products (S–11), returning to his earlier diagram and reading its spatial configuration as an additive relation over distances. This relation forms the basis for constructing an algebraic expression that integrates distance cells in the role table, written in the conventional symbolic notation directly below the table. Under an episodic interpretation, Richard interleaves role table annotation with explanatory model–based reasoning before constructing a single expression that initiates an episode of algebraic reasoning.

S–11 Then, we can say, from the drawing, we can tell that when we add the two separate distances together, we get the total distance of 880 kilometers. So 60t plus 100t... equals 880. Solving this, and I don't think you're after the methods, the algorithms for solving simple linear equations, so I won't take time to do that. But t equals 880 divided by 160 or... oh, about just over 5. And this is something that's good too, is to estimate an answer before you go to the calculator so that you know you're in the right ballpark.

I–11 Yeh.

S–12 So it comes out to 5 and one half (using calculator). And to label the final answer and then to go back and see whether or not it makes any sense. If the first train, by that I mean the 60 kilometer per hour train, goes for... just roughly for 5 hours, let's say, then it will have gone... 60 times 5

> or about 300 kilometers, just over that. The other train, going at 100 for 5 hours, would go 500 kilometers, just over that. The sum of those two, 300 and 500 makes 800, just over that IS the distance that was given, so I would say the answer is probably right.

The resulting equation is manipulated to give a division operation ("880 divided by 60" at S–11), and Richard estimates a value for the unknown time before using a calculator to find the precise value. At S–12, he "goes back" in an explicit verification episode, estimating distances travelled by each train "roughly for 5 hours" and deciding that his precise answer is "probably right." Rather than simply plugging this answer into an earlier equation, Richard's verification carefully reiterates what various algebraic terms refer to in the problem situation, identifying quantities by their type and role in simultaneous travelling events.

Models as idealized problem–solving activity. Even during a public teaching performance, model tactics play a significant role in Richard's solution attempt. Combined with Paul's solution on problem MOD, model tactics appear prevalent across the competence continuum these analyses were designed to sample: algebra students, teachers, and advanced undergraduates all use model tactics in significant ways. Unlike newcomers, teachers tend to use these tactics to generate or evaluate quantitative inferences, but they do not generally find solutions using model tactics. In contrast with Paul's relatively open use of model tactics while solving the presented problem, Richard demonstrates similar tactics as something "they" (i.e., his students) might use. This also contrasts with students' views of model tactics as a "weird way out" of the demand for precision in the face of uncertain algebraic notations. Although Richard's personal position towards model tactics is less apparent.[2] as a relatively elaborate, his performance resembles what others have reported to be a common tutoring strategy: modelling the problem–solving activities of an "idealized novice" (Fox and Karen, 1988).

4.4 Comparison with advanced undergraduates

While a detailed qualitative analysis of a single problem is instructive, the verbal interviews provide a mass of data that can be analyzed quantitatively as well, similar to the quantitative analyses of episodic structure undertaken with written protocols reported in **Chapter 3**. This section uses similar quantitative analyses to contrast the solution attempts of algebra students, algebra teachers, and advanced undergraduates on a comparable set of algebra story problems.

[2]Richard solves problem MOD a second time, showing "what it is that I would have done" (S–14) In his personal solution, he imagines sitting in one train and looking in "my rear view mirror" at the other train, speeding away at 160 kph. Again, model tactics play a generative role in quantitative inference.

Problems MOD, HOS, MRT, and WC provide a reference point for comparing algebra students and teachers with mathematically sophisticated undergraduates studied in the previous chapter. These comparisons follow the earlier analyses of (a) the prevalence of interpretive categories within solution attempts, (b) relations between outcome, tactical content, and errors in final episodes, and (c) the tactical course of making and pursuing quantitative inferences. Problem HOS has a quantitative structure that is isomorphic to problem WT, which was used in the written protocol study. However, its content (hoses filling a pool) was changed to be more accessible than the abstract sense of "a job" in problem WT. Hence, these problems are not strictly comparable. Also, when comparing algebra students and teachers, analyses can be extended across another structural class of problems by including RACE and BAGELS (see Figure 4.1).

4.4.1 Interpretation of verbal transcripts

Verbal transcripts and written material are partitioned into problem–solving episodes, much as in the analysis of written protocols. An analysis of these more detailed solution attempts provides a useful framework for comparing the relative prevalence and significance of problem–solving tactics between students, teachers, and advanced undergraduates. However, since the interview format allows significant interventions on the part of the interviewer, the interpretive framework requires several extensions.

Interviewer interventions. First, a distinction can be made between episodes that participants undertake spontaneously and episodes that they undertake in response to an intervention. As mentioned in the Method section, interventions take several forms in these interviews: (a) prompts to talk aloud, (b) requests for clarification of problem entities or solution activities, (c) recommendations about these entities or activities, and (d) demonstrations of problem–solving activities. Beyond reminders to talk aloud, the more invasive interventions generally concern interpretive categories for tactics, purpose, constraints, errors, and outcomes described in the preceding chapter. Thus, the interviewer might ask a participant to explain the relation between times in problem MOD, might recommend that she treat these times as being the same, or might demonstrate the construction of a diagram for times and show its implications for a quantitative relation.

Interpretive categories and scoring. Interview transcripts were coded by the author using the interpretive framework developed in the preceding chapter. Since a single coder identifies and interprets episodic activity in these interview transcripts, no claim is advanced for reliability across observers. However, care was taken to use interpretive categories in a manner consistent with the scoring of written protocols. On one hand, the verbal transcript simplifies episodic interpretation by providing additional evidence for strategy, tactic, and content. On the other hand, the episodic structure of problem–solving activity appears considerably more complex when written materials are combined with participants' verbal reports.

4.4.2 Episodic structure of solution attempts

None of the solution attempts on reference problems involve demonstrations by the interviewer, but during the solution attempts of both students and teachers the interviewer recommends constraints (e.g., "he's gotta come all the way back"), strategies (e.g., "how would you check a problem like this?"), or particular notations (e.g., "is there a diagram that would go with that?"). Of 99 scored episodes on the reference problems, 24 were preceded by recommendations from the interviewer. The prevalence of interviewer recommendations varies across participants on reference problems but is not more common for students than for teachers. Episodes undertaken after a recommendation by the interviewer are excluded from comparisons of students and teachers with advanced undergraduates unless otherwise indicated.

Prevalence of episodes. There are more spontaneous episodes identified in verbal interviews with students and teachers than in written protocols from advanced undergraduates. Across reference problems, advanced undergraduates average 2.5 episodes per solution attempt, algebra teachers average 4.7 episodes, and algebra students average 6.5 episodes.

The greater prevalence of spontaneous problem–solving episodes has several possible explanations. Without an imposed time limit on the reference problems (8 minutes in the written protocol study), students averaged 12.4 minutes and teachers averaged 5.0 minutes to reach an unassisted final episode. At least for algebra students, longer solution attempts produce more problem–solving episodes. The requirements of talking aloud and giving explanations to the interviewer (some prompted) probably also contribute to a more complex episodic structure in participants' solution attempts. Finally, verbal transcripts show episodes in which participants produce no written material, they verbally elaborate relatively ambiguous written material, or they return to material generated during earlier episodes. None of these episodes could be detected in a written protocol alone. Removing the 8 minute time limit, requiring verbalization, and interpreting detailed narrative materials each help to account for the higher density of episodic activity found during solution attempts in verbal interviews.

These differences might be seen as an artifact of an interaction between the interview procedure and interpretive coding, with relatively little to say about the episodic structure of quantitative inference. On the other hand, the difference in setting is an intentional move along a theoretical sampling dimension, comparing problem solving among undergraduates in class, using a written examination format, with relatively unrestricted problem–solving activity among algebra students and teachers in settings where they usually work on such problems. Under the latter view, the difference in episodic complexity between participants in these verbal interviews and advanced undergraduates is meaningful: solution attempts often fit neither within the time limits nor within the material limits of conventional school or experimental assessments.

Problem–solving attempts. Among advanced undergraduates, the frequent presence of tactical episodes other than algebra was a surprising finding. Their solution attempts did not progress smoothly as a series of algebraic manipulations but instead contained conceptual difficulties that were overcome with a variety of tactical approaches. As in the written protocol study, there is ample evidence for episodic activity other than manipulating algebraic expressions within these verbal interviews. Since the interviews involve only four participants,[3] a comparison with advanced undergraduates can only speak to the presence of various interpretive categories among students and teachers. Table 4.1 collapses across problems to show the percentage of each participant's problem–solving episodes that contain a scored category.

In the verbal interviews, each student and teacher uses algebra at least once while working on the reference problems, but the relative frequency of algebraic tactics within each participant's solution attempts is small, ranging from 6.7% (1 of 15 episodes for Celeste) to 34.4% (11 of 32 episodes for Karen). Thus, for students and teachers alike, the majority of their problem–solving episodes do not explicitly use algebraic tactics. Instead, model–based reasoning is the most prevalent tactic for all but the intermediate algebra student (Karen), who more frequently uses tabular annotations to organize her construction of algebraic expressions (9 of her 12 annotations are "drt" and "trw" charts).

There are no unit or procedure tactics found among students and teachers. Unit tactics were rare among advanced undergraduates, so understandably might not be observed in verbal interviews with only three or four participants. Procedure tactics among advanced undergraduates were most frequently observed on problem WT, where participants found the time working together as an average over individual workers. In the verbal interviews, filling a pool together in problem HOS appears less likely to be interpreted as an average working time.

Like advanced undergraduates, each student and teacher attempts to find a solution on every problem, but episodes for comprehension generally occur most frequently. Again, explicit episodes of verification are infrequent, though each student and teacher makes at least one spontaneous attempt to verify an inferred constraint or solution while working on the reference problems. When episode transitions are collected together as being "on track" (solution or subgoal) versus "off track" (impasse, lost, or wrong) there is large contrast between Karen, who finishes 56.2% of her episodes off track, and Richard, who never transitions out of an episode off track. Celeste and Paul are intermediate and roughly equivalent, finishing 26.7% and 33.3% of their episodes off track, respectively. Looking more closely at problem–solving difficulties, Karen and

[3]Richard (a teacher) did not receive problems MRT or WC, so comparisons between algebra students and teachers rely on a single teacher's solution attempts for these problems (i.e., Paul).

Table 4.1: Percentage of participants' episodes with a scored category.

	Participant (number of episodes)			
Category	Celeste (15)	Karen (32)	Paul (18)	Richard (10)
Tactical content				
Algebra	6.7	34.4	33.3	30.0
Model	53.3	28.1	38.9	40.0
Ratio	26.7	0.0	22.2	10.0
Annotation	13.3	37.5	5.6	20.0
Strategic purpose				
Comprehension	53.3	81.2	44.4	70.0
Solution attempt	33.3	15.6	44.4	20.0
Verification	13.3	3.1	11.1	10.0
Episode transitions				
Solution	13.3	9.4	27.8	20.0
Subgoal	60.0	34.4	38.9	80.0
Impasse	20.0	12.5	16.7	0.0
Lost	6.7	15.6	0.0	0.0
Wrong	0.0	28.1	16.7	0.0
Errors				
Omission	6.7	25.0	22.2	0.0
Commission	6.7	28.1	16.7	0.0
Manipulation	20.0	0.0	5.6	0.0

Paul resemble advanced undergraduates by having more conceptual errors of omission or commission than manipulation errors. In contrast, Celeste's relatively frequent manipulation errors switch the referent of a quantity during arithmetic or algebraic manipulation, while Richard encounters neither conceptual nor manipulative errors.

In summary, the episodic content of problem solving (i.e., tactics, strategy, transitions and errors) among algebra students and teachers is roughly comparable to that found among advanced undergraduates. Nonalgebraic tactics play a significant role in the solution attempts of both teachers and students. Pooling across participants and problems, model-based reasoning tactics are most common (37.3%), followed by algebra (28%), various annotations[4] (22.7%), and ratio tactics (10.7%). Much of the episodic content of problem solving focuses on comprehending problem structure (65.3%) rather than attempting a solution (26.7%), and participants seldom explicitly verify a possible solution (8%).

Model tactics occured one or more times in the solution attempts of 73% of advanced undergraduates and were used by 22.4% to 47.1% of these problem solvers across different problems. In the verbal interview data on reference problems, every participant uses model tactics at least once on every problem, with the exception of Celeste (a student) on problem WC. A higher prevalence of model tactics among students and teachers probably results from the verbal transcript allowing detection of model episodes that do not produce a clear written record. This suggests that the analyses of written protocols in the preceding chapter underestimate the prevalence of model-based reasoning in competent problem solvers' solution attempts.

Final episodes: outcome, tactical content, and errors. Within their "final episodes," advanced undergraduates often found a solution without algebraic tactics. They failed to find a correct solution most often on problems MRT and WT, and their difficulties usually came from conceptual errors. A similar analysis of solution attempts from verbal interviews is possible, although four participants cannot provide a reliable basis for quantitative comparison. Table 4.2 shows tactics and outcomes for unassisted solution attempts by students and teachers on the reference problems. Cells in the table are filled with codes for each problem and show that teachers succeed on 5 of 6 problems attempted (83.3% correct), while students succeed on only 4 of 8 problems attempted (50% correct). As with advanced undergraduates, problem MRT appears the most difficult for both students and teachers.

Both students use model-based reasoning as a solution strategy, but their success with this tactic for finding solutions is varied. Celeste uses model tactics on three of four problems but succeeds only on problem MOD, using an iterative simulation over successive hours of train travel. On problems HOS and MRT, she loses track of correct intermediate state calculations without being able to satisfy limiting constraints. On

[4] 64.7% of annotations in the reference problems are two-dimensional tables or charts that organize quantities for outputs, times, and rates; 23.5% are diagrams; and the rest are recalled formulas or arithmetic expressions.

Table 4.2: Tactics and outcomes for unassisted solution attempts on reference problems.

	Tactic by Outcome†								
	Algebra			Ratio			Model		
Participant	C	I	N	C	I	N	C	I	N
Celeste	—	WC	—	—	—	—	MOD	—	HOS MRT
Karen	HOS WC	—	MRT	—	—	—	MOD	—	—
Paul	HOS	—	MRT	MOD WC	—	—	—	—	—
Richard	MOD HOS	—	—	—	—	—	—	—	—

†Outcomes: C = correct; I = incorrect; N = no solution.

problem WC, Celeste constructs a correct algebraic proportion over working times, but she then switches the reference of an intermediate result and incorrectly presents this as a solution (i.e., 16 boxes filled and checked).

In contrast, Karen uses model tactics to find a final solution only once, conducting an iterative simulation over successive hours of travel on problem MOD. She undertakes this state simulation as a last resort, being unable to choose between incompatible algebraic expressions generated during preceding episodes. Karen succeeds using algebraic tactics on problems HOS and WC, but she reaches an impasse during an algebraic solution attempt on problem MRT after incorrectly attempting to add one–way distances (a conceptual error of commission). Earlier on problem MRT, Karen abandons a model–based iterative simulation undertaken as a solution attempt by spontaneous comparison with her correct model–based solution on problem MOD.

Neither teacher uses model–based reasoning when finalizing a solution attempt, but a detailed examination of Paul's final episodes shows that algebraic tactics do not always proceed smoothly. On problem HOS, Paul recalls and manipulates a formula to find a solution ($1/6+1/2=1/x$), but he then remarks that he could not explain the terms in the formula since "all of this is rote." On problem MRT, he constructs an algebraic expression for composite times, but then decides that a correct solution (16 miles) is wrong because its value is less than the rate of the bus (24 mph). Finally, Paul constructs an algebraic expression with an incorrect rate form on problem WC

(56=x/5+x/2), but he then uses a ratio tactic to scale the resulting 80 boxes down to fit constraints given in the problem statement (i.e., a combined time of 56 minutes).

Three of Paul's solutions involve skilled algebraic manipulation, but in none of these solutions are his algebraic expressions accountable to the problem statement. In contrast, Richard incrementally constructs algebraic expressions for both problems MOD and HOS, using model–based reasoning episodes as explanatory justifications for quantitative relationships (e.g., "... from the drawing, we can tell that when we add the two separate distances together, we get the total distance of 880 kilometers."). He manipulates algebraic expressions only after deciding that they accurately render structural constraints contained in or implied by the problem statement.

In summary, the solution attempts shown as final episodes in Table 4.2 appear to present a progression towards competence in algebra story problem solving. While both students and teachers use nonalgebraic tactics like model–based reasoning during their solution attempts, they differ considerably in how (or whether) they use algebraic tactics to finalize these solution attempts.

- Celeste uses algebraic tactics (i.e., a proportion) infrequently and with difficulty, preferring model tactics but often unable to organize local state calculations in a way that progresses towards a solution.
- Karen frequently constructs or recalls incompatible algebraic expressions, is often uncertain about their relation to other constraints she can infer about a problem's structure, and sometimes returns to model tactics as a solution method.
- Paul also constructs or recalls recalcitrant algebraic expressions, but he works around these expressions using constraints identified with nonalgebraic tactics, sometimes rejecting conceptual errors but also sometimes overturning correct quantitative results.
- Richard uses model and ratio tactics to support inferences about quantitative constraints, carefully justifying terms in algebraic expressions before manipulating these expressions to find unknown values.

Taken as a progression in algebraic competence, the solution attempts of these participants (a) resemble the episodic interpretation made of solution attempts by advanced undergraduates, (b) confirm that algebra story problems present difficulties for a wide range of problem solvers, and (c) show the that the algebraic formalism is a brittle method when used in isolation. The latter point is clear in the contrast between Paul's recall of a relatively opaque algebraic formula ("all of this is rote" on problem HOS) and Richard's carefully justified construction of algebraic expressions on problem MOD.

Tactical course of structural inferences. Among advanced undergraduates, tactical episodes contributed to problem solving by generating constraints, pursuing constraints to determine unknown values, and evaluating constraints in light of other information about problem structure. An analysis of the origin of correct and incorrect

Table 4.3: Percentage of episodes introducing correct or incorrect structural constraints.

		Tactic			
Participant	Origin	Algebra	Ratio	Model	Annotation
Students	Correct	41.7	50.0	47.1	28.6
	Error	33.3	0.0	17.6	42.9
	(n)	(12)	(4)	(17)	(14)
Teachers	Correct	33.3	20.0	45.5	66.7
	Error	33.3	20.0	0.0	0.0
	(n)	(9)	(5)	(11)	(3)
Students	Correct	38.1	33.3	46.4	35.3
and	Error	33.3	11.1	10.7	35.3
teachers	(n)	(21)	(9)	(28)	(17)

structural inferences showed that model tactics were comparable to or slightly better than algebra and ratio tactics for generating correct constraints and were much less likely to introduce conceptual errors. At the same time, model tactics were as likely as algebra and ratio tactics to play an effective role in evaluating constraints by recovering from conceptual errors. Similar analyses were performed for solution attempts in verbal interviews.

Table 4.3 compares the percentage of problem–solving episodes that introduce correct structural constraints with those that introduce conceptual errors. As in the analysis of written protocols, tactical episodes are the unit of analysis. The origin of conceptual material is identified within individual solution attempts, allowing percentages to be interpreted as the relative likelihood that an episode using a particular tactic will introduce a correct constraint or an error within a participant's efforts on a single problem. Like the analyses of category prevalence and final episodes in verbal interviews, these quantitative comparisons exclude episodes that follow interviewer recommendations, though identifying the origin of inferences uses intact solution attempts (i.e., with interviewer recommendations).

For both students and teachers, algebra and model tactics are comparable in the origin of correct structural inferences (Tactics × Correct constraints, not significant), but model tactics are less likely to introduce conceptual errors than are algebraic tactics (Tactics × Conceptual errors, $\chi^2(3) = 5.9$, $p < .12$). Adding episodes from problems RACE and BAGELS to this analysis, algebra tactics are 3.7 times more likely than model tactics to introduce conceptual errors (Tactics × Conceptual errors, $\chi^2(3) = 9.3$, $p < .03$). Model tactics are particularly resistant to errors among teachers,

Table 4.4: Percentage of tactical episodes retracting or replacing an earlier conceptual error.

	Tactic			
Participant	Algebra	Ratio	Model	Annotation
Students	33.3	0.0	29.4	21.4
(n)	(12)	(4)	(17)	(14)
Teachers	22.2	0.0	18.2	0.0
(n)	(9)	(5)	(11)	(3)
Students and teachers	28.6	0.0	25.0	17.6
(n)	(21)	(9)	(28)	(17)

although teachers and students are comparable when making correct structural inferences with model tactics.

The large contrast between students and teachers when introducing conceptual errors with annotation tactics results from Karen's generally unsuccessful use of tables to organize quantities and relations between quantities, essentially a method for constructing algebraic expressions that Richard performs without error. Ratio tactics occur infrequently in both groups. Celeste accounts for all student uses of this tactic, and in two of these episodes she introduces a correct comparison of times. Paul accounts for the only error originating during a ratio episode, when he leaves out the checking time for boxes on problem WC.

Among advanced undergraduates, model–based reasoning showed no particular advantage in recovering from conceptual errors, and the same is true of students and teachers in the current analysis. Table 4.4 shows the percentage of episodes that either remove or replace an earlier conceptual error. An error is removed when a following episode no longer contains that error and is replaced when the following episode introduces a correct constraint in place of that error. As with the origin of structural inferences in verbal interviews, episodes with interviewer recommendations are excluded from this analysis.

Students generally have more errors to recover from than do teachers (see Table 4.1), and as a result they show an increased likelihood of recovering from conceptual errors in all but ratio tactics. In comparison with teachers, students are almost twice as likely to recover from a conceptual error using model tactics. When episodes from students and teachers are combined, algebra, model, and annotation episodes are equally likely to recover from prior errors, while ratio tactics support recoveries less frequently (Tactics $\times$ Repairs, not significant). Though not included in this analysis, the only recovery using a ratio tactic comes from Paul (a teacher), who retracts an error in rate form on problem WC during a prompted verification of an incorrect solution.

Spontaneous problem comparisons. The manipulation of problem ordering designed to elicit analogical transfer in Chapter 3 had little effect on advanced undergraduates in an examination setting. Approximately 10% of competent problem solvers gave clear evidence for undertaking a problem comparison, but most of these led to incorrect inferences about target problem structure, such as transferring an inappropriate algebraic expression.

Part of the original reason for studying analogical problem-solving was the prevalence and success of analogical inferences observed in a relatively open-ended problem-solving interview with a prealgebra student, who was alerted to the possible utility of problem comparisons[5] (Hall, 1987). Episodes of analogical comparison observed with this student (a) were articulate, usually signalled by a verbal exclamation and a shift in attention to retrieved material, (b) were time consuming, and (c) drew on materials both inside and outside the interview setting (e.g., problems his parents had asked him to solve).

As described in Section 4.2, participants in these interviews were periodically asked to compare a problem they had just solved with earlier problems, which either immediately preceded their solution attempt or were drawn from the entire set of preceding problems. Thus, students and teachers were alerted to the possible relevance of comparisons between problems, had unlimited time to pursue any such comparisons, and could easily select earlier work from their own arrangement of materials. One student reported routinely working on homework exercises by searching for related problems in the text or looking at the answers to paired problems at the back of an algebra text. Thus, problem-solving activities resembling analogy to worked examples (e.g., Chi, *et al.*, 1989) were familiar for at least some participants in this study.

Table 4.5 shows all observed instances of spontaneous problem comparison on a common set of problems. A comparison is considered spontaneous if it does not follow directly from a requested comparison. In the table, problem codes are annotated with a "+" if the comparison promotes a correct analogical inference, with a "–" if the comparison promotes an incorrect inference, and with a "(–)" if a candidate analogical inferences is rejected during the comparison. Only the "–" annotations should be considered incorrect inferences, since on closer inspection of these comparisons, rejecting an incorrect conjecture can be as useful as accepting a correct analogical inference. Evaluation of the rejected constraint often either implies a correct constraint or narrows the remaining alternative relations that a participant might consider.

[5] A low frequency of spontaneous analogical comparisons is a common finding in experimental settings where participants are not alerted to the utility of comparisons (e.g., Gick and Holyoak, 1980, 1983; Reed *et al.*, 1985).

Table 4.5: Spontaneous problem comparisons involving correct, rejected, or incorrect analogical inferences about problem structure.

	Participant			
Target	Celeste	Karen	Paul	Richard
MOD	—	—	—	—
HOS	—	CLOS +	—	CLOS +
MRT	—	MOD - HYPO (-) BOATS (-) CLOS (-)	—	n/a
WC	—	SINKS (-) HOS (-)	—	n/a
RACE	TUT+ CLASS+	DANCE (-) CLOS - DANCE -	WC (-) MRT (-)	n/a
BAGELS	BABYSIT +	n/a	—	n/a

Every participant verbally reports at least one spontaneous problem comparison. For example, Karen refers to her solution for problem MOD, saying "I was thinking if I did the table thing like we did before..." as she begins an unsuccessful model–based solution attempt on problem MRT. Comparing students and teachers, spontaneous analogical comparison appear more prevalent among students, although Karen generates more comparisons than all other participants combined. Both groups are relatively successful when making problem comparisons, Karen being the only participant who introduces incorrect constraints by analogy to earlier problems in 3 of 10 spontaneous comparisons (e.g., on problem MRT, as described above). Karen also recalls analogical sources outside the interview setting. For example, on problem MRT she recalls and then rejects a round trip BOATS problem involving river current. Recall of material outside the setting need not draw only on school math problems, as when Celeste chooses fractional parts of an hour for a simulation of baking BAGELS by analogy to charging her employers overtime while babysitting. Finally, Karen constructs an alternative diagrammatic model (HYPO) for an incorrect quantitative constraint on problem MRT, and she then rejects the spontaneously constructed model and its associated constraint by comparison with her diagram for problem MRT.

In summary, spontaneous problem comparisons are both more frequent and more successful by comparison with advanced undergraduates in the written protocol study. First, most of the spontaneous comparisons shown in Table 4.5 could never have been detected on the basis of written materials alone. Second, these analogical comparisons succeed by playing a different role in solution attempts than appeared to be the case among advanced undergraduates. While spontaneous comparisons among advanced undergraduates generally *failed* by importing a complex expression into the target problem, spontaneous comparisons among students are partial, often undertaken to supply or evaluate a single quantitative constraint rather than an entire problem structure. It may be that the kinds of spontaneous problem comparisons detected during verbal interviews were more common during the written protocol study but could not be detected. These kinds of spontaneous comparisons would also be relatively *insensitive* to the problem ordering factor introduced to manipulate participants' chances of making a correct analogical inference. The distribution of sources for spontaneous comparisons in the verbal interviews corroborates this latter interpretation, suggesting that problem solvers may choose among quite diverse materials while comprehending the structure of a new problem.

4.5 Discussion

The qualitative and quantitative analyses presented in this chapter represent the tip of a very large empirical iceberg. Putting aside empirical images of competence that rest on number of problems solved or simple time to solution, this and the preceding study attempt to look deep within the episodic structure of algebra story problem solving among a variety of participants. The findings are surprisingly at odds with existing accounts of "novice" versus "expert" problem solving in mathematics and other domains where formal reasoning is prescribed as a competent way of thinking. This discussion gives an overview of these findings in light of the outstanding issues that opened the chapter.

4.5.1 Qualitative analyses of solution attempts on problem MOD

A detailed qualitative analysis of participants' solution attempts on problem MOD both confirms earlier hypotheses about problem–solving tactics and extends several of these findings with new analytic observations. As in written protocols taken with advanced undergraduates, the solution attempts of algebra students and teachers have an episodic structure around strategy, tactic, and content. In both students and teachers, tactical episodes other than constructing or manipulating algebraic expressions are observed, including model, ratio, and various annotational tactics. Model tactics again play important strategic roles: generating relevant quantitative constraints (e.g., Paul's use of a state diagram to infer that train distances can be added together), determining unknown quantities without introducing algebraic expressions (e.g., Celeste's and Karen's solution attempts using iterative simulation), and evaluating quantitative conjectures (e.g., Karen's retraction of an incorrect algebraic constraint by comparison with a role diagram).

These qualitative analyses also bear on several of the outstanding issues that opened the chapter. Clearly, the episodic structure of problem solving is much richer than observed using written protocols, with the possible exception of Paul's relatively brief ratio solution. While participants' written notations, viewed retrospectively, give a largely sequential view of episodic structure, the verbal interview transcripts show that participants shift between episodes more frequently, often resuming, updating, or revising earlier written notations during a solution attempt. The higher density of episodic activity is corroborated in the quantitative contrast between algebra students, algebra teachers, and advanced undergraduates on a comparable set of problems.

Some of these episodic patterns appear to be opportunistic, as when Paul shifts from a diagram as a model of the first hour of train travel to a whole/part ratio solution — a frequent episodic pattern among advanced undergraduates. Paul uses algebra explicitly only when the interviewer requests that he verify a solution that he considers a settled issue. In contrast, Karen and Richard pursue solutions in a prescribed tactical "style," Richard's performance providing a pedagogical model for Karen's attempt to systematically establish algebraic constraints over given and desired quantities. What is missing in Karen's performance is the careful coupling between model tactics as a source of quantitative inferences and algebraic tactics as a precise manipulative calculus. Ironically, Karen constructs a correct algebraic expression with great difficulty, but is so uncertain about what its terms refer to in the problem situation that she abandons the expression before manipulating it to find a value for travel time.

The algebra students in these interviews, Celeste and Karen, provide a detailed look into the functional anatomy of model-based simulation. In many respects, their activities and notations mirror the dimensional structure of quantities carried in the problem statement. However, these activities and notations manage quantitative inference in a way that simultaneously mirrors the structure of events described by the "story" in a problem statement. In the words of Celeste, "if they all have something in common then that's a lot easier for me to do" (S-23). The consequence of this multiple correspondence is that quantities and calculations take a shape very different from the primarily static, retrospective character of school algebra. This is a fundamental quality of model-based simulation as a solution strategy.

Students' solution attempts also corroborate hypotheses about the functional role of model tactics in generating and evaluating constraints. However, their reflective comments about these tactics show that model activities (e.g., calculating and comparing state values) and notations (e.g., two dimensional tables or one dimensional diagrams) are familiar to students, providing a way to be *certain* about some aspects of problem structure in the face of the explicit demand for *precision* that algebra story problems present. These tactics take on meaning for students: they are private and illegitimate, "a weird way to do it" (Karen, S–26), by comparison with the more public and legitimate tactics of school mathematics, "some other way to do it" (Celeste, S–3).

4.5.2 Quantitative comparisons across competence

Attempting to solve a comparable set of problems, algebra students and teachers show similar patterns of episodic activity in their solution attempts to those found among advanced undergraduates.

- Solution attempts have a complex episodic structure, more often resembling the construction of quantitative constraints rather than direct recall of quantitative structures.
- Model–based reasoning is also prevalent as a nonalgebraic tactic among students and teachers, and even prevalent within each participant's solution attempts, ranging from 28.1% (Karen) to 53.3% (Celeste) of all observed episodes for each participant.
- Model–based reasoning is sometimes used as a final solution method, though not by teachers and not always successfully by students.
- Nonalgebraic tactics compete favorably with algebra for the introduction of correct structural inferences.
- Algebraic tactics are much more likely to introduce conceptual errors than other problem–solving tactics, and this finding is even more pronounced among teachers than among students. This contrast is stronger still if Karen's frequent and error–prone "drt" and "trw" tables are interpreted as algebra tactics.
- Model tactics are comparable to algebra and other tactics for recovering from conceptual errors.
- Spontaneous problem comparisons are more frequent and successful in the verbal interviews, and are used most often among students.

The analysis of verbal interview materials in many ways corroborates the interpretive framework developed using written protocols collected from advanced undergraduates. Complex episodes of problem–solving activity involving varied conceptions of problem structure are not idiosyncratic to advanced undergraduates. Instead, a similar episodic structure can be found in algebra students and teachers, suggesting that this descriptive account of algebra story problem solving holds across very different levels of competence.

There is also evidence within the verbal interviews that the relation between model and algebra tactics changes as problem solvers gain more experience. Comparisons of solution attempts by algebra students and teachers show several important differences.

- Students use model tactics to solve problems directly, either without algebra or when they reach an impasse in algebraic tactics. When used to find a solution, however, model tactics sometimes require state calculations that are difficult to coordinate.
- Teachers also use model tactics, though generally not to find solutions. Instead, their model-based reasoning episodes are combined with algebra and ratio tactics in a way that supports constructing and evaluating quantitative constraints.
- Manipulation skills probably improve with experience, but manipulation errors are a minor source of difficulty across levels of competence.
- Algebraic expressions are narrowly effective as recalled artifacts (e.g., Paul's recall of a "rote" formula for work problems) and quite brittle in isolation from other tactical materials (e.g., Karen's uncertainty over incompatible expressions).

These findings suggest that nonalgebraic tactics, including model-based reasoning, shift from solution strategies among beginners to comprehension and evaluation strategies among more experienced problem solvers. However, competence in algebra story problem solving does not become a matter of routine recall and application of algebraic forms. Instead, competent problem solvers more often construct a collection of quantitative constraints, sometimes finding algebraic expressions that are accountable to what they recognize about problem structure. In contrast, beginning algebraists are able to use model tactics to comprehend and sometimes to solve algebra story problems, but these tactics may not be connected in a useful way to the prescribed ontology of algebra: quantitative roles bound in the relatively opaque structure of algebraic expressions. Acquiring competence, then, appears to be a matter of connecting up diverse notations, some familiar (e.g., diagrammatic or tabular depictions of states) and some taught (e.g., algebraic equalities), to precisely determine unknown values that are accountable to one's understanding of problem structure.

4.5.3 Making mathematics on paper

Quantitative inferences and solution attempts on algebra story problems are remarkably complicated human performances. What is striking in the solution attempts of all four participants studied in this chapter is the role of written notations in their progress on comprehending problem structure, planning calculations, and assessing progress towards a solution. Theories of problem solving in cognitive science usually seal most of this activity inside the head, treated as a cognitive architecture that converts environmental patterns into contingent behavior. This view of problem-solving activity relegates the work setting — papers, desks, beds, kitchen tables — to the role of a prosthetic memory, an adjunct to the cognitive machinery that holds the problem statement or records intermediate cognitive results.

A careful analysis of the episodic structure of problem-solving activities — inferring constraints, reaching impasses, and working around these impasses — shows that written materials shape problem–solving activity as they are constructed. Rather than playing an adjunct role, these materials provide the medium out of which constraints are constructed, activities are carried out, and problem–solving performances are accomplished.

For example, in Celeste's solution attempt on problem MOD, a table of state calculations obscures an irrecoverable conceptual error about rate as a relation between time and distance. This error appears during her transition between gestural/narrative state construction and a written table of calculations and leads Celeste into a problem–solving impasse. After the interviewer supports a repair to the table, she easily resumes an iterative simulation by extending her table down the page to find a precise solution. Thus, at the boundaries of what appears to be a mental performance, written notations on the one hand hide a conceptual error but on the other hand provide a natural structure for activities (i.e., a "plan") in a successful solution attempt.

Analyses reported in this chapter have gone deeper into the content and structure of quantitative inference, showing that notations and activities around those notations provide empirical coverage for much of the episodic structure of participants' solution attempts. In the next and final substantive chapter of the dissertation, these ideas are taken into an analytical framework that asks how diverse notations can be integrated to produce a representational system within the problem–solving setting.

Chapter 5
Reconstructing Applied Quantitative Inference

5.1 Models and quantitative inference

What is model–based reasoning and how does it support quantitative inference? In analyses of solution attempts by advanced undergraduates (Chapter 3), model tactics are one interpretive category among many in a framework for exploring the episodic structure of algebra story problem solving. Participants' problem–solving tactics involve a "model" when they produce states along a dimension defined by an unknown quantity. This operational definition interprets particular kinds of written material according to a prescriptive analysis of problem structure, which is described in Chapter 2 as the "situational context" presented by an algebra story problem. A problem's situational structure is made up of semantic relations between entities that exist within typed dimensions. These entities and relations provide a hypothetical ontology out of which models can be constructed, manipulated, and used to make quantitative inferences.

In Chapter 4, model tactics are a focal point for qualitative and quantitative analyses of verbal interviews. When the interpretive framework for problem–solving episodes is applied across settings (i.e., group testing versus verbal interviews) and across participants with different levels of competence (i.e., advanced undergraduates, algebra students, and algebra teachers), similar episodic patterns are found: model tactics are prevalent, they are sometimes used to finalize solution attempts, and they compete favorably with other tactics when introducing correct constraints and repairing incorrect constraints on problem structure. Model tactics are used to generate and evaluate quantitative constraints at each level of competence, although algebra students are more likely than teachers or advanced undergraduates to attempt solutions using these tactics.

Chapter overview. This chapter attempts to explain how diverse notations, quantities, and activities form a mathematical representation, both for beginning and experienced problem solvers. Activities interpreted as model–based reasoning play a central role in this analysis. On the one hand, model tactics allow beginners to manage the demand for quantitative precision in the face of uncertain mathematical notations.

On the other hand, these tactics persist in significant functional roles among competent problem solvers. Looking more closely at the materials constructed during problem–solving episodes, why should this be the case?

The answer comes in several parts. First, competing frameworks (i.e., cognitive and ecological) for an adequate theory of competent quantitative inference are compared in terms of their central claims about knowledge and representation. As with most going concerns in scientific practice, neither is directly falsifyable in any objective sense, and they may even stand in a symbiotic relation around complex human problem solving. Second, an ecological analysis is constructed on top of interpretive categories developed in preceding chapters. This analysis argues that problem–solving tactics, strategies, content and outcome are critically dependent on the structure of notations and the way these notations carry different kinds of quantities. Results of qualitative and quantitative analyses are presented as support for the epistemological importance of relations between knowledge and setting, treating notations as representational ecologies for competent quantitative inference.

Third, the capacity of model notations for simultaneously managing the correspondence between quantitative and situational structure is developed as an explanation for how model–based quantitative inference works and why this form of inference persists across the competence spectrum sampled by written and verbal protocol studies. The empirical results and analytic arguments in this chapter lead to a conclusion for the dissertation, framing mathematical knowledge as both constructive and dependent upon concrete settings. The structure of tactics, their quantitative content and support for strategic activity are as much a part of the problem–solving setting as they are the subjects of mental representation. What a problem solver knows about mathematics becomes partly a matter of what she can construct in the problem setting at the very point of demonstrating competence. Taking this view seriously, competence becomes a matter of reconstructing knowledge across settings.

5.1.1 Towards a theory of competent quantitative inference

Analyses in the previous chapters bring model tactics to the foreground as empirical evidence for the constructive (versus recalled) nature of quantitative inference. For example, students' successful solution attempts on problem MOD show that model notations and activities simultaneously coordinate between the dimensional structure of quantities and the structure of events portrayed in the problem statement. This is a very delicate form of work: it is fitted to the local circumstances of the problem and physical setting (e.g., motion versus work, exact versus "close enough," beds versus desks), but it is not completely idiosyncratic, since problem solvers with very different backgrounds engage in similar activities (e.g., model tactics for comprehension).

The interpretive framework for these studies ascribes "strategies" and "tactics" to problem solvers, and these allow us to describe relationships between problem–solving activities and outcomes. These interpretive categories are operational definitions that may be applied to written or verbal protocols, but the resulting relationships can be consistent with quite different theoretical explanations for quantitative inference. The question taken up in this chapter is what sort of theory to embed these empirical contrasts within and, more important, whether the empirical interpretation developed in these studies can be used to support theoretical alternatives to existing accounts of competent quantitative inference.

Descriptive and explanatory adequacy. What would a theory of quantitative inference that explains these phenomena look like — i.e., how can we approach traditional standards of descriptive or explanatory adequacy (Chomsky, 1965; Johnson–Laird, 1983)? First, we need to adequately describe what problem solvers are doing, within the observational record available in written protocol and verbal interview studies. This is a narrow observational window (though wider than many), and the studies in Chapters 3 and 4 examine only a subset of the existing variety of algebra story problems (e.g., see Mayer, 1981). With these limitations in mind, the questions to be answered are: how do participants reach the quantitative conclusions that they do, and why are some kinds of inference more difficult than others? The preceding chapters document several interesting empirical patterns in episodic activity, and this chapter provides an additional line of analysis that opens up individual episodes to look for relations between materials, activities, and outcomes.

For the theory to be explanatory, we also need to say how people could acquire this descriptive account of competent quantitative inference. Comparisons across levels of mathematical experience in Chapter 4 show that competence is not so much *what people know*, in the sense of individual conventional notations for expressing quantities and relations, but *how conventions are integrated* together in a mathematical representation that is meaningful — i.e., a problem solver's sense of certainty and precision, at arm's length when the solution attempt starts, are brought together in a final solution. This chapter extends earlier comparisons across levels of competence with a new set of interpretive categories, identified in the qualitative analysis of Chapter 4 (Section 4.3) and operationalized in the next section. In many empirical studies of mathematical problem solving, students use standard quantitative notations (e.g., mathematical symbols) in apparent isolation from more familiar or conventional materials (e.g., concrete or mental models, Greeno, 1989), and this may be particularly true of algebraic expressions (Kaput and Sims–Knight, 1983). To understand how competence could be acquired, we must first determine what distinguishes the construction of a mathematical representation that supports quantitative inference from these apparently disconnected notations.

5.1.2 Cognitive and ecological approaches to quantitative inference

The central issue for any theory of competent quantitative inference is the relation between knowledge, setting, and activity when people make these inferences. There are many relations possible, but I will contrast two broad alternative perspectives that have a special significance for recent studies of mathematical problem solving across settings (i.e., school versus everyday life, as discussed in Chapter 2). The first is what can probably be called the "received view" in contemporary cognitive science and looks within the individual: knowledge structures, carried in the form of mental representations, give rise to more or less competent activity across settings — i.e., a *cognitive perspective*. The second perspective opens up the relation between knowledge and activity by also looking outside the individual for constraints on competence: a person's knowledge, activity, and particular settings are mutually determining — i.e., an *ecological perspective*. While neither perspective denies the existence of the other's focus, neither would allow that the other treats the excluded aspect (knowledge structures versus settings) comprehensively.

In important ways, these perspectives are competing accounts of the causal role of knowledge and setting in human affairs. They diverge over the phenomena of complex human inference, particularly in areas of human reasoning that have traditionally been seen as forms of logical inference (e.g., causal attribution) or procedural manipulation (e.g., mathematical problem solving). Careful empirical studies of these areas as "tasks" in laboratory settings have identified reliable sources of difficulty and changes in performance with experience. But a growing empirical literature, collected in practical settings from classrooms to street markets over the past decade, poses serious challenges to the representativeness of these experimental phenomena and the adequacy of theories based on these phenomena. The challenge amounts to a profound discontinuity between cognitive theory and human practice, as in studies of quantitative inferences in and out of the classroom (Carraher, Schliemann, and Carraher, 1988; Lave, 1988a; Scribner, 1984) or the complex enabling conditions for transfer of training (Campione and Brown, 1990). Although cognitive accounts of complex human reasoning may ultimately "scale out" to encompass these discontinuities, the central theoretical challenge is to deal adequately with the situations in which human cognition and learning actually occur.

An ecological analyses of "situated cognition" appears to require a shift in our working hypotheses about knowledge, from an objective structure possessed by an individual to a relation between the person and the social/material aspects of diverse settings (e.g., Brown, Collins and Duguid, 1988; Greeno, 1989). Although this debate concerns recent developments in cognitive science, the theoretical and methodological issues involved have a long history in the wider disciplinary structure of the behavioral sciences — e.g., Cole (1989) reviews related issues in cultural psychology, and Star (1988) frames recent ecological analyses of science within wider sociological traditions. Studies of algebra story problem solving will not resolve either the local debate in cognitive science or the broader disciplinary status of an ecological perspective on human activity. However, by focusing on the central relation between knowledge, activity, and setting, we can examine very different explanations for constructing and using algebraic representations. The specific question, for each position, is what assumptions are required to explain the relationship between problem–solving strategies, tactics, and outcomes in a way that allows us to define "competence."

Cognitive economies. The cognitive perspective assumes that strategies are abstract, plan–like knowledge structures (e.g., to solve an algebra story problem, find a precise value for the unknown quantity), while tactics are more specific knowledge structures that carry methods for implementing abstract strategies (e.g., to find an unknown quantity, write an algebraic expression that contains that quantity as a variable). Both are mental representations that are recalled in particular settings and give rise to activity that achieves better or poorer outcomes, depending on what knowledge these structures encode. Thus, what one knows resembles a cognitive economy: *knowledge structures are carried in the head, selected in particular settings, and then converted into coordinated activity.* Learning becomes a matter of getting knowledge into the head, usually in the form of explicit verbal instruction and practice, so that it can be carried across settings, and then getting this knowledge into action in an appropriate setting.

This perspective is probably best articulated as what Newell and Simon (1976) call a "physical symbol system," a machine that interprets structured patterns of symbols which designate objects in the external environment. According to the Physical Symbol System Hypothesis, these interpretive machines and some peripheral mechanism for designating objects in the world are "necessary and sufficient means for general intelligent action" (p. 116). Although adherents to this view have begun to explicitly examine the relationship between environmental structure and more abstract constraints on behavior (Anderson, 1988), the designation relation is generally not part of the research program. Instead, competence remains firmly a matter of the knowledge that one holds and can use across settings (Newell, 1982, 1989). The distributed representation and processing movement (e.g., Rumelhart, McClelland, and the PDP Research Group, 1986) does not appear to restructure this theoretical position, though it does call symbolic representations into question as an explanatory level of analysis for mental representations.

Representational ecologies. The ecological perspective similarly advances a causal argument about knowledge and human activity, but the organizing metaphor is one of interaction: *knowledge is constructed as a mutually determining relation between a knower and settings.* This might best be seen as intermediate between the cognitivist decision to attach knowledge of the world (i.e., settings) to mirror–like structures in the mind (Rorty, 1979) and earlier behaviorist theories that restrict mind to the peripheral status of stimulus–response bonds shaped by settings. Cognitivism and behaviorism are opposing poles of the familiar mind–body dualism, and they are polar opposites precisely because each resolves the dilemma of knowing about the world by eliminating the other's pole. The ecological perspective rejects this dualism by distributing knowledge across individual and settings (Costall and Still, 1987).

Because an ecological position takes knowledge as a relation between person and setting, problem–solving strategies can no longer be described in terms of internal knowledge structures alone. Instead, strategies become conventional and routine activities in an ongoing work process — i.e., a student tries to find an unknown value because this is what these problems are conventionally about in school and home work, and this activity is routine in the sense that one need not choose between alternative activities to get on with the conventional work. To the extent that teaching is organized around a view of learning in which the student needs to acquire a standard assortment of knowledge structures and then practice applying these in appropriate settings, the student *should not* choose alternative strategies during school or home work. In contrast, studies in Chapters 3 and 4 show that people do indeed deviate from the knowledge standard of recalled mathematical forms, sometimes in very interesting ways, and these nonstandard activities appear to be *necessary* constructive excursions.

From an ecological perspective, tactics are extemporaneous local resolutions to snags (de la Roche, 1986; Lave, 1988a, 1988c) or impasses that arise in ongoing activity of pursuing strategies. They are constructed out of conventional materials and activities (e.g., drawing, making tables, and calculating) but are nonroutine in the sense that their eventual form cannot be determined in advance. When problems are novel, unanticipated impasses in conventional strategies are frequent, and this requires that the relation between person and setting (i.e., knowledge) is reconstructed across settings, a process of work that cannot simply be recalled. Taking knowledge as a relation between person and setting, passive vessel approaches to learning can be expected to produce inert forms of knowing (Brown, Collins, and Duguid, 1988), rather like the robust disconnection between school mathematics or physics and everyday cognition reviewed by Greeno (1989). Instead, settings have both material and social aspects that define a "practice," and learning occurs as one begins to participate in this practice (Lave and Wenger, to appear). Competence, then, is a particular set of working relations between a person and settings of practice (both material and social), perhaps culminating in full participation within the ongoing work of particular settings.

5.2 Ecologies for quantitative inference

Alternative theoretical accounts of knowledge, as an objective commodity versus a reconstructable relation, probably cannot be resolved by any analysis of hypothetical representations. As ecological analyses bring forward detailed observations of human practice in particular situations, proponents of the cognitive perspective design knowledge structures to handle each new observation. This could be seen as an adaptive relationship, even a research strategy: from careful descriptive studies of human activity in open[1] representational ecologies, we get descriptions of knowledge structures within closed representational systems. But these positions are not so easily reconciled when their explanatory accounts of knowledge inform our choices about important social issues like the design of educational materials and settings (Anderson, Boyle, and Reiser, 1985; Lave, 1988b; Resnick, 1987a, 1987b), assessment of individual competence (Collins, 1987; Newman, Griffin, and Cole, 1989), and the design of environments for complex human activity (Norman, 1988; Norman and Hutchins, 1988; Suchman, 1987, 1988).

This is not to say that cognitive analyses of knowledge structures sufficient for complex activity are not revealing, or that criteria for representational equivalency have not been useful (e.g., Larkin and Simon, 1987). However, many problem-solving tasks have characteristics that these approaches, at present, do not adequately describe: (a) constructing a representation where there was none before, (b) solving problems when no single representation is expressive enough to make the problem well-structured (Simon, 1973; Star, 1989b), (c) solving problems where more than one agent is involved in performing a complex task (Hutchins, in press; Levin *et al.*, 1986), and (d) explaining why some types of inference are particularly robust or efficient. In important respects, algebra story problem solving has each of these characteristics, both for beginners and more experienced problem solvers. The analyses that follow are undertaken in the hope that an ecological approach will open up new descriptive and explanatory accounts of competent quantitative inference in this domain. In particular, we can begin an ecological analysis of quantitative inference by holding the interpretive categories of "strategy" and "tactic" accountable to a view of the notational and quantitative materials generated during problem-solving episodes.

5.2.1 The material basis of quantitative inference

Treating model tactics as an atomic category black-boxes important distinctions about the structure and content of quantitative inference. In the written protocols of advanced undergraduates (Chapter 3), state simulations most commonly appear as a series of calculations, organized as a table in which successive rows correspond to different states. Less frequently, states are constructed as a series of connected extensions to a spatial diagram, showing a single dimension explicitly (e.g., distance or work output) and carrying information from the other dimension (e.g., time) as peripheral annotations (e.g., quantitative labels, expressions, etc.). Model tactics also show a characteristic structure and content in qualitative analyses of solution attempts on problem MOD (Chapter 4). When problem solvers use model tactics to work around

[1]The "open" character of representations in human activity has been described by Simon as the "hopelessness of defining in reasonable compass a problem space that could not, at any time during the problem solving process, find its boundaries breached by the intrusion of new alternatives" (1973, p. 188). This challenge has been taken up recently in the literature on "distributed artificial intelligence" (e.g., Hewitt, 1986; Star, 1989b).

the dilemma of finding a precise value for an unknown quantity, their verbal reports add contextual detail to the role of conventional notations used during different tactics and the kinds of quantities that problem solvers explicitly manipulate.

As argued in Chapter 4, interpretive categories for strategies and tactics give a relatively neutral organizational overlay to written and verbal protocol materials. These descriptive categories do not tell us how models are constructed, how they promote inferences, or how they help to organize problem-solving activities. However, the verbal interviews are richer than the current episodic interpretation admits, and by extending this framework slightly, we can analyze the material setting for quantitative inference in these problems. The following sections examine interview materials by treating the problem-solving setting as a *representational ecology* in which materials are constructed, manipulated, and revised as a solution attempt progresses.

Framing assumptions. In the conventional meaning of an ecology (i.e., relations between organisms and their environment), what are the relations between people (problem solvers) and the environments in which they solve problems? The problem-solving settings studied here are complex ecologies in many respects: they are places where people work, places where they are asked to reproduce standard ways of thinking, and places where their efforts are evaluated by others. Although the verbal interviews do not study many of these aspects directly, the wider sense of ecology as a relation between a person and their work environment holds for advanced undergraduates, algebra students, and algebra teachers alike. The analysis of materials developed in this chapter focuses on a very narrow slice through the work settings of algebra students and teachers: the notational structures that they use and the ways in which they embed different kinds of quantities in these structures.

Ecological analyses of quantitative inference are not so commonplace that we have a clear methodological trail to follow. Studies of quantitative reasoning outside of school settings generally employ detailed observation of inference and calculation embedded in ongoing activities — e.g., buying groceries (Lave, Murtaugh, and de la Rocha, 1984), preparing meals under dietary constraints (de la Roche, 1986; Lave, 1988c), filling package orders (Scribner, 1984), or receiving payment for marketplace transactions (Carraher, Carraher, and Schliemann, 1987). In each case, aspects of the material and social setting figure prominently in the observational record and the interpretation of findings.

Even within the individualized context of working on school math problems, the material setting for quantitative inference is quite complex. Analyses in this chapter focus on specific materials constructed and manipulated by problem solvers during their solution attempts. Materials originate within and are carried across episode boundaries identified in a relatively coarse interpretive framework, and the following assumptions frame an analysis of these materials:

- There is a simple descriptive typology for the *structure and content of materials* that are produced within problem–solving episodes (described in a following section).
- As combinations of notational structure and quantities, these *materials form a representational ecology* within which problem solvers work.
- These materials are a physical setting in which *problem solvers construct tactics* that we interpret as a coherent problem–solving approach.
- The quality of *inferences about problem structure depend upon what local material constructions will afford*, their surfaces offering niches for quantities and their structure helping to organize activity (e.g., calculations).
- When people infer quantitative constraints on problem structure, *competence is a working relation between the person and material setting.*

Clearly, the key assumption for an ecological analysis is that the material setting can, in some sense, "afford" a problem solver with opportunities for inference or manipulative activities (Gibson, 1979; Chapter 4.9 in Reed and Jones, 1982). This is similar to part of Gibson's original intent for the term:

> The affordances of the environment are what it offers the animal, what it provides or furnishes, whether for good or ill... It implies the complementarity of the animal and the environment (Gibson, 1979, p. 127).

However, in using this term, I do not want to require that knowledge of the world must be direct, something for which Gibson's theory of direct perception (or ecological realism) has been roundly criticized (e.g., see Ullman, 1980, and the surrounding commentary). Instead, when problem solvers construct materials during a problem–solving episode, they are changing the local setting, with attendant changes in opportunities for inference and manipulative activity. By thinking of affordances as relations between the person and setting, I mean to introduce a level of analysis at which complex constructive problem solving can be examined.

As mentioned earlier, there may be no analysis that will distinguish between cognitive and ecological approaches to complex human activity. In fact, an effective ecological analysis will describe regularities in human activity that can act as specifications for extending a hypothetical collection of knowledge structures. Without trying to settle theoretical claims, the following analyses examine (a) the relative prevalence of materials used by students and teachers, (b) the material construction of problem–solving tactics, (c) relations between materials and strategic activities, (d) relations between materials and local difficulties, and (e) the relative likelihood of introducing inferences about problem structure (correct, conceptual errors, and repairs) when using different materials.

5.2.2 (Re)Constructing material designs

The added detail in verbal interviews makes it possible to develop a more systematic account of notations and quantities used during model tactics, and to ask how these relate to other problem–solving tactics and conceptions of problem structure. To do this, new categories for "notational structure" and "quantitative ontology" are added to the interpretive framework for problem solving episodes. These categories describe the local material setting of problem–solving episodes, and they were applied to verbal interview transcripts along with the interpretive framework as described earlier (Section 4.2 in Chapter 4).

Identifying local materials. Participants produce a variety of materials (enacted, narrated, or written) during an episode. Along with tactics, strategies, and other interpretations of episodic content, the following categories were used to identify conventional materials:

- *Gestures.* Usually also narrated but without written material, the participant uses hand or arm movements to depict aspects of problem structure. For example, Celeste moves her hands apart and narrates arithmetic operations to enact the first and second hours of travel on problem MOD (Chapter 4, Section 4.3.1, S–1 and S–3).
- *Narrative description.* Without written material, the participant gives a verbal report of operations on quantities or relations among problem elements. For example, Richard describes travellers with synchronized watches as a justification for inferring that travel times are equal on problem MOD (Section 4.3.4, S–8).
- *Expressions.* The participant explicitly writes either arithmetic or algebraic expressions. For example, Karen writes incompatible algebraic equalities before attempting a solution on problem MOD (Section 4.3.2, S–7).
- *Tables.* Written material (quantities, labels, arithmetic or algebraic expressions) and narrative descriptions of these materials are organized around the two dimensional structure of a table. For example, Karen completes a table of state calculations in a successful solution attempt on problem MOD (Section 4.3.2, S–26).
- *Diagrams.* Much as with tables, diagrams provide spatial locations for other notational material. For example, Paul constructs a diagram of successive states in problem MOD (Section 4.3.3, S–5), reading an arithmetic relation from the first state (i.e., adding distances) and later extending segments labelled as states towards the given distance apart.

These interpretive categories are are not independent, since tables and diagrams provide spatially organized niches or surfaces for placing or finding other notations (e.g., quantities or expressions). Thus, notations with more complex structure may afford opportunities for quantitative inference, much as local ecological conditions provide niches for organisms and their activities. The surfaces of more complex notations also afford configural information about their contents, as when labels on congruent segments in a diagram are read to imply that quantities contained in those labels should have equal values. While configural information is also available in tables, their structure provides less in the way of semantic constraints on quantitative inference. For example, the contents of adjacent cells in a table may be related, but their spatial adjacency provides no constraint on what that relation might be.

Notational structure. As ecological niches for inference and activity, notations afford configural constraints through their conventional structure. By conventional structure I mean the shape that materials take as they are constructed across space (e.g., successive rows in a table) and over time (e.g., constructing a sequence of states). Structural conventions need not be written, as when Richard's narrative description of time intervals on problem MOD describes a spatial shape (i.e., congruent intervals bounded by common places), while his narration follows a temporal structure (i.e., "If they leave at the same... then at some specific time later..." in Section 4.3.4, S–8). These shapes are conventional in the sense that different people construct and interpret them in similar ways.[2]

By moving from types of notation to their conventional structure, I am framing a specific hypothesis about the material basis of quantitative inference: the structure of conventional notations simultaneously renders and affords important relations between quantities. To pursue this hypothesis, we need to reinterpret descriptive types of notation as conventional structural forms:

- *Flat expressions.* Written arithmetic or algebraic expressions have a superficially flat structure, as when Paul narrates and then writes an arithmetic expression during a whole/part ratio on problem MOD (Section 4.3.3, S–2). Clearly, these notations are not flat under interpretation, but their written shape is linear, compressed, and referentially redundant (i.e., multiple occurrences of the same term).
- *Narratives.* Spoken materials have a conventional structure in two senses: (a) they often describe other structural shapes and (b) this descriptive talk has a temporal and articulatory shape of its own. For example, both are significant aspects of Richard's narrative model for equal times, described above. A detailed investigation of narrative structure is beyond the scope of this analysis, which treats narrative material without writing as a distinct category. However, narratives often provide configural material beyond the relatively flat structure of expressions, rather like a spoken scenario.
- *Two dimensional (2D) tables.* Tables organize quantities, expressions, and calculations around locations in a dimensional array. For example, Celeste organizes state calculations in a 2D table during her solution on problem MOD (Section 4.3.1, S–4).

[2]Livingston (1986) makes similar points about the "lived work" of constructing and interpreting mathematical proofs.

- *Scenes.* Diagrams organize quantities around locations in an iconic drawing of typed entities in a single dimension. For example, Paul uses adjacent line segments emanating from a common place when inferring that train distances are added on problem MOD (Section 4.3.3, S–2).

These structural categories oversimplify problem-solving activities by assuming that episodes showing conventional notations are mutually exclusive. Instead, more complex structures (i.e., 2D tables and scene diagrams) embed less complex structures (i.e., expressions and narrative material), meaning that problem solvers redundantly structure their activities with quantities both within and across episode boundaries. In order to analyze these materials, I also simplify the situation by pulling out what appears to be the primary form of notation (e.g., a diagram, with or without significant narration) and categorizing its conventional structure.

Quantitative ontology. People use quantities and quantitative relations in different ways during their solution attempts, and these can be described as a local quantitative ontology. During any particular episode, they may either use quantities to describe a "state" within the problem's situational structure, "compare" quantities associated with different states, or present a post hoc summary of the "role" that quantities play in the problem's quantitative structure.

- *State.* Model–based simulation tactics use intermediate states to organize calculations around a prospective partitioning of one dimension, as with time in Celeste's solution of problem MOD (Section 4.3.1, S–4).
- *State comparison.* Ratio tactics usually compare quantities arising from distinguished states within the problems' structure, as in Paul's whole/part comparison when solving problem MOD (Section 4.3.3, S–2).
- *Role.* Algebraic tactics typically use terms in expressions to summarize composite roles in a retrospective view of problem structure, as with train distances in Karen's various algebraic expressions on problem MOD (e.g., $A + B = 880$).

An alternative interpretation of quantitative roles as summaries is that they are "final states" in a temporal or exploratory history of problem–solving activities. In keeping with this view, state comparisons typical of ratio tactics may connect state and role ontologies for quantity, arranging initial and final state values in what amounts to a structural analogy with a simple manipulative calculus (i.e., cross multiply and simplify). These kinds of solution attempts may sit at a boundary between arithmetic calculation and algebraic manipulation, as in Paul's solution to problem MOD (Section 4.3.3). Celeste also makes frequent and relatively successful use of arithmetic and algebraic proportions across interview sessions, and she often produces these materials when asked to use algebra.

The following analyses reserve the ontological category of "state" for a prospective, partitioned sense of quantity that includes state comparisons. In contrast, the category of "role" captures a retrospective, composite view of quantity. Within any particular episode, a problem solver might view a problem's structure as a collection of states or as a set of related roles. Notations and quantitative ontology usually change over the course of a participant's solution attempt(s). For example, Paul's solution for problem MOD (Section 4.3.3) shifts from a written diagram used to infer the relation between distances in a single state, to a comparison of states using an arithmetic expression to divide the final distance apart by the distance apart after the first hour, and finally to a retrospective summary of quantitative roles using a recalled algebraic expression. Likewise on problem MOD, Karen shifts from diagrams, tables, and algebraic expressions that carry role quantities to a model-based simulation that organizes state values in a table (Section 4.3.2).

5.3 Qualitative analyses of material designs for inference

When categories for notational structure and quantitative ontology are combined, we have the materials for a much richer interpretation of differences between problem solvers and their problem-solving tactics. Only a few of these material combinations are "standard" activities in school mathematics. By standard I mean the prescribed use of conventional materials, as when algebra students are taught to manipulate expressions carrying quantitative roles. Similarly, ratio tactics using arithmetic or algebraic expressions to compare state quantities are sometimes explicitly taught in connection with algebra story problem solving. In contrast, complex notational structures carrying state quantities are decidedly "nonstandard" (i.e., 2D or scene notations), since they seldom[3] appear as instructions for how to solve algebra story problems. The surprising findings of Chapters 3 and 4 are that nonstandard materials play such a critical role in competent problem solving, even after years of instruction with standard materials.

5.3.1 Material designs within episodes

Figure 5.1 shows examples of complex notational structures — i.e., 2D or scene structures — that carry either state or role quantities. Examples are drawn from solution attempts on problem MOD, analyzed in detail in Section 4.3 of Chapter 4. Each notation involves written quantities, embedded in the spatial structure of a notation under different conventions. These are "material designs" in the sense that people construct the particular combination of notation and quantity within the episode. From an ecological perspective, these materials provide "niches" for information given or implied in the problem statement. 2D tables provide an organized

[3]While particular values for related variables are often shown in 2D tables as part of the conventional approach to graphing functions, these materials are not standard for instruction on applied problem solving.

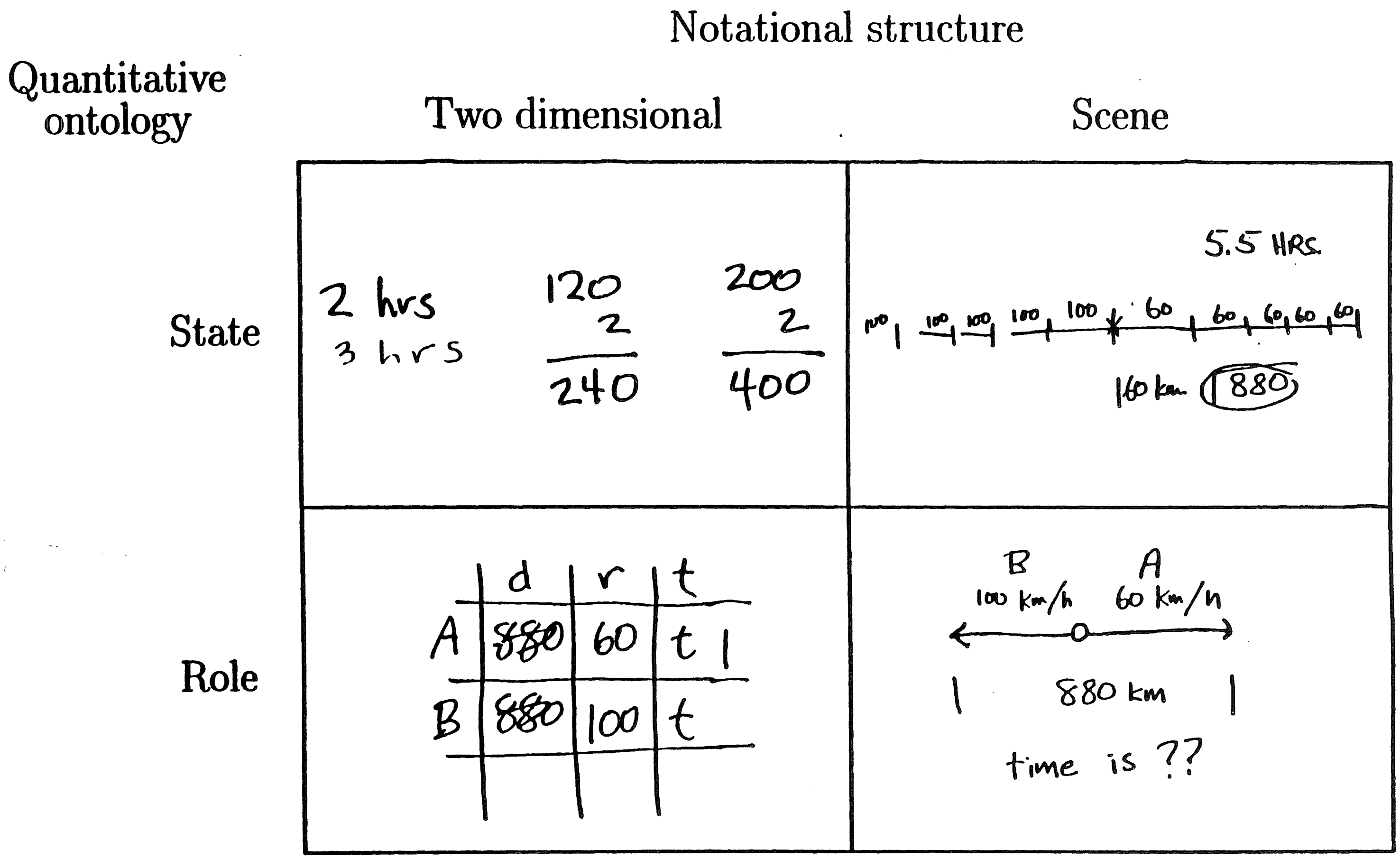


Figure 5.1: Complex notational structures and quantitative ontologies in participants' solution attempts on problem MOD.

array of cells that can be filled with different kinds of quantities, just as diagrammatic scenes provide iconic components whose surfaces can be labelled with different kinds of quantities. The ecological hypothesis is that when problem solvers label or fill niches in these notations, the resulting configural arrangements afford different opportunities for quantitative inference and manipulation. The following paragraphs examine the affordances of different material designs, using examples from Figure 5.1.

2D state tables. Celeste's table of state calculations (upper left in Figure 5.1) places quantities in cells of a 2D array: rows organize successive time intervals in the simulation, while columns separate quantities for faster and slower trains. Row and column position in the array designate properties of state quantities carried in each cell, and these properties often appear explicitly as row or column labels. Celeste's table uses this convention weakly by placing time units in successive rows, and a more explicit example of these conventions can be found in the third episode of Figure 3.2 in Chapter 3. Before starting the table, Celeste's calculations correctly use addition, embedded in a narrative and using hand movements to describe adjacent and opposite–directed motion during first and second hours of train travel.

After writing these results as the first row of the table, Celeste considers the third hour in her simulation and mistakenly chooses to double distances accumulated after two hours. This shift from addition to doubling introduces a conceptual error and coincides with a shift from one form of notation (narrative/gestural) to another (2D state table). In keeping with the hypothesis that material settings afford inferences about problem structure, this change in setting affords Celeste a different understanding of linear rates. At this point in her solution attempt, conventions for extending the table (i.e., adding rows and filling their contents) organize state values and calculations adequately, but nothing in the structure of the table gives explicit evidence that doubling, as an operation for transitioning to the next state, violates linearity. In particular, the conventions for adding rows to the table do not require a linear relation across quantitative dimensions, and the array of constant values that results gives no configural clues for judging linearity (i.e., disproportionate spatial extension). Treating the next row in the table as the next state, Celeste continues without detecting her error and reaches an impasse within the configural structure of the table (i.e., "I'd take the half way out" at S–4, then "you're just going back to where you started" at S–5 in Section 4.3.1).

As a material design, 2D state tables clearly support attempts to find precise solutions. They also provide positive and negative affordances for both inference and manipulation:

+ a positional classification for the quantitative dimension of cell values (e.g., time or distance) by applying row and column labels (constrains inference, organizes manipulation)
+ a replicable form for local relations like equal times and composite distances (constrains manipulation)

+ a local site for calculation, extending quantitative relations to the next row and state, in the case of iterative simulation (organizes manipulation)
+ a local site for comparing the current state with limit constraints, as well as choosing a generating value for the next state in a heuristic simulation (organizes inference and manipulation)
+ low overhead cost for revising the history of a simulation, simply by adding new rows (flexible manipulation)
- no configural constraints on choosing relations between state quantities from either dimension (inference)
- no configural constraints on choosing quantitative relations that transition between states, essentially a local version of rate as a role quantity (inference)
- no feedback on the linearity of rates or their composition (manipulation)
- no configural constraint or explicit location for the limit constraint (inference and manipulation)

State scenes. Paul's prompted completion of a drawing for train separation is shown in the upper right of Figure 5.1. The scene starts with distances travelled after the first hour, drawn as collinear and adjacent segments that are labelled with state values for distance. Elapsed time is not explicitly rendered, though Paul does narrate a temporal interval while constructing the state scene ("you start off... this is one hour's" at S–2 in Section 4.3.3). This starting scene promotes correct inferences about adding distances and equal times ("So they go 160 kilometers in 1 hour" at S–2) and quickly leads to a ratio solution that compares starting and final states (i.e., shown below his diagram as a flat arithmetic expression).

When asked to complete the scene, Paul constructs four following hours as segments connected to the prior state's scene. These extend away from the common origin and are labelled with individual values for distances. Although Paul does not construct these state scenes as a solution attempt, it is difficult to imagine how he could replicate a conceptual error like Celeste's successive doubling operation. This would require that he double the number of segments constructed between each successive state or that he change the value of labels for successive segments in the diagram. Either activity would violate the configural sense of moving at a constant speed that Celeste retrospectively narrates but is unable to verify in her 2D table of state calculations. Thus, different notational structures carry the same kinds of quantities with very different consequences.

Both 2D state tables and state scenes can organize calculations to determine precise solutions, but the configural properties of scenes afford a wider class of quantitative inferences:

+ configural constraints on quantitative relations in the rendered dimension, including local and limiting relations (constrains inference)

+ quantities can be carried as labels on their iconic objects of reference, imposing a common type in the rendered dimension (constrains inference, organizes manipulation)

+ a local site for calculation, extending quantitative relations as connected components of the scene for iterative simulation (organizes manipulation)

+ explicit feedback on the linearity of individual rates and their composition, within conventions for the shape and number of scene components (constrains inference, organizes manipulation)

- high overhead cost for revisions to the history of a simulation, erasing and redrawing iconic states (manipulation)

- multiple sites for state values, located at places or segments, and multiple renderings for segment labels, as local or cumulative values (inference, manipulation)

- partial rendering of the implicit dimension (e.g., counting connected segments to determine cumulative time) violates configural constraints for all but congruent relations in the explicit dimension (inference, manipulation)

- iconic components do not support metric precision for local or limit value comparisons, as in measuring cumulative distance (inference)

- iconic quality of configural constraints is restricted to linearly ordered dimensions like motion, output, or time (inference, manipulation)

2D role tables. Karen's 2D "dirt table" arranges role quantities in the lower left of Figure 5.1, providing cells for quantities that are typed by columns and related to events by rows. The structure of her table is much like Celeste's state table, but the contrast between role and state quantities is rendered by literally closing off any possibility of state activity at the bottom of the table. Also, the dimensions of the table do very different work in each material design. In Celeste's 2D table of state quantities, the horizontal dimension (columns) renders one quantitative measure (distance) in a way that separates the contribution of different trains, while the vertical dimension (rows) renders the other quantitative measure (time). Rate is distributed across the table as a history of simulation. In contrast, Karen's 2D role table mixes three quantitative measures together in the horizontal dimension (i.e., d, r, and t), while the horizontal dimension renders different events in the situational structure of the problem. These are very different material designs, both because of the different kinds of quantities carried in the table and because of the way the structure of the table is used to render relevant aspects of problem structure.

Karen manages to insert given values for rates into appropriate cells, but then struggles with where to place the given distance apart (880 km). Without introducing algebraic expressions relating train distances (e.g., D_1 and $880 - D_1$), the role table provides no correct niche for the limiting value, and Karen mistakenly inserts 880 into both distance cells for trains A and B. Transitioning between this annotation and algebra tactics (i.e., role expressions), she constructs several incompatible expressions and ultimately abandons both tactics in favor of model-based iterative simulation. The point at which Karen abandons her 2D role table illustrates how the same notational structure supports very different material designs. As shown in Section 4.3.2 of Chapter 4, she tries to insert a single hour into cells for role quantities: "Oh! Well time is 1 (writes 1 in time cell for train A)... No" (S–21). The conventional structure of the 2D role table will not hold state values for time in a way that is coherent with role values for rate and distance: "that would just be 60 plus 100, and that's not 880... 60 times 100, no that's not it either" (S–22). Abandoning the role table as a material setting for state quantities, Karen reconstructs an earlier role scene to show a state in train separation after the first hour, and then finds a solution in a 2D table of state quantities somewhat better organized than Celeste's solution (see Figure 4.3).

Unlike complex notational structures carrying state quantities, 2D role tables afford no manipulative capacity for determining precise values. However, they do afford additional constraints on the dimensional organization of table entries, even while they have many of the drawbacks of 2D state tables:

+ positional classification of dimensional and event type for role quantities (constrains inference)
+ recall of relations between adjacent columns when column labels carry a memorized formula (organizes inference)
+ implicit convention for recognizing when a sufficient collection of quantitative constraints are available or when necessary quantities are missing (organizes inference)
- no configural constraints on quantitative relations between quantities in cells (inference)
- no unique position possible for composite quantities (inference)
- no manipulative capacity

Role scenes. Richard constructs a diagrammatic scene carrying role quantities (lower right in Figure 5.1). This role scene explicitly uses the spatial configuration of collinear segments to show a composite relation between train distances and a "brace" for the given total (880 km). Unlike Paul's state scene, which could be used to organize calculations in a solution attempt,[4] Richard's role scene supports inferences about quantitative relations but cannot be directly manipulated to determine a precise value for the unknown time.

In contrast with Karen's difficulties when trying to find locations for quantities in a 2D role table and relations between these quantities as algebraic constraints, Richard uses this role scene to justify placing values and expressions in a 2D role table that is

[4]Celeste uses state scenes for iterative simulation on several problems; Paul uses a state scene to record a heuristic simulation on problem MRT; and comparable materials are used by the advanced undergraduates studied in Chapter 3.

almost identical to Karen's (see Figure 4.5 in Chapter 4). He then generates algebraic expressions by carefully reading cell contents out of the two dimensional structure of the table. Richard infers that unknown times are equal after a narrative account of their common boundaries, and he infers that distances add on the basis of their configural position in the role scene (i.e., "from the drawing, we can tell that when we add the two separate distances together, we get the total distance of 880 kilometers" in Section 4.3.4, S–11). Thus, he overcomes the relative configural opacity of a 2D role table by constructing narrative scenarios or diagrammatic scenes that render one dimensional relations between quantitative roles.

As with 2D role tables, role scenes afford no manipulative capacity for determining precise values. However, much like state scenes, their configural properties afford some forms of quantitative inference:

+ configural constraints on quantitative relations in the rendered dimension (constrains inference)
+ quantities can be carried as labels on their iconic objects of reference, imposing a common type in the rendered dimension (constrains inference)
- no explicit rendering of the implicit dimension, though peripheral expressions are often used (inference)
- iconic components do not support metric precision for value comparisons (inference)
- iconic quality of configural constraints is restricted to linearly ordered dimensions like motion, output, or time (inference)
- no manipulative capacity

In summary, different combinations of notational structure and quantitative ontology within an episode can be interpreted as the construction of material designs, and these lead to very different forms of quantitative inference in algebra story problem solving. 2D tables provide a positional typology for state or role quantities, but they provide very little explicit information about how these quantities are related. By comparison, scene diagrams for states or roles also provide organized locations for selected quantities (e.g., distances in Paul's state diagram), but they are less complete, often leaving entire classes of dimensional entities implicit (e.g., times in Paul's state diagram). However, the scene notations shown in Figure 5.1 render explicit configural information about quantities and their relations that is absent in 2D notations. Thus, the configural properties of scene notations have an advantage for some forms of quantitative inference.

While complex notational structures carrying states are nonstandard (i.e., not schooled), they are effective because there are conventions governing their use, and people usually come to the problem–solving setting aware of these conventions. Nonstandard material designs afford relevant inferences when their configural structure carries quantities in a way that justifies particular quantitative relations (e.g., same/equal, adjacent/additive). But material designs need to do more than afford relevant structural inferences if their users are to obtain precise results, and complex notations carrying state quantities often provide an effective (if not accepted) means for obtaining precision. For example, 2D tables or scene diagrams showing successive states support solution attempts because one can extend the written structure of the table or diagram through the simulated dimension to create new states, organize the calculations required for finding quantities in these states, and record the values that result within the local structure of the notation. That is, the conventions governing the construction of successive rows of a table or connected components of a diagram can carry quantities in what we have called model-based simulation.

5.3.2 Combining materials across episodes

Stepping back from the local arrangement of notation and quantity within episodes, the conventional written structure of nonstandard materials allows problem solvers to work around disruptions introduced by a need for precise calculation in a complex quantitative structure. But complex notational structures are not invented during algebra story problem solving. Instead, tables and diagrams are conventional notational systems in many settings, both inside and outside of school, and well before students are introduced to formal algebraic notation. For example, we might assume that pre–algebra students are able to read and use calendars, purchase and find seats in auditoriums, or memorize multiplication tables. Each designed artifact locates items (e.g., planned events, seats, or products) within multiple, ordered lists or classification schemes (e.g., days of the week versus weeks of the month). Similarly, we might assume that pre–algebra students are able to construct and interpret a simple map giving directions to some location. These drawings locate items in a configural scene, using distinguished places, labels, and topological relations of connectivity, relative position, etc.

Calendars or maps are not solutions to algebra story problems. However, their material design rests on a set of conventions that, when suitably reproduced and combined, are sufficient to carry quantities and support some forms of quantitative inference. The central empirical finding of "model–based reasoning" in beginning and competent algebra story problem solving can be explained in terms of constructing combinations of conventional materials. Although complex notations for 2D tables and diagrammatic scenes can take on more specialized conventions in the practice of mathematics (e.g., Cantor's diagonal method or Cartesian graphs for linear functions), people come to algebra instruction already fluent[5] in a set of conventions for

[5] This is a tacit "representational practice" not unlike relatively invisible aspects of scientific work recently described by sociologists of science (Fujimura, 1987; Lynch, 1988; Star, 1989a). In effect, the conventions around nonstandard materials allow them to act as "immutable mobiles" that are transportable across problem settings and locally fitted to particular problem settings (Latour, 1986).

constructing and interpreting these structures. Both support the construction of standard algebraic materials, but they also allow problem solvers to obtain quantitative precision without standard materials. This sense of conventions draws on sociological studies of constructing and interpretating artistic performances (Becker, 1982), and the idea that nonstandard material combinations "work around" algebraic precision draws on recent studies of routine computing use in diverse work settings (Gasser, 1984). Both lines of analysis stress the improvisational character of getting work done within and around formal systems.

Structural inference versus precise calculation. How do the conventions of different material designs contribute to making structural inferences and organizing quantitative calculations? Within an episode, any particular material design may support inference or calculation to a different extent, and it is often when the affordance for these activities is mismatched that people transition into another episode and construct another material design. These transitions sometimes occur when episodes go "off track" (i.e., the participant gets lost, reaches an impasse, or decides that they are wrong about some aspect of problem structure). For example, on problem MOD Karen cannot find a cell in her 2D role table that will sensibly carry a state quantity (the first hour of travel). Unable to pursue the state within a material design for role quantities, she reconstructs an earlier role scene into a state scene and confirms that the trains are separated by 160 km after one hour ("Well yeh. He'd have gone 60 and he'd have gone... that far in one hour." at S–23 in Section 4.3.2). Writing this inference as an association between values (1 = 160), Karen transitions into an episode where she constructs a 2D state table and finds a precise solution ("I'll do this til I hit 880!" at S–30).

Transitions between material designs also occur when people are "on track" and working towards standard algebraic representations. For example, on problem MOD Richard combines different materials across episodes in his solution attempt, working around the relative configural opacity of a 2D role table by taking advantage of explicit configural relations in both a narrative scenario describing coincident time intervals as role quantities ("If they leave at the same time..." at S–8 in Section 4.3.4) and a role scene showing composite distances ("from the drawing we can tell..." at S–11). These materials identify relevant structural constraints, and Richard incrementally assembles these in a material design with robust manipulative capacity — i.e., standard algebraic expressions.

Conventional ways of working around problems. State quantities generally have an accessible manipulative calculus for students (i.e., model–based iterative or heuristic simulation), provided that these quantities and local quantitative relations between them can be embedded within given constraints. For example, Karen's 2D state table on problem MOD allows her to accumulate local values for time and distance "til I hit 880!" Successive states are constructed by adding rows at the bottom of a 2D table, performing a common set of arithmetic calculations ($160 \times k$, where k is the "next" value for time), and writing the results in cells of each new row. These continue until the calculated result equals or exceeds the given distance apart (880). A nonintegral remainder (e.g., "you've got 80 more to worry about" at S–30) is a relatively minor disturbance within the organizational structure of the table, since local calculations can be scaled to "fit" within given constraints (e.g., "its just another half an hour" at S–31).

This is the sense in which model tactics are constructed as workarounds to standard algebraic solutions. The event structure of the problem text acts as a specification for a nonstandard material design: one can get started in a complex notation without committing to quantitative roles or their relations, relations are afforded in the course of constructing the design, and the conventions for manipulating components of the design (e.g., adding a row or connected scene) carry the organization for a solution attempt just as they carry quantities (e.g., a label on a segment or the contents of a table cell). However, workarounds using nonstandard material designs can break down: (a) when the problem describes events that are not states within a global constraint — e.g., "A small hose can fill a swimming pool in 6 hours..." in problem HOS;[6] (b) when limiting constraints are difficult to evaluate — e.g., equal distances in composite but unequal times, in problem MRT; or (c) when written materials are so weakly structured that progress is difficult to monitor. For example, during a solution attempt using model–based heuristic simulation on problem MRT, Celeste rearranges the dimensional structure of a 2D state table at subsequent states and loses track of progress towards a precise solution.

Role quantities carried in complex notational structures (i.e., 2D tables or scene diagrams) cannot be directly manipulated to find solutions. That is, there are no conventional activities with components of either notation that physically organize the calculation of precise values. A 2D role table shows composite values or variables for unknown values, but extending the table in any direction violates the dimensional meaning of events (vertical dimension) or role quantities (horizontal dimension). Similarly, a role scene shows relations between quantities but cannot be extended to obtain more precise information about these quantities. As a result, constraints that are identified using either material design must be carried into other materials that have a manipulative capacity. By design, the manipulative conventions of specialized notations like arithmetic or algebraic expressions support precise calculations that are rigorously consistent. However, these relatively flat notations must be constructed (or given) before manipulation is possible. While complex material designs for role quantities do no afford precise calculation, they do carry role quantities or even expressions in their cells or labels, and these can be assembled (perhaps with difficulty)

[6]See Appendix A for the texts of problems mentioned in this section.

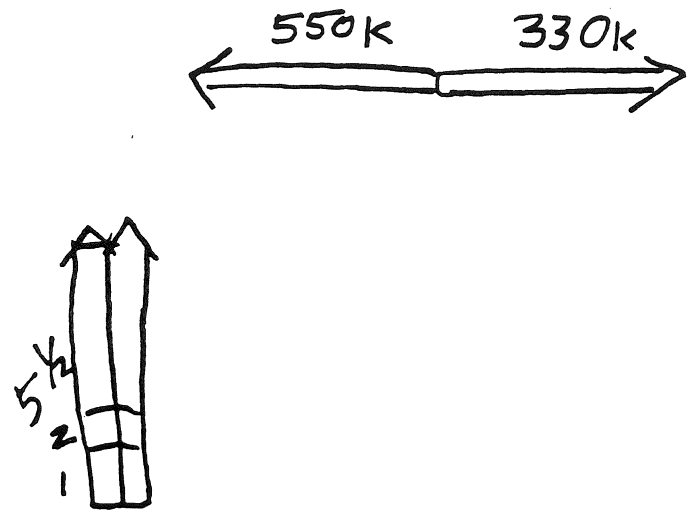


Figure 5.2: Tutorial in one dimensional scenes on problem MOD.

into a collection of related expressions. Within an ecological view of the material setting, this assembly transitions into a different material design.

5.3.3 Learning to combine standard and nonstandard materials

In order to construct an effective mathematical representation, one needs to be able to do more than use conventional materials in isolation. In fact, the studies reported in Chapters 3 and 4 show that people with very different levels of ascribed competence are able to use both standard and nonstandard materials, although algebra students have a much more difficult time combining these in a way that leads to standard algebraic expressions. At least part of what one needs to learn about algebra story problem solving is how to combine these materials to determine precise values within the standard notational structure of algebraic expressions.

Collaborative combinations of conventional materials. What follows is a series of protocol excerpts from interviews with Celeste (an algebra student) that show the results of a relatively simple tutorial intervention. Towards the end of her third session (of four), the interviewer presents a convention for drawing and labelling diagrammatic scenes for each quantitative dimension. Although the intervention and subsequent collaboration are brief, Celeste incrementally changes her approach to subsequent problems.

> I–49 I'll give this to you, and you can use it if you want. In this trains problem we saw we could draw (draws an elongated rectangle with an arrow on the end)... here's a distance going in that direction. They're travelling in opposite directions, directly away from each other (draws another). So this one's gonna go 550 kilometers (labels above)...
>
> S–49 And the other one's gonna be 330...

As show at the top of Figure 5.2, horizontal and adjacent directed segments for distance are drawn at the top of the page (I–49), and Celeste easily finishes labelling these with values, found earlier while solving the problem. Below and to the left, vertical and congruent directed segments are drawn for identical 5 1/2 hr intervals of travel (S–54). In this case, the relation between quantitative states and roles in the time dimension is described by analogy to a clock running to produce "piled up hours" and shown as (upward) directed segments in the scene.

> I–51 Now we can draw times in the same way. I can draw a segment for the time. Let's draw the times going up and down, as if its a clock where you just keep piling hours on top...
>
> S–51 Oh yeh.
>
> I–52 So here's a bunch of hours piled up, let's say the clock is running in this direction (draws an arrow at the top), its running upwards. So here's 5 and 1/2 hours.
>
> S–52 And here's where they started (indicates base).
>
> I–53 They start down at the bottom, that's the starting time.
>
> S–53 Ok.
>
> I–54 And so here's another 5 and 1/2 hours. Its actually the same 5 and 1/2 hours...
>
> S–54 Um hmm... for the other train, its exactly the same.

This collaborative construction of dimensional scenes for related distances and times depends on shared conventions for drawing, labelling, and reading diagrams that Celeste is already capable of using (e.g., her abbreviated diagram in Figure 4.2). However, the tutorial occurs after Celeste has solved the problem using a 2D table of state calculations (i.e., model–based iterative simulation), and directed segments are labelled with constant values rather than algebraic expressions. Thus, the tutorial does not actually use these role scenes to solve a problem or to construct a standard algebraic representation.

Constructing standard materials, after the solution. On the next problem (RACE, the last of the comparison problems), Celeste finds a solution by iteratively constructing states that extend a segment diagram (i.e., another model-based iterative simulation). However, immediately after finding a solution (S–106, below), she spontaneously draws a scene showing overlapped distances and writes an algebraic expression that captures the relation between role quantities for these distances.

> (RACE) Frank and Joan both plan to run in the West End race. Joan is faster and can run 10 kilometers per hour, while Frank only runs 8 kilometers per hour. Frank cheats by starting the race 5 kilometers ahead of Joan so that they will cross the finish line together. If both runners start the race at the same time, how long do they run?
>
> S–106 Ok, they went like this... (draws two equal length, same direction segments, one slightly above the other). Like that.
>
> I–107 Ok, so now they're going in the same direction. Its really the same distance, the way you've written it.
>
> S–107 Right. So it would be, like D1 equals D2.
>
> I–108 Ok, but we know that's not quite true.
>
> S–108 Well, yeah. Ohhh! That sounds like one of my algebra problems, because you go... like, um... D1 equals D2 plus 5 (writes this). Because 5 is the amount he cheated. And then you have to figure that out.
>
> I–109 Good. Ok. Now you already know what those values are, why don't you just write them in underneath.
>
> S–109 You could say, well, D1 is 25 and D2 is 20 plus 5.

At this point, the conventions for constructing and interpreting scenes and expressions that carry role quantities come together on the same surface, without prompting by the interviewer. Celeste also constructs a scene for equal times and writes an algebraic expression that describes these role quantities (i.e., $T_1 = T_2$). However, she still solves novel problems with 2D tables or scenes that carry state quantities, and only combines materials to construct a standard algebraic representation after the fact.

Constructing solutions with standard materials. During the fourth (and final) session, Celeste starts a solution attempt by constructing role scenes and then assembling algebraic expressions around components in these scenes (see Figure 5.3).

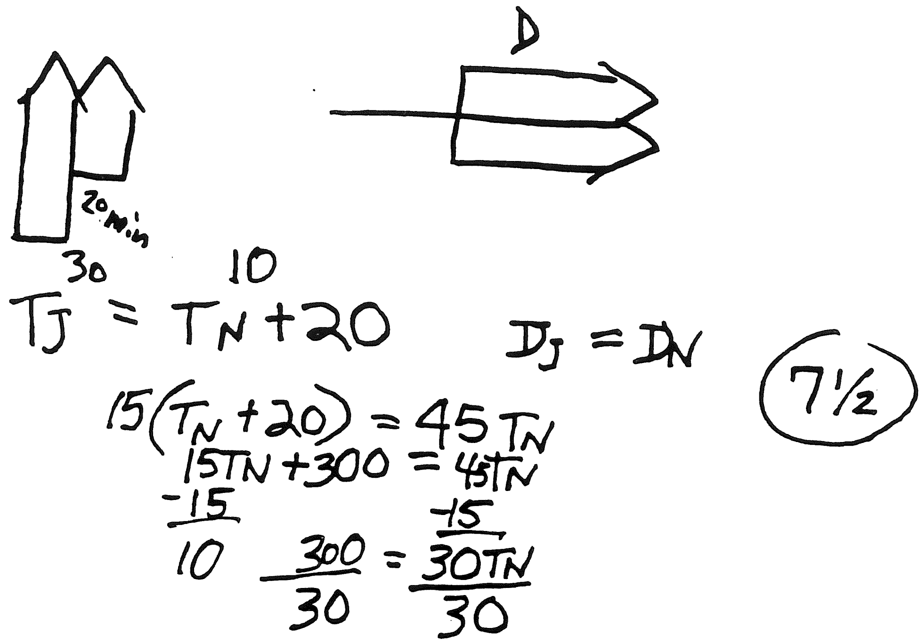


Figure 5.3: Celeste's solution to problem COMMUTE.

(COMMUTE) Nancy lives in Angleton and works in Beeville, while Jeff lives in Beeville and works in Angleton. Jeff cycles to work at 15 miles per hour, while Nancy drives to work at 45 miles per hour. For each to be at work on time, Jeff must leave his home 20 minutes before Nancy leaves her home. How far does each travel from home to work?

S–78 Ok. ... Ok, like for the diagram things. Let me start out doing it. Um... they're gonna go... here (draws two vertical, same-finish segments).

I–78 So what are you drawing? Times here?

S–79 Yeh. Yeh, Jeff has to leave his house 20 minutes before her, so he's got a head start here. Well, no wait... Ok, so... since he rides his bike, he goes slower. So the distance... they're going the same distance, but she's going faster. And he has kind of a head start.

I–79 So which time is his?

S–80 This (indicates shorter time).

Celeste constructs a scene in which collinear, directed segments finish at the same time, but she initially associates the shorter segment with Jeff's "head start." This would be a sensible reading of the scene as a model of distance, though she has already stated that "they're going the same distance" (S–79). Prompted to focus on the duration of Jeff's head start, Celeste correctly associates commuters with time segments and labels their difference (i.e., 20 min).

S–89 And so he goes 20 minutes earlier, and she goes 20 minutes later.

I–89 Ok, show me the distances now.

S–90 So... the distances here are gonna be... Wait, they're going the same direction (draws congruent segments, same direction). So they go the exact same distance... say 15 miles or something.

I–90 That's right, yeh. So what can you do with that?

S–91 (both laugh) You have to solve the problem now! Ok, um... she goes faster though. This is the distance (writes D above segments). So Nancy's time is gonna be... wait. Jeff's time is gonna be equal to Nancy's time plus 20, because he left 20 minutes earlier. (writes $T_J = T_N + 20$) Like minus... plus 20. So... hmm... its kinda like this problem.

[... Celeste retrieves prior problem, BUS–FLY...]

Despite misreading the direction of travel, Celeste draws a scene showing distances, narrates the relation between them ("So they go the exact same distance..." at S–90), and then constructs an algebraic expression for travel times beneath her original role scene. The retrieved problem, BUS–FLY (see Appendix A), is a quantitative isomorph of COMMUTE. On this problem, Celeste abandoned a solution attempt using a 2D table of state calculations, and the interviewer guided her through a collaborative solution using role scenes to construct algebraic expressions.

S–95 Yeh, they went... um... ok, so you could go that... like here (indicates algebraic expression over times in BUS–FLY), her distance or Jeff's distance, say. You say... well I already have this...

I–95 The time.

S–96 Both distances are the same, the distance of Jane or Jeff and the distance of Nancy are exactly the same. (writes $T_J = T_N$)

I–96 Ok, you probably want to call these D, rather than T.

S–97 (laughs) Yeh, right (overwrites with $D_J = D_N$). So the distances are the same, and so you'd take 15 miles per hour, and the time of Jeff. I mean... yeh, that is the time of Jeff, and the time of Nancy plus 20 equals 45 miles per hour... 45 time of Nancy, cause that's how fast she goes. So it'd be... 15 Tn plus 7 300... that's not right... equals 45 Tn. And then... minus 15 minus 15... whoops... 30 Tn 300 equals... over 30 over 30. So it would be 10. I mean... yeh, 10.

Starting with nonstandard materials (scenes for role quantities), Celeste constructs a pair of algebraic expressions in which terms refer to components of role scenes and operators refer to relations in the problem situation (e.g., "plus 20, because he left 20 minutes earlier" at S–91). After literally retrieving an earlier solution to a related problem, she uses a multiplicative relation between distance, rate, and time (either recalled or transferred from BUS–FLY) and manipulates the resulting algebraic expression to determine a precise value for Nancy's time.

I–97 So what's 10.

S–98 10 is the number... the time of Nancy to get there. And so now you have to find out the time of Jeff. You have the time of Nancy, so it takes Jeff 30 minutes, because the time of Nancy plus 20 is 30. Time that Nancy took... 10 minutes, and you already know that Jeff had a... cause it fits.

I–98 Sure, if you add 10 and 20 you're gonna get his time, that's exactly right. What are you supposed to find?

S–99 The... how far does each travel? Oh, well... in 10 minutes... You gotta find out how far Nancy goes in 10 minutes. Because how far Nancy goes in 10 minutes is how far they both go, or how far Jeff goes in 30 minutes. So if he goes 15 miles per hour, then he goes 7 and 1/2 miles to work. Because 1/2 of 15... for half an hour, he goes 7 and 1/2.

I–99 Ok, and if he goes 7 and 1/2, how far does...

S–100 She goes 7 and 1/2 too!

I–100 Its the same, yeh. Right.

S–101 Neat. That was short, yeh. That was great. Yeh! If you just find everything else, you can just find that really easy.

In her earlier solution attempts, Celeste seldom constructed algebraic expressions without assistance, and her attempts to hold algebraic terms accountable to the problem situation showed that the expressions she did construct were relatively opaque. Although Celeste is able to find precise solutions using state simulation, these solution attempts are outside the standard notational structure and quantitative ontology of algebra (i.e., role expressions) and so cannot pass as competent problem solving within the algebra curriculum. Furthermore, as with her doubling error in an iterative simulation on problem MOD, she often encounters difficulties when using nonstandard materials designs. Although constructed out of conventional notations like tables or segment diagrams that organize familiar arithmetic calculation, the "weird way out" is often difficult as well.

The tutorial demonstration that opens this section is an intervention along several dimensions that are relevant to an analysis of constructing algebraic representations as material designs. First, the demonstrated materials almost exclusively use role quantities, in contrast with Celeste's predominate use of state quantities in her spontaneous solution attempts. Second, the intervention uses the text of algebra story problems as a specification for constructing independent scenes in each quantitative dimension, unlike role scenes observed in spontaneous solution attempts, which explicitly render only a single quantitative dimension. Third, the demonstrated materials are fundamentally incomplete as instructions for how to construct an algebraic representation of related linear functions, since rates are not explicitly present as two dimensional entities (e.g., "intensives" in the domain analysis of Shalin and Bee, 1985).

Despite a shift in quantitative ontology and an incomplete design for linear functions, the demonstration and subsequent collaboration allow Celeste to gradually integrate constructing and labelling role scenes into her problem–solving activities. In doing so, she shifts from constructing standard materials, *after* an existing solution, to constructing algebraic expressions on top of scene structures as a way of finding a precise solution. Thus, the strategic arrangement of her problem–solving episodes changes as well as the material designs contained within them. Perhaps most fundamental, Celeste shifts from viewing quantities as prospective properties of states to viewing quantities as a retrospective summary of roles that are embedded in a coherent combination of material designs (e.g., Jeff's time, rather than the first hour of train travel). After a relatively brief intervention and some collaborative work in combining conventional materials, Celeste takes up the demonstration and produces a competent algebraic solution to a novel problem.

This is encouraging for an argument that material designs are constructed out of conventional notations and that competence amounts to different constructive patterns for reaching standard algebraic representations. Still, Celeste's progress remains within the collaborative scaffolding of the verbal interview setting, and her shift from state to role quantities may well be both content–specific (i.e., to motion problems) and a short–lived adaptation to the social circumstances of the interviews. The tutorial is not a study of algebra instruction, but it does show that conventions for scene notations and role quantities can be combined, even by a beginning algebra student, to reach a standard algebraic representation.

5.4 Quantitative analyses of material designs for inference

In Chapter 4, I argued that competence cannot be entirely a matter of using individual tactics, since both algebra students and teachers use the same kinds of tactics. Instead, differences in competence must be a matter of how these tactics are combined, rather like selecting from a set of conventional materials and tools to design an artifact.

Problem solvers succeed when they combine tactics in a way that connects relevant structural constraints with manipulative activities that precisely determine unknown values. These successful combinations sometimes involve algebraic expressions, but people also find solutions without any evidence of using formal algebra. Model-based solutions to algebra story problems tell us about how people manage complex quantitative relationships, but they do not match a prescriptive view of competence in the algebra curriculum.

Competent problem solvers manage to construct an effective set of algebraic expressions, either before or after solving the problem. For example, Paul (a teacher) uses algebra on request to verify a solution on problem MOD, but he finds the solution by comparing states with a ratio tactic (Section 4.3.3 at S–7). In contrast, algebra students are able to manipulate existing expressions, but they have great difficulty constructing these materials. Given some support, however, they may be able to combine available materials in a way that helps to construct standard algebraic expressions. For example, Celeste progressively adopts a demonstration of a particular combination of materials — i.e., role scenes are used to construct algebraic expressions — and with collaborative support, she reorganizes conventional aspects of her quantitative inferences into standard curricular forms.

Even at different levels of competence, problem solving is not usually a matter of progressively refining a single representation of problem structure. Instead, successive problem–solving episodes use different notational and quantitative materials, and the particular arrangement of these materials is constructed more often than recalled. This is an important empirical observation that is at odds with most existing accounts of competent quantitative inference at this level of mathematics. The constructed aspect of quantitative inference, even on seemingly innocent school math problems, raises an interesting theoretical possibility: what a problem solver appears to "know" about a mathematical concept may depend on how she structures the setting while solving any particular problem. This is the sense of a representational ecology described in the introduction to this chapter.

We can describe these constructed and organizational aspects of a representational ecology by using the notational and quantitative materials observed within individual problem–solving episodes. If competence depends on constructing a setting that affords knowing about mathematics, then problem solvers with otherwise similar backgrounds should appear more or less knowledgeable depending on how they construct particular notational structures and embed different kinds of quantities in these structures. Under an ecological interpretation, these constructions are changes to the setting, each providing a new opportunity for knowing about related linear functions.

5.4.1 Local representational ecologies

This section extends the interpretation of problem–solving episodes to look at (a) the prevalence of different material combinations of notation and quantity, (b) their comparative use by algebra students and teachers, and (c) relations between these materials and problem–solving tactics, strategies, and outcomes. The prevalence of different materials gives empirical constraints on a descriptive theory of inference, telling us which kinds of notational structures embed which kinds of quantities for what purpose. The comparative use of these materials by algebra students and teachers helps to explain what needs to be learned to achieve the curricular sense of competent algebra story problem solving.

Episodes as a unit of analysis. Quantitative comparisons reported in this section use the problem–solving episode as a unit of analysis, pooling together episodes taken from different participants on the same set of problems. Thus, contrasts between algebra students and teachers are not comparisons of individuals. However, relations between interpretive categories are similar whether comparing individual participants or comparing students or teachers using pooled episodes. For example, in Table 4.1 of Chapter 4, participants use the same tactics but do so with differing prevalence. Results are similar when comparing algebra students and teachers on the larger set of comparison problems: students and teachers again use the same tactics but use them with different prevalence (Participants × Tactics, $\chi^2(3) = 8.5$, $p < 0.04$). Although both groups use model tactics equally often, students are more likely than teachers to use annotations (29% versus 8.8% of their episodes), less likely to use ratios (6.5% versus 20.6%), and somewhat less likely to use algebra (24.2% versus 32.4%).

Quantitative analyses in this section use episodes from six "comparison problems," combining reference problems used in Chapter 4 with two additional problems, RACE and BAGELS. These are drawn from the structural class shown at the upper right of Figure 4.1 in Section 4.2, and texts for these problems are reproduced in Appendix A. Using episodes as the unit of analysis allows statistical tests of the reliability of relationships between categories — e.g., does the likelihood of a conceptual error depend on notational structure? However, the relatively large number of episodes from verbal interviews come from only four participants, so any relationships reported here are tentative. These analyses explore the material setting of competent quantitative inference, holding differences in tactical or strategic categories accountable to the actual materials that problem solvers speak about or write down while solving problems.

We can approach these questions by treating combinations of notational structure and quantitative ontology as material designs for different problem–solving tactics. A design, in this sense, is a commitment by a problem solver to a particular way of organizing and interpreting particular types of quantities. This opens up the elusive problem solving injunction to "choose a good representation," with the added provisions (a) that an effective representation of problem structure may require several such commitments, all held together at the same time, and (b) that much of the

The tutorial demonstration that opens this section is an intervention along several dimensions that are relevant to an analysis of constructing algebraic representations as material designs. First, the demonstrated materials almost exclusively use role quantities, in contrast with Celeste's predominate use of state quantities in her spontaneous solution attempts. Second, the intervention uses the text of algebra story problems as a specification for constructing independent scenes in each quantitative dimension, unlike role scenes observed in spontaneous solution attempts, which explicitly render only a single quantitative dimension. Third, the demonstrated materials are fundamentally incomplete as instructions for how to construct an algebraic representation of related linear functions, since rates are not explicitly present as two dimensional entities (e.g., "intensives" in the domain analysis of Shalin and Bee, 1985).

Despite a shift in quantitative ontology and an incomplete design for linear functions, the demonstration and subsequent collaboration allow Celeste to gradually integrate constructing and labelling role scenes into her problem–solving activities. In doing so, she shifts from constructing standard materials, *after* an existing solution, to constructing algebraic expressions on top of scene structures as a way of finding a precise solution. Thus, the strategic arrangement of her problem–solving episodes changes as well as the material designs contained within them. Perhaps most fundamental, Celeste shifts from viewing quantities as prospective properties of states to viewing quantities as a retrospective summary of roles that are embedded in a coherent combination of material designs (e.g., Jeff's time, rather than the first hour of train travel). After a relatively brief intervention and some collaborative work in combining conventional materials, Celeste takes up the demonstration and produces a competent algebraic solution to a novel problem.

This is encouraging for an argument that material designs are constructed out of conventional notations and that competence amounts to different constructive patterns for reaching standard algebraic representations. Still, Celeste's progress remains within the collaborative scaffolding of the verbal interview setting, and her shift from state to role quantities may well be both content–specific (i.e., to motion problems) and a short–lived adaptation to the social circumstances of the interviews. The tutorial is not a study of algebra instruction, but it does show that conventions for scene notations and role quantities can be combined, even by a beginning algebra student, to reach a standard algebraic representation.

5.4 Quantitative analyses of material designs for inference

In Chapter 4, I argued that competence cannot be entirely a matter of using individual tactics, since both algebra students and teachers use the same kinds of tactics. Instead, differences in competence must be a matter of how these tactics are combined, rather like selecting from a set of conventional materials and tools to design an artifact.

Problem solvers succeed when they combine tactics in a way that connects relevant structural constraints with manipulative activities that precisely determine unknown values. These successful combinations sometimes involve algebraic expressions, but people also find solutions without any evidence of using formal algebra. Model-based solutions to algebra story problems tell us about how people manage complex quantitative relationships, but they do not match a prescriptive view of competence in the algebra curriculum.

Competent problem solvers manage to construct an effective set of algebraic expressions, either before or after solving the problem. For example, Paul (a teacher) uses algebra on request to verify a solution on problem MOD, but he finds the solution by comparing states with a ratio tactic (Section 4.3.3 at S–7). In contrast, algebra students are able to manipulate existing expressions, but they have great difficulty constructing these materials. Given some support, however, they may be able to combine available materials in a way that helps to construct standard algebraic expressions. For example, Celeste progressively adopts a demonstration of a particular combination of materials — i.e., role scenes are used to construct algebraic expressions — and with collaborative support, she reorganizes conventional aspects of her quantitative inferences into standard curricular forms.

Even at different levels of competence, problem solving is not usually a matter of progressively refining a single representation of problem structure. Instead, successive problem–solving episodes use different notational and quantitative materials, and the particular arrangement of these materials is constructed more often than recalled. This is an important empirical observation that is at odds with most existing accounts of competent quantitative inference at this level of mathematics. The constructed aspect of quantitative inference, even on seemingly innocent school math problems, raises an interesting theoretical possibility: what a problem solver appears to "know" about a mathematical concept may depend on how she structures the setting while solving any particular problem. This is the sense of a representational ecology described in the introduction to this chapter.

We can describe these constructed and organizational aspects of a representational ecology by using the notational and quantitative materials observed within individual problem–solving episodes. If competence depends on constructing a setting that affords knowing about mathematics, then problem solvers with otherwise similar backgrounds should appear more or less knowledgeable depending on how they construct particular notational structures and embed different kinds of quantities in these structures. Under an ecological interpretation, these constructions are changes to the setting, each providing a new opportunity for knowing about related linear functions.

5.4.1 Local representational ecologies

This section extends the interpretation of problem–solving episodes to look at (a) the prevalence of different material combinations of notation and quantity, (b) their comparative use by algebra students and teachers, and (c) relations between these materials and problem–solving tactics, strategies, and outcomes. The prevalence of different materials gives empirical constraints on a descriptive theory of inference, telling us which kinds of notational structures embed which kinds of quantities for what purpose. The comparative use of these materials by algebra students and teachers helps to explain what needs to be learned to achieve the curricular sense of competent algebra story problem solving.

Episodes as a unit of analysis. Quantitative comparisons reported in this section use the problem–solving episode as a unit of analysis, pooling together episodes taken from different participants on the same set of problems. Thus, contrasts between algebra students and teachers are not comparisons of individuals. However, relations between interpretive categories are similar whether comparing individual participants or comparing students or teachers using pooled episodes. For example, in Table 4.1 of Chapter 4, participants use the same tactics but do so with differing prevalence. Results are similar when comparing algebra students and teachers on the larger set of comparison problems: students and teachers again use the same tactics but use them with different prevalence (Participants × Tactics, $\chi^2(3) = 8.5$, $p < 0.04$). Although both groups use model tactics equally often, students are more likely than teachers to use annotations (29% versus 8.8% of their episodes), less likely to use ratios (6.5% versus 20.6%), and somewhat less likely to use algebra (24.2% versus 32.4%).

Quantitative analyses in this section use episodes from six "comparison problems," combining reference problems used in Chapter 4 with two additional problems, RACE and BAGELS. These are drawn from the structural class shown at the upper right of Figure 4.1 in Section 4.2, and texts for these problems are reproduced in Appendix A. Using episodes as the unit of analysis allows statistical tests of the reliability of relationships between categories — e.g., does the likelihood of a conceptual error depend on notational structure? However, the relatively large number of episodes from verbal interviews come from only four participants, so any relationships reported here are tentative. These analyses explore the material setting of competent quantitative inference, holding differences in tactical or strategic categories accountable to the actual materials that problem solvers speak about or write down while solving problems.

We can approach these questions by treating combinations of notational structure and quantitative ontology as material designs for different problem–solving tactics. A design, in this sense, is a commitment by a problem solver to a particular way of organizing and interpreting particular types of quantities. This opens up the elusive problem solving injunction to "choose a good representation," with the added provisions (a) that an effective representation of problem structure may require several such commitments, all held together at the same time, and (b) that much of the organizational structure of these designs is available in the problem–solving setting — i.e., constructed on paper, in space, and across time by the problem solver.

Under an ecological interpretation, particular material designs should support problem–solving activities differently. In an analysis of problem–solving episodes, this appears in the comparative use of tactical materials for different purposes (i.e., comprehension or solution attempts) and with different outcomes (e.g., transitioning out of episodes, making structural inferences, etc.). When material designs are only or never used for some purpose, we have a clear constraint on the kinds of activities that they afford their users. When the prevalence of material combinations differs across levels of competence (i.e., algebra students versus teachers), we have constraints on what must be learned to construct and interpret quantitative inferences in the problem–solving setting. Hence, the broad interpretive questions are: how do different material designs support quantitative inference and (b) how can differences in these designs be used to explain competence and how it might be acquired?

Problem solvers, problems, and materials for design. If people are constructing material designs when given algebra story problems, we can start an exploration of representational ecologies by looking for material differences between levels of competence, on the one hand, and between different aspects of problem structure, on the other. While earlier analyses showed that people with different mathematical backgrounds use the same tactics, they may not combine notational and quantitative materials in the same way when pursuing these tactics. Likewise, material designs may vary with different aspects of problem structure — i.e., the content domain of the problem (motion or work) or its normative structure (relations within and across dimensions).

As a baseline observation, algebra students and teachers use each type of notational structure (i.e., expressions, narrative, 2D and scene) and each kind of quantity (i.e., state and role). Thus, competence cannot be a simple matter of *knowing about* a particular notation or type of quantity. Instead, students and teachers use these materials in different ways. Considering only the prevalence of different notations, students use fewer expressions, less narrative material, and many more 2D notations than teachers (Notational structure × Participant, $\chi^2(3) = 12.6$, $p < .006$). Both groups use role quantities more frequently than state quantities (Quantitative ontology × Participant, not significant).

Algebra students and teachers construct different material designs within problem–solving episodes. Table 5.1 shows the percentage of episodes with different material combinations for students, teachers, and both groups combined.[7] Students frequently use 2D notations to carry state quantities (e.g., Celeste's "weird way out" of problem MOD in Section 4.3.1 of Chapter 4), while teachers never use 2D state tables and seldom use 2D role tables (e.g., Richard's "chart" on problem MOD in

[7]Percentages are calculated within panels for each group — e.g., 22.6% of students' episodes use a 2D table to carry state quantities.

Table 5.2: Notational structure and quantitative ontology in episodes from problems with different content (motion versus work).

	Problem Content (number of episodes)	
Materials	Motion (n = 62)	Work (n = 34)
Notational structure*		
Expression	30.6	41.2
Narrative	9.7	26.5
2D	30.6	20.6
Scene	29.0	11.8
Quantitative ontology**		
State	29.0	58.8
Role	71.0	41.2

Use of materials depends on content domain: $^{*}p < .04$, $^{**}p < .01$.

only a small collection of algebraic schemata that cover a formal view of the most common categories of algebra story problems. With the exception of 2D state calculations, however, algebra students and teachers resemble each other more than either group resembles a normative account of remembering mathematical forms — e.g., characteristic formulas and procedures for "extracting the appropriate equations from the text" (Hinsley *et al.*, 1977, p. 98). While both groups recall formulas and other problem–specific information (e.g., "dirt" or "chart" annotations for motion problems), much of their episodic activity uses quite different materials (e.g., narrative, 2D, or scene notations carrying state quantities).

Materials appearing in problem solving episodes may also depend upon characteristics of the problem being solved. There are no reliable differences in the materials used for different problem structures (i.e., problem classes in Figure 4.1 of Chapter 4), but the content domain of these problems does influence which materials appear within episodes. Table 5.2 shows the percentage of episodes using different materials for motion or work problems. In the top panel, complex written notations (i.e., 2D and scene) are more common for motion problems, while narratives and standard expressions are more common for work problems. In a separate analysis (not shown), this pattern is stronger among students than teachers. The lower panel of Table 5.2 shows that work problems most often elicit state quantities (58.8%), while motion problems most often elicit role quantities (71%). A separate analysis shows that this pattern is strongest among teachers.

The appearance of local materials across problem solvers and problems is consistent with an ecological hypothesis that problem solvers construct quantitative inferences, rather than simply recalling standard mathematical forms and applying them to a problem. There may be no reliable differences between materials constructed across structural classes of problems because these structures are not apparent to problem solvers as they work through a solution attempt. That is, problem solvers are not recalling the structure of problems so much as they are constructing inferences about problem structure within the material setting. This is not to say that problem solvers never recall materials that are useful to them, but that they seldom recall a complete quantitative structure. As the analysis of spontaneous problem comparisons in Chapter 4 demonstrated (see Figure 4.5), recalled material need not even contain a correct quantitative inference to be useful.

Both algebra students and teachers use complex notational structures, and these tend to follow the content domain of the problem being solved. When problems involve collinear motion, problem solvers often construct scenes out of directed segments and are then able to coordinate quantities, inferences about quantitative relations, and calculation around the components of these scenes. Output in work problems less directly resembles collinear segments, and there is a corresponding increase in alternative materials when people construct and pursue quantitative inferences. This is particularly true of narrative material, which increases threefold over that found with motion problems (algebraic expressions also increase). The twofold increase in state quantities for work problems can be interpreted similarly, since problem solvers often reason about quantities within the structure of events (e.g., the amount of a pool filled after an hour in problem HOS) rather than recalling standard algebraic expressions carrying role quantities.

The material design of problem–solving episodes tracks the content of problems, and their construction changes with one's relation to the algebra curriculum. These analyses of episodic materials expand our descriptive account of competence in several important ways: (a) successful problem solving *requires* nonstandard materials, regardless of one's level of competence, (b) problem solving consists of activities that construct and interpret these materials, and (c) one apparently becomes competent by learning to combine nonstandard and standard algebraic materials while constructing a problem representation. The sections that follow examine how conventional materials are used to construct tactics and whether these material designs shape problem–solving outcomes by determining what activities a problem solver can attempt and what inferences they can make about problem structure.

5.4.2 Making: (re)constructing problem–solving tactics

As an interpretive category, problem–solving tactics imply a particular view of problem structure — e.g., as a model of train separation or as a collection of related algebraic terms for problem MOD. The notational and quantitative materials introduced in this

Table 5.3: Notations and quantities used during problem–solving tactics.

Tactics (episodes)	Quantitative ontology	Notational structure: Expressions	Narrative	2D	Scene
Algebra	State	7.7	0.0	0.0	0.0
(n = 26)	Role	92.3	0.0	0.0	0.0
Model*	State	0.0	13.2	34.2	23.7
(n = 38)	Role	0.0	7.9	0.0	21.1
Ratio	State	36.4	36.4	0.0	0.0
(n = 11)	Role	9.1	18.2	0.0	0.0
Annotation	State	0.0	0.0	4.8	0.0
(n = 21)	Role	9.5	4.8	57.1	23.8

Notational structures carry different kinds of quantity: $^*p < .02$.

chapter allow us to ask how these tactical viewpoints are constructed. In particular, how are nonstandard materials constructed, how are they combined with standard materials to form a representation of problem structure, and how are problem–solving outcomes shaped by these constructive activities? When problem solvers work outside the algebraic formalism, notations and quantities are not combined without intention or in a completely idiosyncratic fashion. Instead, these materials are used to construct different problem–solving tactics and strategies as problem solvers work to span the gap between text and equations (or precise values).

Table 5.3 combines episodes from students and teachers to show how different notations and quantities are used to construct problem–solving tactics. Separate analyses show that students and teachers construct and use tactics in different ways, and differences that appear to be reliable are discussed in the text. The first question is how notation and quantity are combined to construct what we interpret as problem–solving tactics.

- Algebraic tactics exclusively use expressions, both among algebra students and teachers. These usually involve quantitative roles but sometimes compare states in an algebraic proportion.
- Model tactics never use arithmetic or algebraic expressions alone. Instead, they use narration, 2D tables, or scenes to organize quantitative inferences, usually about state quantities (71.1% of model episodes use state quantities). Within these notations, students and teachers construct model tactics very differently.

  - The majority of students' model tactics (52%) are simulations that organize state quantities in 2D tables. Students also use scenes, most often to depict relations between role quantities.
  - In contrast, teachers never use 2D notations to construct model tactics. The majority of their model tactics use scenes (69.2%), most often to depict relations between state quantities.

- Ratio tactics never use 2D or scene notations, use narratives more often than written arithmetic expressions (54.6% of ratio episodes), and most often involve a state comparison (72.8% of ratio episodes) rather than a comparison between role quantities.
- Annotation tactics almost exclusively use quantitative roles (95.2% of annotations), and these are most often embedded in a 2D table (e.g., Karen's "dirt" table or Richard's "chart" on problem MOD). Next most common are diagrammatic scenes carrying role quantities (23.8% of annotations). As observed earlier, annotations of any sort are rare among algebra teachers (3 of their 34 episodes).

There are two striking patterns in the material design of problem–solving tactics, and they are found across levels of competence. First, algebra and ratio tactics never use 2D tables or scenes; instead they use either expressions or narrative materials alone. In contrast, model and annotation tactics most often rely on more complex notational structures like 2D tables or scenes, and they seldom use narratives or expressions alone to carry quantities. When people work outside the standard algebraic formalism, the material design of tactics adds layers of configural structure to standard notations for quantitative relations. The second major pattern cuts across notational structures: model and ratio tactics most often involve state quantities, while algebra and annotation tactics most often involve role quantities. Since ratio tactics use standard expressions to compare state quantities, a prospective, partitioned sense of quantity spans both standard and nonstandard approaches to quantitative inference.

Although algebra students and teachers both use model–based reasoning, they construct this tactic out of different materials. As in Chapter 4, students often use model tactics to find precise values by organizing state calculations as a form of iterative or heuristic simulation. 2D tables of state quantities are the medium out of which these calculations are usually constructed, although state scenes can also accommodate simulation. In contrast, teachers most often use model tactics to generate or evaluate quantitative relations, and they seldom use these tactics to calculate precise values. Scenes carrying state quantities are a material design for these explorations.

Table 5.4: Notations and quantities used during problem–solving strategies.

Strategy (episodes)	Quantitative ontology	Notational structure: Expressions	Narrative	2D	Scene
Comprehension**	State	0.0	12.9	6.5	8.1
(n = 62)	Role	24.2	8.1	19.4	21.0
Solution attempt**	State	19.2	0.0	34.6	11.5
(n = 26)	Role	34.6	0.0	0.0	0.0
Verify	State	12.5	12.5	12.5	12.5
(n = 6)	Role	37.5	12.5	0.0	0.0

Notational structures carry different kinds of quantity: $^{**}p < .005$.

5.4.3 Doing: material designs, strategies and manipulative capacity

The interpretive framework used in preceding chapters assumes that people use tactics to achieve some purpose or strategy. That is, a tactical view of problem structure supports activities that have consequences for problem–solving outcomes. Just how this occurs is at the heart of the theoretical contrast carried through this chapter: whether activity is determined *a priori* by what one knows (i.e., one's mental representation), or whether setting and activity interact to influence what one knows. The materials produced within an episode are important for either account of strategies used to make quantitative inferences and solve problems. From an ecological perspective, the material design of tactics afford different activities and different understandings of problem structure. The first can be found in relations between materials and problem–solving strategies (examined in this section); the second can be found in relations between materials and problem–solving outcomes — i.e., episode transitions, errors, and the episodic course of inferences (examined in the next section).

The relation between materials and strategies is also important for describing differences in competence. Unlike problem–solving tactics, there are no reliable differences between algebra students and teachers in the prevalence of various strategies (Participant × Strategy, $\chi^2(2) = 1.9$, not significant). Thus, competent quantitative inference cannot be attributed to simple differences in the adoption of strategies like routinely verifying determined values. A more useful contrast may be found by examining how materials support problem–solving strategies.

Table 5.4 shows which combinations of notation and quantity are used during episodes with different strategies (i.e., comprehension, solution attempt, or verification). Again, separate analyses show that students and teachers differ, and these contrasts are discussed in the text.

- Comprehension episodes use every type of notation and most often involve role quantities (72.6% of comprehension episodes).
  - When scene notations are used for comprehension, students almost always examine role quantities (90.9%), while teachers most often examine state quantities (57.1%).
- Solution attempts most often involve expressions (53.8%), never use narrative materials alone, and never use 2D tables or diagrammatic scenes to carry role quantities. These are common constraints on finding precise values, but students and teachers use different materials for solution attempts.
  - 90.9% of teachers' solution attempts use expressions, and these more often carry role quantities (54.5% of solution attempts).
  - 60% of students' solution attempts use 2D tables to organize state quantities, followed by role expressions (20%) and state scenes (13.3%).
- Verification episodes are infrequent (less than 10% of all episodes) and show no striking material patterns in either group.

Algebra students regularly use 2D tables of state quantities to organize calculations during solution attempts, something never done by teachers. Algebra teachers sometimes undertake extended state calculations (e.g., Paul's retrospective simulation using a diagrammatic scene on problem MOD, Section 4.3.3), but seldom as a solution attempt. These contrasts are corroborated in Table 5.5, which shows how students and teachers use material designs to finalize solution attempts. While the number of episodes is small, students clearly attempt solutions using more varied notations, including 2D tables and scenes, while teachers exclusively use standard algebraic expressions. Although not statistically reliable, students usually attempt to find precise solutions using state quantities, while teachers most often use role quantities.

Apart from differences in the way that students and teachers attempt solutions, it is interesting that of all episodes involving role quantities, only written expressions are used in solution attempts. Narrative material, 2D tables, and diagrammatic scenes never carry role quantities in a way that supports a solution attempt; instead, these material combinations are used primarily as a comprehension strategy. These quantitative findings corroborate qualitative comparisons made in Section 5.3, where material designs were seen to support strategic activities very differently. Role expressions and 2D tables of state quantities both have conventional and effective manipulative capacities for determining precise values during solution attempts (e.g.,

  - The majority of students' model tactics (52%) are simulations that organize state quantities in 2D tables. Students also use scenes, most often to depict relations between role quantities.
  - In contrast, teachers never use 2D notations to construct model tactics. The majority of their model tactics use scenes (69.2%), most often to depict relations between state quantities.

- Ratio tactics never use 2D or scene notations, use narratives more often than written arithmetic expressions (54.6% of ratio episodes), and most often involve a state comparison (72.8% of ratio episodes) rather than a comparison between role quantities.
- Annotation tactics almost exclusively use quantitative roles (95.2% of annotations), and these are most often embedded in a 2D table (e.g., Karen's "dirt" table or Richard's "chart" on problem MOD). Next most common are diagrammatic scenes carrying role quantities (23.8% of annotations). As observed earlier, annotations of any sort are rare among algebra teachers (3 of their 34 episodes).

There are two striking patterns in the material design of problem-solving tactics, and they are found across levels of competence. First, algebra and ratio tactics never use 2D tables or scenes; instead they use either expressions or narrative materials alone. In contrast, model and annotation tactics most often rely on more complex notational structures like 2D tables or scenes, and they seldom use narratives or expressions alone to carry quantities. When people work outside the standard algebraic formalism, the material design of tactics adds layers of configural structure to standard notations for quantitative relations. The second major pattern cuts across notational structures: model and ratio tactics most often involve state quantities, while algebra and annotation tactics most often involve role quantities. Since ratio tactics use standard expressions to compare state quantities, a prospective, partitioned sense of quantity spans both standard and nonstandard approaches to quantitative inference.

Although algebra students and teachers both use model-based reasoning, they construct this tactic out of different materials. As in Chapter 4, students often use model tactics to find precise values by organizing state calculations as a form of iterative or heuristic simulation. 2D tables of state quantities are the medium out of which these calculations are usually constructed, although state scenes can also accommodate simulation. In contrast, teachers most often use model tactics to generate or evaluate quantitative relations, and they seldom use these tactics to calculate precise values. Scenes carrying state quantities are a material design for these explorations.

Table 5.4: Notations and quantities used during problem–solving strategies.

Strategy (episodes)	Quantitative ontology	Notational structure: Expressions	Narrative	2D	Scene
Comprehension** (n = 62)	State	0.0	12.9	6.5	8.1
	Role	24.2	8.1	19.4	21.0
Solution attempt** (n = 26)	State	19.2	0.0	34.6	11.5
	Role	34.6	0.0	0.0	0.0
Verify (n = 6)	State	12.5	12.5	12.5	12.5
	Role	37.5	12.5	0.0	0.0

Notational structures carry different kinds of quantity: $^{**}p < .005$.

5.4.3 Doing: material designs, strategies and manipulative capacity

The interpretive framework used in preceding chapters assumes that people use tactics to achieve some purpose or strategy. That is, a tactical view of problem structure supports activities that have consequences for problem–solving outcomes. Just how this occurs is at the heart of the theoretical contrast carried through this chapter: whether activity is determined *a priori* by what one knows (i.e., one's mental representation), or whether setting and activity interact to influence what one knows. The materials produced within an episode are important for either account of strategies used to make quantitative inferences and solve problems. From an ecological perspective, the material design of tactics afford different activities and different understandings of problem structure. The first can be found in relations between materials and problem–solving strategies (examined in this section); the second can be found in relations between materials and problem–solving outcomes — i.e., episode transitions, errors, and the episodic course of inferences (examined in the next section).

The relation between materials and strategies is also important for describing differences in competence. Unlike problem–solving tactics, there are no reliable differences between algebra students and teachers in the prevalence of various strategies (Participant × Strategy, $\chi^2(2) = 1.9$, not significant). Thus, competent quantitative inference cannot be attributed to simple differences in the adoption of strategies like routinely verifying determined values. A more useful contrast may be found by examining how materials support problem–solving strategies.

Table 5.4 shows which combinations of notation and quantity are used during episodes with different strategies (i.e., comprehension, solution attempt, or verification). Again, separate analyses show that students and teachers differ, and these contrasts are discussed in the text.

- Comprehension episodes use every type of notation and most often involve role quantities (72.6% of comprehension episodes).
  - When scene notations are used for comprehension, students almost always examine role quantities (90.9%), while teachers most often examine state quantities (57.1%).
- Solution attempts most often involve expressions (53.8%), never use narrative materials alone, and never use 2D tables or diagrammatic scenes to carry role quantities. These are common constraints on finding precise values, but students and teachers use different materials for solution attempts.
  - 90.9% of teachers' solution attempts use expressions, and these more often carry role quantities (54.5% of solution attempts).
  - 60% of students' solution attempts use 2D tables to organize state quantities, followed by role expressions (20%) and state scenes (13.3%).
- Verification episodes are infrequent (less than 10% of all episodes) and show no striking material patterns in either group.

Algebra students regularly use 2D tables of state quantities to organize calculations during solution attempts, something never done by teachers. Algebra teachers sometimes undertake extended state calculations (e.g., Paul's retrospective simulation using a diagrammatic scene on problem MOD, Section 4.3.3), but seldom as a solution attempt. These contrasts are corroborated in Table 5.5, which shows how students and teachers use material designs to finalize solution attempts. While the number of episodes is small, students clearly attempt solutions using more varied notations, including 2D tables and scenes, while teachers exclusively use standard algebraic expressions. Although not statistically reliable, students usually attempt to find precise solutions using state quantities, while teachers most often use role quantities.

Apart from differences in the way that students and teachers attempt solutions, it is interesting that of all episodes involving role quantities, only written expressions are used in solution attempts. Narrative material, 2D tables, and diagrammatic scenes never carry role quantities in a way that supports a solution attempt; instead, these material combinations are used primarily as a comprehension strategy. These quantitative findings corroborate qualitative comparisons made in Section 5.3, where material designs were seen to support strategic activities very differently. Role expressions and 2D tables of state quantities both have conventional and effective manipulative capacities for determining precise values during solution attempts (e.g.,

Table 5.5: Notations and quantities used during final episodes.

Materials	Participants (number of episodes) Students (n = 11)	Teachers (n = 8)
Notational structure*		
Expression	45.5	100.0
Narrative	0.0	0.0
2D	36.4	0.0
Scene	18.2	0.0
Quantitative ontology		
State	63.6	37.5
Role	36.4	62.5

Use of notations depends on participant, $^{*}p < .05$.

accumulating hours until trains are 880 km apart). In contrast, 2D tables or diagrammatic scenes that carry role quantities have no conventional manipulative capacity for determining precise values.

The relation between problem–solving tactics and strategies depends on what people construct inside the material setting of an episode: distinct notational structures that carry different senses of quantity. These constructions outline a chasm as one moves from nonstandard to standard designs for quantitative inference, a progression that involves two material boundaries: from more to less complex notational structures and from state to role quantities. Each combination of materials offers a particular view of problem structure and each supports particular activities. For algebra story problem solving, nonstandard materials outside the algebraic formalism (a standard material design) are *required* because neither material alone is sufficient for constructing and using a mathematical representation. For example, constructing a role scene can promote otherwise implicit relationships between quantities (e.g., Karen's inference about composite distances on problem MOD, Section 4.3.2), but some other material design will be required in order to find precise values for related quantities (e.g., Karen's 2D table of state calculations on problem MOD).

Algebra students and teachers sometimes construct different material designs for model tactics, as in students' exclusive use of 2D tables of state quantities to organize model–based simulations. These material differences are understandable by looking beyond local episode boundaries: students' and teachers' activities diverge primarily at the point of constructing algebraic expressions, before these expressions are manipulated to determine precise values. Constructing a coherent collection of algebraic expressions is difficult across levels of competence, but is particularly so for students, who can generally manipulate given expressions without errors. On the informal side of this material chasm, algebra students work around demands for algebraic precision by constructing alternative materials. These nonstandard material designs carry given and calculated values in a notational structure that supports precise calculation, and they become the representational ecology of model–based simulation. Algebra teachers also construct state simulations, but they seldom use these materials to find precise solutions without standard algebraic materials. These differences agree with the qualitative and quantitative comparisons of Chapter 4.

5.4.4 Knowing: local outcomes and the origin of inferences

Relations between materials (notations and quantity), tactics, and strategy describe the setting for problem–solving activities. That materials afford different opportunities for activity is one part of the evidence for viewing participants' solution attempts as the construction of a representational ecology. Another line of evidence for this view of problem solving is the influence of materials on problem–solving outcomes. That is, not only does the setting determine what a person can do, but its material form also influences what a person appears to know about problem structure. One place to look for evidence for this hypothesis is in empirical relations between materials and local difficulties around episode boundaries. These include subjective transitions out of episodes (i.e., a participant's sense of being on or off track) and various types of errors that influence activities within the episode (i.e., conceptual errors of omission or commission and manipulative errors).

Materials and local outcomes in problem solving. Particular material designs contribute disproportionately to local difficulties. Table 5.6 shows the percentage of episodes with various local outcomes while participants are using different material combinations of notation and quantity. The slightly unorthodox convention of staggering table entries allows rows and columns to be read more easily as comparisons of the likelihood that different material designs will lead to difficulties. Reading across the first panel, we can compare different materials for the likelihood of transition off track; reading down a column, we can compare the relative strengths and weaknesses of a particular material design for encountering local difficulties. For example, expressions and scenes both carry role quantities, and we might wonder which notational structure is more likely to lead to local difficulties with this kind of quantity. Reading across the third row of each panel (i.e., Off track, Omission, etc.) shows that subjective and conceptual difficulties are much more likely when role quantities are carried by standard algebraic expressions (e.g., 59.3% of episodes with role expressions end off track). Again, differences between algebra students and teachers are discussed in the text.

- Participants are most likely to transition off track when using expressions or 2D notations and least likely to do so when using narrative or scene notations (Notational structure × Transition, $\chi^2(3) = 11.8$, $p < .009$).

- Type of quantity alone has no reliable relation to local transitions, but the contrast between expressions and scenes is strongest for role quantities (i.e., role expressions are 7.7 times more likely than role scenes to end off track).
- The relation between notational structure and transitional difficulty is similar across groups, though teachers most often have difficulties with expressions (87.5% of their off track episodes), while students most often have difficulties with 2D tables (46.4% of their off track episodes).

- Omission errors are also most likely with expressions and 2D notations, but the relationship is not statistically reliable. Again, role scenes are comparatively resistant to local difficulties.
- Commission errors never occur when participants use scene or narrative notations but are most prevalent when they use expressions (Notational structure × Transition, $\chi^2(3) = 15.9$, $p < .002$).
  - As with transitions off track, role expressions are particularly error prone.
  - The relation between notational structure and errors of commission is strongest among students. A similar pattern holds for teachers, who seldom use 2D notations, never use these to carry state quantities, and never encounter local difficulties of any kind with this notation.
- Manipulation errors are generally infrequent and are not related to the materials used by students or teachers.

Local difficulties are least likely when problem solvers use scene or narrative materials and are most likely when they use expressions or 2D notations. This pattern holds for algebra students and teachers, though students alone have difficulty with 2D notations. From an ecological view, how one constructs the material setting of problem solving not only influences activity (i.e., what strategy one pursues) but also influences what one appears to know about problem structure (i.e., the likelihood of local difficulties and conceptual errors).

Within this broad pattern, there are two interesting contrasts in the way notations carry different kinds of quantity. First, expressions and scenes carry role quantities with very different local outcomes, expressions much more frequently leading people off track or showing conceptual errors. However, this does not appear to be the case when these notations carry state quantities. One explanation for this contrast is that quantitative relations are more easily inferred when the surrounding notation provides strong configural constraints (i.e., the iconic properties of scenes discussed in Section 5.3) or when the ontology of quantities is embedded in the situation depicted in the problem (i.e., a prospective, partitioned sense of quantity). Thus, the relative opacity of expressions may be reduced when they carry state quantities.

Second and exclusively among students, 2D notations lead to different kinds of local difficulties when they carry different kinds of quantities. Students are most likely to

Table 5.6: Percentage of episodes with local difficulties while using different materials.

Materials: Notational Structure by Quantitative Ontology

	Expressions (33)†		Narrative (15)		2D (26)		Scene (22)	
Difficulty	State (6)	Role (27)	State (9)	Role (6)	State (14)	Role (12)	State (9)	Role (13)
Off track*	51.5		20.0		50.0		13.6	
	16.7		22.2		50.0		22.2	
		59.3		16.7		50.0		7.7
Omission	21.2		0.0		23.1		9.1	
	16.7		0.0		7.1		11.1	
		22.2		0.0		41.7		7.7
Commission*	36.4		0.0		19.2		0.0	
	0.0		0.0		28.6		0.0	
		44.4		0.0		8.3		0.0
Manipulation	9.1		0.0		3.8		4.5	
	33.3		0.0		7.1		11.1	
		3.7		0.0		0.0		0.0

†Values in parentheses are the number of episodes using each material category.
*Local outcome depends on notational structure, $p < .009$.

leave out relevant constraints (i.e., conceptual errors of omission) when using a 2D table to organize role quantities. In contrast, they are most likely to introduce incorrect constraints (i.e., conceptual errors of commission) when using 2D tables to organize state calculations. In keeping with earlier arguments about the affordance of different notational structures, 2D role tables provide typological constraints on quantities (e.g., distance versus time), but they provide few (if any) constraints on how or whether embedded quantities should be related. This may explain why conceptual errors of omission predominate with these material designs, particularly among students who may not appreciate the relation between "dirt" tables and recurring algebraic formulas. Similarly, 2D state tables provide weak constraints for inferring quantitative relations, but they do provide an organizational structure for the calculations involved in state simulation. However, in order to use these material designs to calculate precise values, students must make commitments to quantitative relations that are not required when constructing and filling a 2D role table. This may explain why conceptual errors of commission are so common with 2D state tables.

Do local difficulties simply occur whenever problem solvers attempt solutions? Although manipulative errors are most common during solution attempts (Strategy × Manipulative errors, $\chi^2(2) = 9.7$, $p < .005$), there are no reliable relations between strategy and transitioning off track or having a conceptual error. Thus, the material basis of local difficulties cannot be attributed to the difficulty of attempting solutions. Instead, there is a tradeoff between configural support for relevant inferences versus manipulative capacity: what a problem solver might know and do in a particular material setting. As shown in Table 5.4, narratives, role scenes, and 2D role tables are never used in solution attempts. But even when material designs cannot be used in solution attempts, there are still marked advantages for particular notational structures. Narratives never lead to conceptual or manipulative errors, local difficulties which every other material design encounters. Scenes are somewhat less likely than narratives to lead problem solvers off track, but both narrative and scene notations lead to fewer local difficulties than 2D tables when these notations carry role quantities.

Patterns in local difficulties can also be understood as a contrast between standard (explicitly schooled) and nonstandard material designs for making quantitative inferences. As nonstandard material designs, scene and narrative structures are least likely to produce conceptual errors, regardless of the kinds of quantities they carry. Thus, problem–solving episodes outside of the algebraic formalism are least often responsible for conceptual errors. In contrast, more than two thirds of conceptual errors appear with standard role expressions or 2D role tables (e.g., Karen's "dirt" chart), and these materials are more likely than any others to involve conceptual errors. 2D state tables, used exclusively by students and suitable for precise calculation, involve fewer conceptual errors than standard role expressions but more conceptual errors than either scenes or narrative materials. Thus, the likelihood of local difficulties increases as one moves towards materials that are designed more for quantitative precision than for quantitative inference.

error; both are counted as "repairs." Similar to the presentation of local outcomes in the last section, the percentage of episodes introducing inferences while using particular materials are arranged to facilitate relevant comparisons. Thus, reading across the first panel, we can compare different materials for the origin of correct constraints; reading down a column, we can compare the relative strengths and weaknesses of a particular material design (i.e., the relatively likelihood of introducing correct constraints, conceptual errors, or structural repairs). Differences between algebra students and teachers are again reported in the text.

- Correct inferences about problem structure (e.g., that times are equal in problem MOD) most often originate during episodes in which problem solvers produce narrative materials and least often originate when they use standard expressions (Notational structure × Correct constraint, $\chi^2(3) = 10.1$, $p < .02$).
  - The same pattern holds for students and teachers, though the advantage for narrative material is more pronounced among students (i.e., 6 of their 7 narrative episodes introduce correct constraints, 85.7%).
  - Students produce 24 of the 26 episodes containing 2D notations. Among these, 2D role tables more frequently introduce correct constraints.
- Conceptual errors of omission or commission originate almost exclusively during episodes in which problem solvers use standard expressions or 2D tables (Notational structure × Conceptual error, $\chi^2(3) = 14.1$, $p < .003$). For scenes carrying role quantities, the only error of omission occurs when a student (Karen) draws a diagram of round trip distance segments that are not equal on problem MRT.
  - Expressions frequently introduce conceptual errors in both groups, and these are the only conceptual errors among teachers (6 of their 15 episodes using expressions, 40%).
  - Again, students use more 2D notations than teachers, and they alone introduce conceptual errors with these materials (9 of their 24 episodes using 2D tables, 37.5%). The majority of these originate in 2D role tables (i.e., Karen's "dirt" or "trw" tables).
  - Although not statistically reliable, notations carrying role quantities are more likely to introduce conceptual errors than those carrying state quantities (Quantitative ontology × Conceptual errors, $\chi^2(1) = 2.5$, $p < .12$).
- Repair of conceptual errors does not reliably depend on notational structure or quantitative ontology. However, a comparison of expressions and 2D tables shows a marked asymmetry in the likelihood of repair: only role expressions repair prior conceptual errors, while only state tables repair these errors.

Standard materials of algebra instruction, expressions and 2D tables carrying role quantities, are the loss leaders for making quantitative inferences about problem structure. This is particularly true of role expressions, which seldom introduce correct constraints (29.6%) and more frequently introduce conceptual errors (40.7%). To their credit, role expressions are as likely as other materials to support structural repairs. The contributions of role expressions are similar for both algebra students and teachers. Among students, 2D tables carrying role quantities are as likely to introduce correct constraints as they are to introduce conceptual errors (41.7% for each), and they never support structural repairs.

In sharp contrast, nonstandard material designs are generally much more effective for making quantitative inferences about problem structure. Regardless of the type of quantity involved, narrative episodes are more likely than any other materials to introduce correct constraints, they never introduce conceptual errors, and they sometimes support structural repairs. Scene notations are somewhat less effective at introducing correct constraints, but they seldom introduce conceptual errors and are comparable to or surpass other materials for supporting structural repairs. In comparison with 2D role tables, tables carrying state quantities less often introduce correct constraints, but they also less often introduce conceptual errors and sometimes repair prior conceptual errors. For complex written materials, these findings corroborate qualitative observations of material designs made in Section 5.3 and uncover a surprising contribution from narrative episodes, for both algebra students and teachers.

5.5 Discussion

This chapter uses materials found within episodes to reconsider problem solving as the construction of material designs for quantitative inference. By adding material categories for notational structure and quantitative ontology to the episodic framework of preceding chapters, we obtain an interesting methodological synergy: (a) the episodic framework provides an interpretive vocabulary for detecting constructive inferences and their combination during solution attempts, and (b) the materials contained within these episodes provide a novel descriptive account of the problem–solving setting. Qualitative and quantitative analyses examine the relation between material designs and the interpretive categories of Chapters 3 and 4: problem–solving tactics, strategies, outcomes, and the origin of inferences. Each material design provides a partial view of problem structure, and individual designs must usually be combined to construct an effective mathematical representation. An ecological analysis of interactions between the inferential and manipulative capacity of different material designs helps to explain what people do and know when demonstrating mathematical competence. This analysis identifies different constructive packages for algebra story problem solving that have implications for theories of competent quantitative inference and instruction.

structure. This is particularly true of role expressions, which seldom introduce correct constraints (29.6%) and more frequently introduce conceptual errors (40.7%). To their credit, role expressions are as likely as other materials to support structural repairs. The contributions of role expressions are similar for both algebra students and teachers. Among students, 2D tables carrying role quantities are as likely to introduce correct constraints as they are to introduce conceptual errors (41.7% for each), and they never support structural repairs.

In sharp contrast, nonstandard material designs are generally much more effective for making quantitative inferences about problem structure. Regardless of the type of quantity involved, narrative episodes are more likely than any other materials to introduce correct constraints, they never introduce conceptual errors, and they sometimes support structural repairs. Scene notations are somewhat less effective at introducing correct constraints, but they seldom introduce conceptual errors and are comparable to or surpass other materials for supporting structural repairs. In comparison with 2D role tables, tables carrying state quantities less often introduce correct constraints, but they also less often introduce conceptual errors and sometimes repair prior conceptual errors. For complex written materials, these findings corroborate qualitative observations of material designs made in Section 5.3 and uncover a surprising contribution from narrative episodes, for both algebra students and teachers.

5.5 Discussion

This chapter uses materials found within episodes to reconsider problem solving as the construction of material designs for quantitative inference. By adding material categories for notational structure and quantitative ontology to the episodic framework of preceding chapters, we obtain an interesting methodological synergy: (a) the episodic framework provides an interpretive vocabulary for detecting constructive inferences and their combination during solution attempts, and (b) the materials contained within these episodes provide a novel descriptive account of the problem–solving setting. Qualitative and quantitative analyses examine the relation between material designs and the interpretive categories of Chapters 3 and 4: problem–solving tactics, strategies, outcomes, and the origin of inferences. Each material design provides a partial view of problem structure, and individual designs must usually be combined to construct an effective mathematical representation. An ecological analysis of interactions between the inferential and manipulative capacity of different material designs helps to explain what people do and know when demonstrating mathematical competence. This analysis identifies different constructive packages for algebra story problem solving that have implications for theories of competent quantitative inference and instruction.

5.5.1 Constructing material designs for quantitative inference

From an ecological perspective, a mathematical representation is not simply a knowledge structure applied to a problem. Instead, problem solvers construct representations within the work setting, assembling conventional notational structures and particular kinds of quantities into what I have called a material design. These are local changes to the material setting, they are combined across the temporal and material boundaries of episodes within solution attempts, and these combinations make up a representational ecology for problem-solving. Different patterns of material construction are the primary characteristic that distinguishes more from less competent problem solving.

Material designs. In episodes drawn from a common pool of comparison problems, algebra students and teachers use the same kinds of notations and quantities. Notations include relatively flat expressions for arithmetic and algebra, temporally structured narratives, and relatively complex written structures for two dimensional tables or diagrammatic scenes. Both groups also use two different quantitative ontologies: a prospective view of quantities within states in the problem's event structure and a retrospective view of quantities as roles in the problem's quantitative structure.

As in Chapter 4, episodes from both groups fall outside the standard material designs of algebra instruction (i.e., state or role expressions and 2D role tables). 54.9% of algebra students' episodes construct nonstandard material designs (i.e., narratives, scenes, or 2D state tables), compared with 49.9% of teachers' episodes. Although algebra teachers never use 2D state tables, material excursions from standard algebraic designs suggest that algebra students and teachers resemble each other more than either resembles a view of competence as recalling characteristic equations or translating from words to equations.

Construction versus recall of material designs. Since algebra story problems provide only a text, people must either recall or construct the materials they use to solve these problem. Algebraic formulas (e.g., $D = R \times T$) and the arrangement of 2D role tables (e.g., "dirt" tables) are sometimes directly recalled, but problem solvers more often must construct the material design of local episodes, both to identify quantitative constraints and to determine precise solutions. This claim for local construction is consistent with the finding that material designs follow the content domain of problems, scenes being more frequent in motion problems and narratives or expressions more frequent in work problems. Material designs that are not part of standard instruction appear to be a constructive response to being unable to recall standard mathematical forms (e.g., Karen's "God, I don't remember formulas or anything for this" at S-2 in Section 4.3.2 of Chapter 4).

Material design of local tactics. Analyses in this chapter show that problem–solving tactics are constructed out of different materials. Standard material designs (i.e., state or role expressions and 2D role tables) are used only for algebra or ratio tactics, while nonstandard material designs (i.e., narratives, scenes, or 2D state tables) are most often used for model tactics. With the exception of 2D role tables, complex notational structures are the material basis for tactical excursions outside of schooled algebra. The material design of ratio tactics, using expressions or narrative to carry state quantities, suggests that they may be intermediate between nonstandard and standard approaches to algebra.

People construct model tactics using notational conventions available outside of schooled mathematics (e.g., tables or diagrams) to carry quantities that are given or implied in the problem text. A qualitative analysis of narrative, gestural, and written work during these episodes suggests (a) that the text itself acts as a specification for constructing the model, (b) that the conventional structure and activities of familiar notations (narrative, 2D, or scene) provide a material setting in which to carry quantities, and (c) that the resulting material design can organize inferences and calculation over these quantities. Whether the design carries state or role quantities, both students and teachers can use models to identify and evaluate quantitative inferences. When the material design of a model carries state quantities, prospectively embedded in the structure of events described in the story and distributed over the dimensional structure of the notation, people sometimes undertake "model–based simulation" (e.g., constructing states as connected components in a labelled scene). Although algebra teachers seldom do so, algebra students and advanced undergraduates (Chapter 3) use these material designs to work around the problem of obtaining algebraic precision.

5.5.2 Affordances for inference and manipulation of quantities

Under an ecological view of problem solving, each type of material design affords different opportunities for inference and precise quantitative manipulation. As a result, there are tradeoffs in what any individual design can accomplish, and this helps to explain both the episodic structure of solution attempts and differences in levels of problem–solving competence.

Material designs and quantitative inference. Qualitative analyses of complex notational structures (Section 5.3) show that both 2D and scene notations afford quantitative inferences, although configural properties should give scenes an advantage over the positional array of 2D tables. This contrast is confirmed by quantitative analyses of local difficulties and the origin of structural inferences (Section 5.4.4). These analysis also provide an unexpected finding that narrative materials perform better than either complex written notation when introducing correct quantitative inferences and suppressing conceptual errors. In sharp contrast, standard role expressions most often lead to local difficulties or conceptual errors.

If the material setting influences what one can do or can understand about problem structure, these are the kinds of differences that we should expect. Scene and narrative materials are relatively free from local difficulties, and this can in part be explained by their configural properties, as written on paper or described during a narrative "tour" of problem structure. Drawing a scene showing related distances as labelled segments makes an explicit spatial commitment to the relation between train distances and their composite distance apart, as the scene is constructed and regardless of whether the drawer knows to add these distances at the outset of the episode (e.g., Karen's diagram at the top of Figure 4.3 in Chapter 4). Likewise, narrating an observer's view of train travel with synchronized watches requires that travellers start and stop recording times together, as the story is told and regardless of whether its teller knows that the times are equal at the beginning of the narrative (e.g., Richard's justification for equal times at S–8 in section 4.3.4 of Chapter 4).

In each case, the configural arrangement of quantities simultaneously corresponds to relevant aspects of situational and quantitative structure. The structure of the scene or narrative scenario has an *iconic correspondence* (Smith, 1987)[8] with relations between event boundaries in the story presenting the algebra problem. At the same time, the scene also has an iconic correspondence with quantities in a common interpretation of arithmetic relations between quantities (i.e., points and intervals on a number line). When the correspondence relations for a material design are simultaneously iconic with represented worlds of events and quantities, the design is a "model" of both represented worlds. As constructed on paper, this simultaneity is a powerful constraint that may explain advantages in quantitative inference observed for narrative and scene notations.

By a similar analysis, state or role expressions show the least affordance for quantitative inference because their relatively compressed structure places a homogeneous pool of symbolic entities (constants, variables, and operators) in correspondence with heterogeneous elements of the problem's situational structure. Using Smith's terms, symbols and operators in expressions have a *reified correspondence* to entities and relations in the situation. Despite what may be an iconic correspondence for individual quantities presented by an algebra story problem (or at least their surface syntax), the standard material designs of algebra do not afford the simultaneous iconic correspondence with events and quantities that scene or narrative materials enjoy. Continuing the analysis, 2D notations should afford quantitative inferences less well than scenes, since they stand in *typological correspondence* with events and quantities. However, 2D notations should afford more than standard expressions, since relations between quantities are at least constrained by their type (either from events or

[8]Following Palmer (1978), Smith catalogues a set of correspondence relations that a representing world (source) may take to a represented world (target). In an *iconic correspondence*, each object, property, and relation in the source corresponds to some object, property, and relation in the target.

measured dimensions). Thus, affordances for quantitative inference can be partly explained in terms of the conventional structure of material designs.

Material designs and manipulative capacity. Analyses also show that material designs have very different affordances for manipulating quantities. Without a conventional structure to support precise calculation, narratives, 2D role tables, and role scenes are never used as solution attempts. In contrast, expressions carrying state or role quantities and complex written notations carrying state quantities all afford conventional manipulative activities for precise calculation. As a result, these material designs frequently appear in solution attempts — e.g., 64.3% of algebra students' 2D state tables are used in solution attempts, while 54.5% of teachers' role expressions are used as solution attempts.

No single material design affords the inferences and manipulative capacity required for constructing and using a mathematical representation. Local mismatch between affordance for inference and manipulation produces much of the complex episodic structure of algebra story problem solving. Transitions out of episodes "off track" or with conceptual errors are most common with standard material designs, but transitions also occur when problem solvers are unable to manipulate correct quantitative constraints. Algebra teachers usually carry structural inferences into standard role expressions, though they sometimes attempt solutions outside of standard algebraic materials (e.g., a state scene during heuristic simulation). Algebra students most often work around the need for algebraic precision by using nonstandard alternatives to algebra in their solution attempts (54.6% of their "final episodes" use 2D or scene notations), despite making frequent attempts to construct algebraic expressions.

Thus a material chasm opens between algebra students' solution attempts and a curricular sense of competence usually demonstrated by teachers. The chasm is not between "words and equations" so much as between nonstandard and standard material designs, both constructed within the local setting of a solution attempt. The chasm is also not a matter of missing notations or quantities: with the exception of 2D state tables, algebra students and teachers regularly construct the same materials. Instead, what distinguishes competent algebra story problem solving from students' workarounds is different ways of combining material designs that afford quantitative inference with those that afford precise algebraic manipulation.

Materials and knowing about mathematics. Material designs appearing in solution attempts are complex generative performances in which problem solvers with different backgrounds pursue quantitative inferences. When conventional materials are assembled into local designs, the setting for further inferences about quantitative structure changes, and different manipulative possibilities are introduced. What one knows about the situational and quantitative structure of a problem depends upon how one designs the setting to hold quantities. In ways that are important for problem–solving competence, knowledge about mathematical structure is locally constructed.

Material designs contribute differently to local difficulties encountered within problem–solving episodes. Both algebra students and teachers are least likely to encounter local difficulties when using narratives or scenes. In contrast, both are more likely to end episodes "off track" or have conceptual errors when using expressions, and particularly when these expressions carry role quantities. Students alone have difficulties with 2D notations, which are as likely as not to end off track. Students also introduce errors of *omission* with 2D role tables (i.e., leaving out necessary constraints) and introduce errors of *commission* with 2D state tables. This difference in the type of local errors corroborates qualitative observations that (a) the notational structure of 2D tables provides weak constraints on quantitative inference (e.g., primarily quantitative types), and that (b) state simulation requires explicit relations between local quantities and limiting constraints.

Material designs also contribute differently to the origin of correct constraints, conceptual errors, and repair of prior conceptual errors. As in Chapters 3 and 4, the standard materials of algebra (i.e., role expressions) frequently introduce conceptual errors but have strong informal competitors for introducing correct inferences and repairing prior errors. Narrative and scene materials surpass expressions for introducing correct constraints, they introduce fewer errors than expressions, and they compete favorably for repairing prior conceptual errors. 2D state tables, used exclusively by students, are comparable to role expressions in their contributions to quantitative inference (both positive and negative). 2D role tables, a standard material design used primarily by students, introduce conceptual errors as often as correct constraints, and they never repair prior conceptual errors.

5.5.3 Competent algebra story problem solving

If materials produced during problem–solving episodes are tools for identifying and pursuing quantitative relations, people across levels of competence appear to be able to use the individual tools of algebra story problem solving. Unfortunately, the conventional skills for using a particular tool are not enough when one must construct a package of these tools that is customized to a particular situation (i.e., a problem). Again, materials that are most supportive of structural inference (i.e., narratives) are least supportive of precise calculation.

Packages for quantitative inference. Materials assembled during solution attempts form three distinct "packages" for quantitative inference. The first is a standard instructional package, apparently designed to meet the constructive and manipulative requirements of algebra story problem solving. 2D role tables, as taught, are to help students organize inferences about quantitative relations, which are then to be read directly out of the structure of the elaborated table. It is ironic that "the standard method that I think most of us teach around here" (Richard at S–4 on problem MOD, Section 4.3.4) is seldom used by teachers in the verbal interview protocols and is never observed among advanced undergraduates[9] in Chapter 3.

The second package is used by teachers and advanced undergraduates, both well past formal instruction. Neither group makes significant use of 2D role tables, though both use complex notational structures to infer quantitative relations and to manipulate state quantities. Narrative and scene materials carry quantities in a configural structure that affords inferences about relations between quantities (e.g., equality and additive structures), and these provide a setting in which standard materials with more robust manipulative capacity can be constructed. Thus, nonstandard packaging by competent problem solvers does not abandon algebraic materials, but competent problem solvers construct expressions around nonstandard material designs.

The third package comes from algebra students, who use nonstandard material designs to work around the curricular demand for algebraic precision. Students construct state simulations, using 2D tables or scene diagrams as conventional and relatively error-resistant materials to find precise values. Unlike teachers, students' solution attempts often occur within the same problem-solving episode where they identify relevant structural constraints (e.g., simulation using a state scene or 2D state table). Students are well aware that this is a nonstandard (i.e., "weird") way of solving algebra story problems. What they do not explicitly recognize is the correspondence between what they can construct (i.e., nonstandard material designs) and what they have been taught to manipulate (i.e., standard algebraic materials).

Repackaging algebra story problem solving. What is required to reach the curricular sense of competence is movement across this material chasm: (a) from complex notational structures that support local inferences to compressed and relatively opaque notations that offer a rigorous manipulative calculus, and (b) from a prospective view of quantity embedded in the situation depicted by the problem to a retrospective summary of quantitative relations. Considering the configural opacity of 2D role tables and role expressions for the content domains and dimensional structure of algebra story problems about related functions, this standard instructional package may be a poor choice of materials for teaching the construction of algebraic representations.

Since algebra students spontaneously construct conventional but nonstandard designs, an alternative approach would be to encourage them to assemble standard algebraic materials around the notational structure provided by these designs. The tutorial intervention described in Section 5.3.3 demonstrates a orthogonal configural arrangement or role scenes as a notational structure for constructing standard algebraic expressions. The expectation was that familiar conventions for individual role scenes could be combined in a two dimensional notation, and that the configural relations between scene components could be used as niches for constructing local algebraic relations between quantitative roles. By building on conventional structures that are

[9]Written protocols were not analyzed using categories for notation and quantity, but none contain written material that resembles a 2D role table.

Material designs contribute differently to local difficulties encountered within problem–solving episodes. Both algebra students and teachers are least likely to encounter local difficulties when using narratives or scenes. In contrast, both are more likely to end episodes "off track" or have conceptual errors when using expressions, and particularly when these expressions carry role quantities. Students alone have difficulties with 2D notations, which are as likely as not to end off track. Students also introduce errors of *omission* with 2D role tables (i.e., leaving out necessary constraints) and introduce errors of *commission* with 2D state tables. This difference in the type of local errors corroborates qualitative observations that (a) the notational structure of 2D tables provides weak constraints on quantitative inference (e.g., primarily quantitative types), and that (b) state simulation requires explicit relations between local quantities and limiting constraints.

Material designs also contribute differently to the origin of correct constraints, conceptual errors, and repair of prior conceptual errors. As in Chapters 3 and 4, the standard materials of algebra (i.e., role expressions) frequently introduce conceptual errors but have strong informal competitors for introducing correct inferences and repairing prior errors. Narrative and scene materials surpass expressions for introducing correct constraints, they introduce fewer errors than expressions, and they compete favorably for repairing prior conceptual errors. 2D state tables, used exclusively by students, are comparable to role expressions in their contributions to quantitative inference (both positive and negative). 2D role tables, a standard material design used primarily by students, introduce conceptual errors as often as correct constraints, and they never repair prior conceptual errors.

5.5.3 Competent algebra story problem solving

If materials produced during problem–solving episodes are tools for identifying and pursuing quantitative relations, people across levels of competence appear to be able to use the individual tools of algebra story problem solving. Unfortunately, the conventional skills for using a particular tool are not enough when one must construct a package of these tools that is customized to a particular situation (i.e., a problem). Again, materials that are most supportive of structural inference (i.e., narratives) are least supportive of precise calculation.

Packages for quantitative inference. Materials assembled during solution attempts form three distinct "packages" for quantitative inference. The first is a standard instructional package, apparently designed to meet the constructive and manipulative requirements of algebra story problem solving. 2D role tables, as taught, are to help students organize inferences about quantitative relations, which are then to be read directly out of the structure of the elaborated table. It is ironic that "the standard method that I think most of us teach around here" (Richard at S–4 on

Chapter 6
Conclusions

6.1 Contrasting views of applied mathematical problem solving

This dissertation provides a collection of analytic and empirical studies of "applied quantitative inference" observed when people with different mathematical backgrounds attempt to solve algebra story problems. The inferences required to deliver precise solutions to these problems are "applied" only in the conventional educational sense that students are expected to apply formal mathematical concepts to problems given in school or on standardized assessments outside of school. These are not occasions for "everyday" quantitative reasoning as they might be encountered in life after school. Considered as problem–solving tasks from a traditional cognitive science perspective, algebra story problems are "ill–structured" (Simon, 1973) in that there is no clearly defined set of operations that take one from an initial state, the problem as presented, to a final state, a precise value for an unknown quantity. The central questions are how people manage to structure this gap between words and precise solutions and what role algebraic equations play in that work.

Cognitive and ecological analyses of competent quantitative inference. It is my hope that this dissertation has contributed new materials for competing theoretical accounts of mathematical reasoning, taking algebra story problem solving as a representative case of complex human problem solving. As reviewed in Chapter 2, quite different accounts come from traditional cognitive studies of applied mathematical problem solving, on the one hand, and a diverse collection of "ecological" analyses of quantitative inference in everyday settings on the other. More traditional cognitive analyses generally take algebraic equations as a primary normative ontology for problems, the knowledge required for competence, the activities of problem solving, and the nature of solutions. In contrast, ecological analyses locate problems and their solutions in the ongoing activities of people involved in material and social settings. These approaches to quantitative inference diverge over the the relation between "problems" and their "solutions," the distribution and use of "knowledge" in constructing both, and ascriptions of "competence" on the basis of problem–solving activity (Section 2.5).

Descriptively adequate interpretations of problem solving. Studies presented in the preceding chapters work towards a descriptively adequate theoretical framework for the problem–solving activities supporting quantitative inference. Although cognitive and ecological approaches to these kinds of problems have underscored the difficulty of school math problems and produced interesting observations of quantitative reasoning in practice, neither provides a detailed descriptive account of the origin of correct or incorrect quantitative inference, whether recalled directly from memory or constructed out of the local material and social setting.

For example, Reed et al., (1985) reported that approximately half of a group of college algebra students who found a correct solution on a representative algebra story problem used a "generate–and–test strategy" without writing algebraic expressions. Instead of taking these activities as legitimate forms of inference, follow on studies manipulated materials and instructions to facilitate analogical comparisons, with the result that students apparently abandoned successful strategies in favor of syntactic (and error–prone) equation matching. In contrast, Carraher and Schliemann (1987) found an almost identical form of "rated addition" among construction foremen, who were able to solve unfamiliar target problems involving linear proportions. Their analysis did not systematically examine how similar strategies could be used to manage more complex quantitative relations (e.g., related linear proportions) or how these activities related to other forms of quantitative inference. In this dissertation, detailed explorations of episodic structure in solutions to algebra story problems are used to develop a descriptive account of quantitative inference that relates "nonstandard" problem–solving activities to more "standard" forms of mathematical representation.

6.2 Major contributions

The major analytic and empirical contributions of the dissertation start with a prescriptive analytic framework for describing problem structure, aspects of which are used to describe how inferences about quantitative relations can be constructed and how precise calculations might be organized during problem solving. These lead to a relatively open–ended interpretive framework for analyzing the detailed structure of individual solution attempts, which is applied to written and verbal problem–solving protocols collected with advanced undergraduates, algebra students, and algebra teachers. Finally, relations between problem–solving tactics and various outcomes are reinterpreted within an ecological analysis of local episodes in which different "material designs" support quantitative inference.

Analysis of problem structure. A variety of categorization schemes have been proposed for algebra story problems, but these generally fail to consider interactions between the quantitative and situational structure of typical problems. This dissertation describes a prescriptive analytic framework that combines a *quantitative network formalism* (Section 2.4.1) devised by Greeno and colleagues (Greeno *et al.*, 1986; Shalin and Bee, 1985) with an analysis of corresponding *situational structure* based on entities and relations within and across measured dimensions (Section 2.4.2). A prescriptive analysis of generating mathematical representations shows that these frameworks interact in interesting and useful ways.

- Quantitative relations and larger structures can be generated or evaluated by constructing their corresponding situational structures in a dimensional model of the problem.
- A dimensional model describing the situational context presented by a problem can be used in a form of simulation to find precise solutions without using formal algebraic representation.
- The dimensional structure of such a model can be used to assemble and partially manipulate a standard collection of algebraic expressions.
- Hypothetical interactions between ontological categories for quantity and situation are used to organize exploratory studies of problem solving in the remainder of the dissertation.

An interpretive framework for episodic structure. Obtaining an adequate description of "solutions" in applied mathematical problem solving is a difficult theoretical and practical problem. On the one hand, adopting a prescriptive view of problem–solving activities may leave out the very materials or tactics that support competent performance (e.g., treating algebraic expressions as an exclusive outcome). On the other hand, these materials and activities may be quite diverse, burdening an analysis that looks closely at their interactions. This dissertation draws on traditional methods of protocol analysis to propose the *problem–solving episode* as a fundamental unit of analysis (Section 3.3). Episodes are identified as coherent strategic, tactical, and conceptual activities within an enclosing solution attempt. As a coherent segment of activity, the episode allows relatively open–ended investigation of relations between ascribed strategies, observed tactical materials, and problem–solving outcomes.

The episodic structure of algebra story problem solving. The core of this dissertation reports two exploratory studies of applied quantitative inference. Problem solvers with quite different backgrounds were asked to show their work and/or talk aloud when solving a common set of algebra story problems involving related output–per–time rates. Their solution attempts were interpreted according to the framework for episodic structure developed in Section 3.3, and resulting protocol data were subject to quantitative analysis (Sections 3.4 and 4.4).

- In both written and verbal protocol studies, problem solving episodes without explicit algebraic notations were quite common, including various annotations, construction and use of dimensional models of problem structure, and nonalgebraic uses of ratios.
- Model–based reasoning was a surprisingly common activity across levels of competence: from newcomers (algebra students) to career educators (algebra

teachers) to institutionally–certified competent problem solvers (advanced undergraduates in computer science and engineering).

- Model and ratio tactics also appeared in "final" problem–solving episodes, where people either offered a solution or failed to solve the problem.
- The primary source of difficulty across groups was conceptual errors, in which the problem solver either omitted a relevant structural constraint or introduced an incorrect constraint. Manipulative errors in arithmetic and algebra were less common and often were repaired within the episode.
- Model tactics played a central role in applied quantitative inference: generating quantitative inferences, determining precise values for unknown quantities, and evaluating conjectured constraints.
- Analyses tracking the origin of structural inferences showed that model tactics introduced more correct inferences and fewer conceptual errors than algebraic tactics. Model tactics competed favorably with other tactics for repairing (i.e., retracting or replacing) conceptual errors.
- In the study of advanced undergraduates, manipulating problem structure and the sequence of presenting problems did not produce hypothesized positive or negative transfer effects. Closer examination of written annotations showed several cases of negative transfer that were unrelated to these manipulations.
- In contrast, when problem solvers had more control over their work in a verbal interview setting, spontaneous problem comparisons were common and largely successful, even when the retrieved solution provided a poor or misleading analogical source.

Qualitative analyses of solution attempts by algebra students and teachers in a verbal interview setting (Section 4.3) largely corroborated these findings and extended the interpretive framework for episodic structure in several ways.

- The details provided by verbal interviews showed episodic structure to be non–linear and dense by comparison with analyses of episodes in written protocols.
- Model–based simulation, as written and narrated, simultaneously reproduced relevant aspects of quantitative and situational structure. In general, the temporal and spatial structure of notations constructed during episodes appeared to influence subsequent activities and outcomes.
- Particularly for students, solution attempts proceeded as a form of articulation, working around the tension between being *certain* (understanding the problem) and being *precise* (finding a particular value) under various external constraints (e.g., timeliness, prescribed activity, etc.).
- Participants' descriptions of nonalgebraic tactics suggested that these were generally private and illegitimate problem–solving activities: the "weird way to do it" by comparison with "some other way to do it" for students.

- Model tactics shifted from solution strategies among algebra students to comprehension or evaluation strategies among teachers, who seldom used model–based reasoning as a solution strategy.

An ecological reconstruction of mathematical representation. Detailed observations of the temporal and material production of problem–solving episodes in Chapters 3 and 4 were used to propose a novel account of constructing mathematical representations as the integration of "material designs" (Section 5.2). Material designs are local to episodes and combine differently structured notations (expressions, narratives, two dimensional tables, or scenes) with different views of quantity (state or role). As such, each provides an incomplete perspective on problem structure, but each also affords the problem solver with different inferential and manipulative capacities.

Local designs can be "standard" school–taught materials like algebraic expressions or formula–specific tables; alternately, they can be "nonstandard" materials like diagrammatic scenes showing a succession of quantitative states. Different local designs are integrated across episode boundaries to construct a mathematical representation that supports more or less effective problem–solving activity. The problem–solving episodes identified in verbal interview protocols were reinterpreted using local material designs to organize questions about constructing problem–solving tactics, manipulating local materials to determine precise values for unknown quantities, and demonstrating competent mathematical problem solving (Sections 5.3 and 5.4).

- Both algebra students and teachers use materials from each subcategory of notational structure and quantitative ontology, and both frequently combine these into nonstandard material designs that differ from traditional school mathematics. Unlike students, however, algebra teachers never construct two dimensional tables carrying state quantities.
- Qualitative analyses of verbal protocols suggest that material designs are usually constructed, rather than recalled, and that they may even be constructive resolutions to being unable to recall mathematical forms.
- "Model tactics" are constructed exclusively out of complex notational structures (i.e., narratives, 2D tables, or scenes), and these usually carry state quantities.
- Qualitative analyses of representative cases of material designs suggest that correct inferences should be afforded better by the configural aspects of scenes than by the typed array of 2D tables. The relative advantages of different notations are also examined as different types of correspondence relations provided by material designs onto event and quantitative structure (Section 5.5).
- A quantitative analysis of local difficulties and the origin of structural inferences confirms the advantage for scenes, but unexpectedly shows that narrative episodes perform better than either complex written notation. Expressions, which provide neither configural nor typing constraints on inference, generally perform below any other notation.

circumstances it could provide a useful instructional model for constructive aspects of problem solving (see also Hall, 1989b). As with any model used in teaching, there are problems of registration: the model may cover some aspects of the target domain well but cover other aspects poorly. This proposal addresses relations and operations possible within a representation of the situational structure of compound algebra story problems, and the correspondence of these aspects to relations and operations possible with a representation of quantitative structure. Combined with a quantitative model like that proposed by Greeno *et al.* (1986), their joint contribution could prove more effective than either used alone.

Figure 6.1 shows paired graphical representations of situational and quantitative structure for the MRT problem. At the top of the figure, a *dimensional frame* displays orthogonal output (in this case, distance) and time dimensions, with entities arranged along those dimensions by their respective situational relations: times are adjacent and distances congruent. At the bottom of the figure, a quantitative network (Shalin and Bee, 1985) shows the common distance found by applying motion rates to component times. Each representational device illustrates important aspects of competence in this problem–solving domain.

In contrast with translation rules or tabular arrangements, the illustrative medium of dimensional frames provides a spatial abstraction for compound rate problems that promotes a physical justification for essential quantitative constraints. Time segments add because they are adjacent within the vertical dimension, while distance segments are equal because they are congruent within the horizontal dimension. As noted in earlier descriptions of quantitative structure (Section 2.4.1), substructures corresponding to these constraints must be constructed before using the quantitative network to find a solution — e.g., the additive triad over time extensives that centers the quantitative network in Figure 6.1. The ability to appropriately select and place these quantitative substructures appears to require a substantial investment in training time (Greeno *et al.*, 1986). It may be that a well–designed illustration around the idea of dimensional frames could effectively support the acquisition and use of a quantitative network illustration.

In contrast with a set of algebraic equations, quantitative networks provide a spatial abstraction for variables and equivalence relations that makes the global structure of what would otherwise be a linear encoding more apparent. Rather than writing a set of equations with repeated variable names or constants, a notation that can obscure the role of quantitative entities and make the applicability of algebraic operations difficult to recognize, the quantitative network directly captures the notion of shared variables or constants and multiple ways of reaching a particular unknown. The network provides a visually inspectable form of algebraic calculus, essentially constraint propagation, that may prove easier for students to learn than more traditional instructional methods (i.e., algebraic operations on linear equations). Thus, the two illustrative media provide interdependent representational stages intermediate between a problem text and a correctly manipulated set of algebraic constraints.

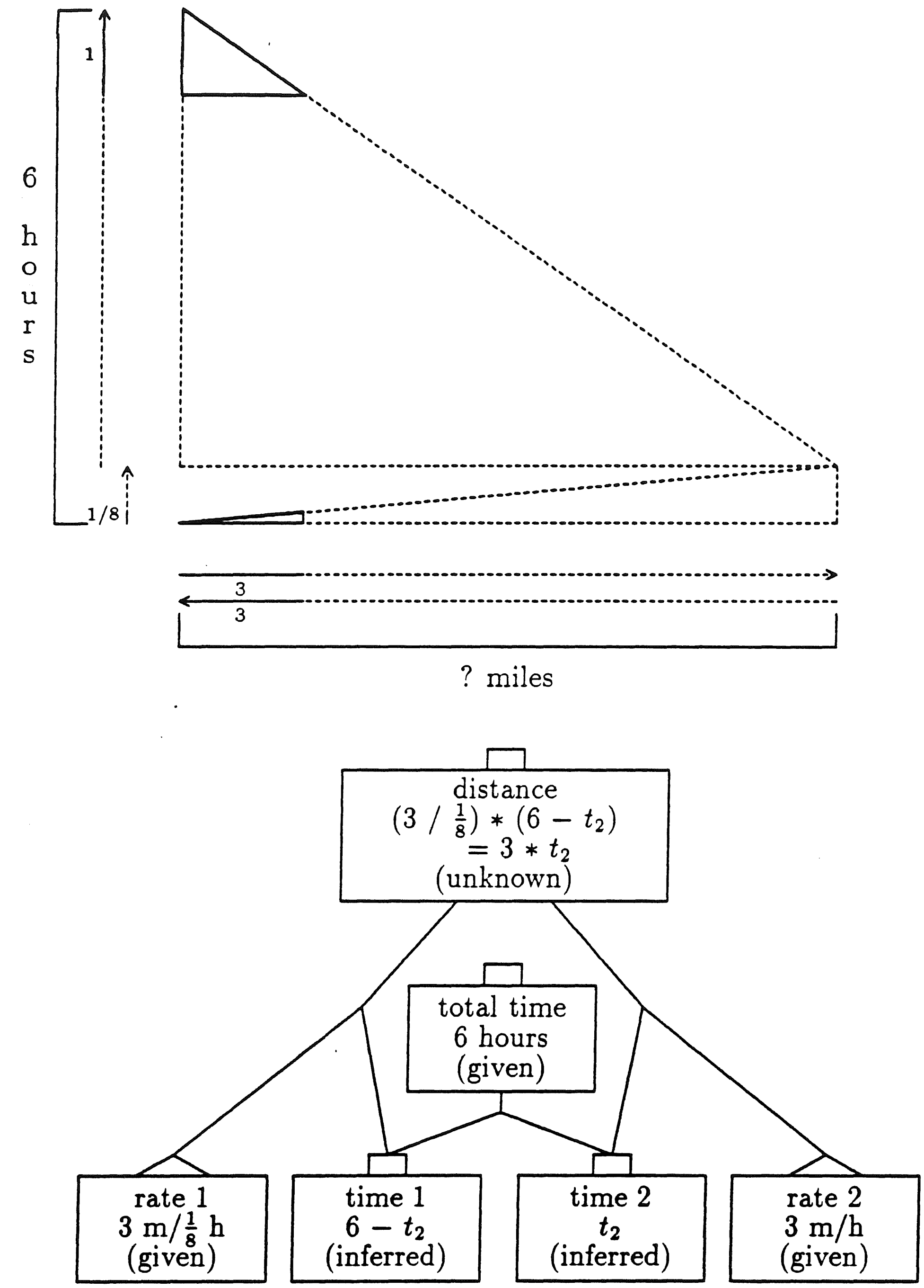


Figure 6.1: Combining interactive illustrations: a two–dimensional frame and a quantitative network for problem MRT.

- Some material designs do not afford any form of manipulation for obtaining precise values (e.g., narratives or 2D role tables). Mismatches between affordance for inference and precise manipulation provide one explanation for the complexity of episodic structure in applied problem solving.
- Differences between "novices" and "experts" cannot be explained on the basis of knowing *about* individual material designs as problem–solving tactics alone. Instead, multiple designs *are* "packaged" together by competent problem solvers in a way that exploits the relative benefits of individual designs (i.e., inference versus manipulation).
- Anecdotal evidence, provided by a tutorial intervention with an algebra student (Section 5.3.3), suggests that more effective representational packages can be constructed around material designs that are already familiar to students.

6.3 Future work

There are a number of directions in which to extend this work. I will mention two in passing and then describe a proposal for designing an educational illustration in some detail. First, the ecological analyses of constructing material designs for quantitative inference could meaningfully be extended to the written protocols described in Chapter 3, and these data could then support a similar form of analysis to that undertaken in Chapter 5. It would be interesting to see if relations between material designs, local outcomes, and the origin of structural inferences hold among this larger group of advanced undergraduates.

Second, the burden of data analysis faced throughout completion of this dissertation has produced a rather large corpus of detailed episodic traces of problem–solving activities. Since these data are organized around a relatively well–structured language for episodic content, they provide an interesting database against which to compare different approaches to constructing and managing qualitative interpretations of students' activities. These range from relatively simple indexing schemes over protocols and analytic annotations in a hypertext format to more elaborate uses of machine learning techniques to induce descriptions of individual or aggregate patterns in solution attempts (e.g., Garlick and VanLehn, 1987).

6.3.1 Material designs as educational illustrations

In the preceding chapters, I have interpreted the relative prevalence and consequence of conceptual versus manipulative errors as evidence that people have difficulty in assembling the quantitative structure of algebra story problems, even long after they have mastered the algebraic formalism. Likewise, the prevalence and functional role of model–based reasoning were interpreted as evidence that even mathematically–sophisticated problem solvers explore the situational context of these problems in an attempt to construct or repair a representation that will support a solution. Based on these findings and their interpretation, I now examine several implications for teaching mathematical problem solving.

The primacy of conceptual errors and use of model–based reasoning, in some cases to recover from these errors, suggest that instruction based solely within the mathematical formalism will be inadequate for solving non–routine or applied problems. Textbook instruction in algebra story problem solving typically addresses this issue by providing some suggestions for "... translating from words to appropriate algebraic forms" (Kolman and Shapiro, 1981, p. 64). These range from direct translation rules taking textual phrases into equations (e.g., rewrite *twice*ju as 2×) to the construction of tables that organize quantitative entities and their interrelationships around known formulas. The desired result is a set of simultaneous linear equations amenable to algebraic operations. While these suggestions provide a sort of organizational strategy for the student's problem–solving activity, they fall well short of specifying how quantitative relations, particularly those that are only implied by the problem text, can be identified, arranged as entries in a table, or effectively used. Instead, the results of these studies point to persistent problem–solving difficulties that the traditional algebra curriculum addresses weakly if at all.

How might these components of competent problem solving be taught more effectively? I will argue that the situational context of an algebra story problem, and in particular the correspondence between situational relations and quantitative constraints, should be a legitimate object of teaching in the algebra curriculum. This is clearly appreciated in other problem–solving curricula. For example, consider the utility of force diagrams for solving statics problems in physics. Students who ignore or incorrectly construct force diagrams can be expected to manipulate equations or formulas without visible signs of progress. This is quite similar to Paige and Simon's (1966) finding that "auxiliary representations" helped students to detect impossible algebra story problems, sometimes before writing any equations at all. The question, then, is whether there might not be a similar organizing representation for algebra story problem solving? There have been some suggestive precedents: Gould and Finzer (1982) described an animated computational environment that allowed students to make guesses about rate problems in a one–dimensional world of motion, and Greeno (1983) described an effective instructional technique in which students use an electric train set to help calculate solutions to compound motion problems. The intent in both cases was to provide students with an interactive illustration[1] as part of their problem–solving instruction.

As one possibility among many, I describe an illustration that draws directly from the analysis of situational structure presented earlier and consider under what

[1]Ohlsson (1987) gives a prescriptive methodology for constructing *interactive illustrations* as well as a particular illustration, called "Rectangle World," for the ratio sense of rational numbers.

circumstances it could provide a useful instructional model for constructive aspects of problem solving (see also Hall, 1989b). As with any model used in teaching, there are problems of registration: the model may cover some aspects of the target domain well but cover other aspects poorly. This proposal addresses relations and operations possible within a representation of the situational structure of compound algebra story problems, and the correspondence of these aspects to relations and operations possible with a representation of quantitative structure. Combined with a quantitative model like that proposed by Greeno *et al.* (1986), their joint contribution could prove more effective than either used alone.

Figure 6.1 shows paired graphical representations of situational and quantitative structure for the MRT problem. At the top of the figure, a *dimensional frame* displays orthogonal output (in this case, distance) and time dimensions, with entities arranged along those dimensions by their respective situational relations: times are adjacent and distances congruent. At the bottom of the figure, a quantitative network (Shalin and Bee, 1985) shows the common distance found by applying motion rates to component times. Each representational device illustrates important aspects of competence in this problem–solving domain.

In contrast with translation rules or tabular arrangements, the illustrative medium of dimensional frames provides a spatial abstraction for compound rate problems that promotes a physical justification for essential quantitative constraints. Time segments add because they are adjacent within the vertical dimension, while distance segments are equal because they are congruent within the horizontal dimension. As noted in earlier descriptions of quantitative structure (Section 2.4.1), substructures corresponding to these constraints must be constructed before using the quantitative network to find a solution — e.g., the additive triad over time extensives that centers the quantitative network in Figure 6.1. The ability to appropriately select and place these quantitative substructures appears to require a substantial investment in training time (Greeno *et al.*, 1986). It may be that a well–designed illustration around the idea of dimensional frames could effectively support the acquisition and use of a quantitative network illustration.

In contrast with a set of algebraic equations, quantitative networks provide a spatial abstraction for variables and equivalence relations that makes the global structure of what would otherwise be a linear encoding more apparent. Rather than writing a set of equations with repeated variable names or constants, a notation that can obscure the role of quantitative entities and make the applicability of algebraic operations difficult to recognize, the quantitative network directly captures the notion of shared variables or constants and multiple ways of reaching a particular unknown. The network provides a visually inspectable form of algebraic calculus, essentially constraint propagation, that may prove easier for students to learn than more traditional instructional methods (i.e., algebraic operations on linear equations). Thus, the two illustrative media provide interdependent representational stages intermediate between a problem text and a correctly manipulated set of algebraic constraints.

transform and then add motion rates in this problem restructures the dimensional frame to have single segments on both time and output dimensions — e.g., $\frac{9}{24}$ hours for each "return trip" mile. The corresponding quantitative network would require only three entities: a time extensive (6 hours, given) results from multiplying the combined rate intensive ($\frac{9}{24}$ hours per mile, inferred) by an unknown extensive for round trip distance. This is a sensible change in representation only because the time segment given in the "goal state" of the problem is presented as a composed whole (i.e., "... he was gone for 6 hours" in the text of problem MRT), and round trip distance segments are congruent. Thus, representational choices in the dimensional frame provide justification for construction of a simplified quantitative network.

Second, problem–solving activity (e.g., iterative simulation) within the dimensional framework can actually help to recover from prior conceptual errors. For example, consider a problem solver who first attempts a solution within the algebraic formalism and omits the constraint that distances are the same (i.e., the same variable). Finding two simultaneous linear equations in three variables, this person reaches an impasse. Choosing model–based reasoning for the purpose of comprehension in the next episode, they immediately face a representational decision in the distance dimension: should positionally distinct or identical spatial segments be chosen? Certainly, the possibility of an incorrect choice remains, but when making this choice in the algebraic formalism of the prior episode, the consequences of an incorrect representational decision were less apparent. Correctly choosing congruent distance segments in the dimensional frame could allow a problem solver to achieve a solution within the model–based reasoning episode, or to return to the algebraic formalism with a more complete representation.

In summary, choosing an apt combination of situational and quantitative models for instructional purposes is a challenging problem. This suggestion for the dimensional frame as an illustrative mechanism would require further refinement in order to be effectively integrated with an algebraic illustration, as discussed above. Nonetheless, this approach is interesting in several respects. First, it is consistent with an empirical picture of episodic problem–solving behavior in people with quite different mathematical backgrounds. Taking these findings as evidence for competent problem solving, such an illustration might support what problem solvers actually do during attempts to solve non–routine problems. The proposal is based on a characterization of these attempts and an analysis of common problem–solving difficulties. Second, although the solution of a particular class of problems may become routine with practice, the ability to construct an algebraic representation will continue to be important for novel problems or problems that have become unfamiliar with the passage of time. Being able to construct a representation in the algebraic formalism, based on the *constraint–generating* inferences I have described as one role for model–based reasoning, may never become entirely routine. Last, combined illustrative media may be of some practical value in delivering instruction on algebra story problem solving, whether provided through computer–based instruction or a traditional algebra curriculum.

Appendix A
Algebra Story Problems

The following problems were constructed as instances within structural classes formed by crossing different time and output relations (see Figure 4.1). Within each structural class, two motion problems and one work problem are included.

Adjacent output, same time

(MOD) Two trains leave the same station at the same time. They travel in opposite directions. One train travels 60 km/h and the other 100 km/h. In how many hours will they be 880 km apart?

(CLOSURE) Tom can drive to Bill's house in 4 hours and Bill can drive to Tom's house in 3 hours. How long will it take them to meet if they both leave their houses at the same time and drive toward each other?

(HOS) A small hose can fill a swimming pool in 6 hours and a large hose can fill it in 2 hours. How long will it take to fill the pool if both hoses are used at the same time?

Same output, sequential time

(MRT) George rode out of town on the bus at an average speed of 24 miles per hour and walked back at an average speed of 3 miles per hour. How far did he go if he was gone for six hours?

(DANCE) Martha left home for the dance studio, riding her bicycle at 8 miles per hour. After reaching the studio, she realized that her dancing shoes were still at home. She called her son, who left home immediately, driving at 24 miles per hour. If their combined travel time was 1 and 1/3 hours, how far was it from home to the studio?

(WC) Randy can fill a box with stamped envelopes in 5 minutes. His boss, Jo, can check a box of stamped envelopes in 2 minutes. Randy works filling boxes. When he is done, Jo starts checking his work. How many boxes were filled and checked if the entire project took 56 minutes?

Anchor–overlap output, same time

(RACE) Frank and Joan both plan to run in the West End race. Joan is faster and can run 10 kilometers per hour, while Frank only runs 8 kilometers per hour. Frank cheats by starting the race 5 kilometers ahead of Joan so that they will cross the finish line together. If both runners start the race at the same time, how long do they run?

(BAGELS) Fred and Ethel begin making bagels at 9 in the morning. Fred can make 24 bagels each half hour, while Ethel only makes 18 bagels. At what time will Fred have made 40 more bagels than Ethel?

(BOARDS) Huck and Tom agree to paint opposite sides of the same fence. Tom can paint 10 boards on his side in an hour, while Huck can paint 14 boards on his side. Tom secretly paints 24 boards on his side the night before. If the boys finish the fence together the next day, how many hours did they work together?

Same output, same–finish time

(BUS–FLY) Karen leaves the county airport at noon, driving 60 miles per hour to a political meeting at the state capitol. Rudolph, realizing that she left her speech behind, catches a plane two and one half hours later. The plane flies at 180 miles per hour and lands at the capital just as Karen arrives. What is the distance from the county airport to the state capitol?

(CROSS) Nancy lives in Angleton and works in Beeville, while Jeff lives in Beeville and works in Angleton. Jeff cycles to work at 15 miles per hour, while Nancy drives to work at 45 miles per hour. For each to be at work on time, Jeff must leave his home 20 minutes before Nancy leaves her home. How far does each travel from home to work?

(CAKES) Tim and Matt pack fruitcakes at the Country Vittles Food Company. Each packs 75 cakes an hour during a normal day. Tim shows up for work on time at 8:00 am Friday and begins packing fruitcakes. Matt, tired from a night out, comes to work three hours late and packs 120 cakes per hour so that he will pack as many cakes as Tim by the end of the day. How many fruitcakes does Matt pack on Friday?

Anchor-overlap output, sequential time

(DELIVERY) Albert and John, of the Fleet Delivery service, can pedal their bikes at 20 and 28 kilometers per hour, respectively. Albert leaves Sandstone with a package for Baytown, but has to stop in Rockport because his chain breaks. He calls John, who leaves immediately from Sandstone and grabs the package as he passes Albert in Rockport. John travels another 20 kilometers to deliver the package. How far is Sandstone from Baytown if their total travel time was 5 hours?

(RUN–BIKE) Ned lives along the route of the charity fun–run, exactly 6 kilometers from the start. Watching television at home, he sees his favorite runner, Hilda, start the race averaging 8 kilometers per hour. Just as Hilda crosses the finish line, Ned begins cycling along the race route in order to congratulate her in the winner's circle. If he cycles at 10 kilometers per hour and reaches the finish line 3 hours after the start of the race, how long was the race?

Appendix A
Algebra Story Problems

The following problems were constructed as instances within structural classes formed by crossing different time and output relations (see Figure 4.1). Within each structural class, two motion problems and one work problem are included.

Adjacent output, same time

(MOD) Two trains leave the same station at the same time. They travel in opposite directions. One train travels 60 km/h and the other 100 km/h. In how many hours will they be 880 km apart?

(CLOSURE) Tom can drive to Bill's house in 4 hours and Bill can drive to Tom's house in 3 hours. How long will it take them to meet if they both leave their houses at the same time and drive toward each other?

(HOS) A small hose can fill a swimming pool in 6 hours and a large hose can fill it in 2 hours. How long will it take to fill the pool if both hoses are used at the same time?

Same output, sequential time

(MRT) George rode out of town on the bus at an average speed of 24 miles per hour and walked back at an average speed of 3 miles per hour. How far did he go if he was gone for six hours?

(DANCE) Martha left home for the dance studio, riding her bicycle at 8 miles per hour. After reaching the studio, she realized that her dancing shoes were still at home. She called her son, who left home immediately, driving at 24 miles per hour. If their combined travel time was 1 and 1/3 hours, how far was it from home to the studio?

(WC) Randy can fill a box with stamped envelopes in 5 minutes. His boss, Jo, can check a box of stamped envelopes in 2 minutes. Randy works filling boxes. When he is done, Jo starts checking his work. How many boxes were filled and checked if the entire project took 56 minutes?

Anchor–overlap output, same time

(RACE) Frank and Joan both plan to run in the West End race. Joan is faster and can run 10 kilometers per hour, while Frank only runs 8 kilometers per hour. Frank cheats

Appendix B
Instructions for Written Protocol Study

The following are two sets of instructions contained in a workbook of algebra story problems used to collect written problem-solving protocols. The first set of instructions occupied the first page of the booklet and were explained to participants as a group. Afterwards, they attempted to solve four problems, each on a separate page of the booklet. Their work on prior problems was not accessible without turning back to earlier pages in the booklet, and they were neither encouraged for forbidden to do this. A set of instructions for explaining their solution attempts appeared on the last page of the booklet, and these were explained to the participants after they attempted to solve all the problems. After the second set of instructions, participants wrote brief explanations for each of their solution attempts.

Instructions

Your work on this exercise will be ANONYMOUS: once turned in, your work cannot be associated with you in any way. Your attendance in class will be checked after this exercise is over.

On the pages which follow, you will see a series of algebra story problems. Please solve each of these problems.

SHOW ALL OF YOUR WORK in the space provided below the problem.

Write only on the problem page.

Work from top to bottom, writing new material below previous material.

When showing your work, please DO NOT ERASE! If you make a mistake, simply mark through the mistake with a single line.

You will be allowed 8 minutes to solve each problem. When you find an answer to a problem, DRAW A BOX AROUND YOUR ANSWER.

Please DO NOT PROCEED to the next problem until instructed to do so. If you finish early, please sit quietly until instructed to continue to the next problem.

Please do not turn page until instructed to do so.

Instructions for Giving Explanations

You can assist us in interpreting your work. For EACH of the four problems you have just seen, please do the following:

1. Explain your work on the BACK OF THE PRECEDING PAGE. DO NOT make any changes to the work you have shown on the problem sheet.

   For each VARIABLE which you used, please tell us what that variable refers to in the problem text. For example:

C_1 is the cost of John's lunch.	C_1

   EXPLAIN the important equations, relations or facts which you used in solving the problem. Some examples are:

a known formula	$e = mc^2$
total cost is the sum of each item's cost	$C_t = C_1 + C_2$
the second event started 10 minutes later	$T_2 = T_1 + 10$

   If you have time, show us what you would tell a friend to help him/her solve the problem.

2. Again, please give us these explanatory comments on the BACK OF THE PRECEDING PAGE for each of the four previous problems. You will be given 20 minutes to explain your work for all the problems.

Please do not turn page until instructed to do so.

References

Allen, J.F. (1983). Maintaining knowledge about temporal intervals. *Communications of the Association for Computing Machinery, 26*(11), 832–843.

Allen, J.F. (1984). Towards a general theory of action and time. *Artificial Intelligence, 23*(2), 123–154.

Anderson, J.R. (1987) Methodologies for studying human knowledge. *Behavioral and Brain Sciences, 10*, 467–505.

Anderson, J.R. (1988) The place of cognitive architectures in a rational analysis. *Proceedings of the 10th Annual Cognitive Science Society Meetings.* Hillsdale, NJ: Lawrence Erlbaum, Associates, 1–10.

Anderson, J.R., Boyle, B.J., and Reiser, B.J. (1985) Intelligent tutoring systems. *Science, 228*, 456–461.

Becker, H.S. (1982) *Art worlds.* Berkeley, CA: University of California Press, Ltd.

Bobrow, D.G. (1968). Natural language input for a computer problem–solving system. In M. Minsky (Ed.) *Semantic information processing* (pp. 135–215). Cambridge, MA: MIT Press.

Brachman, R.J. (1979). On the epistemological status of semantic networks. In N.V. Findler (Ed.) *Associative networks: Representation and use of knowledge by computers* (pp. 3–50). New York: Academic Press.

Brown, C.A., Carpenter, T.P., Kouba, V.L., Lindquist, M.M., Silver, E.A. and Swafford, J.O. (1988) Secondary school results for the fourth NAEP mathematics assessment: algebra, geometry, mathematical methods, and attitudes. *Mathematics Teacher, 81*, 337–347.

Brown, J.S., Collins, A. and Duguid, P. (1988) Situated cognition and the culture of learning. Bolt, Beranek and Newman Laboratories Incorporated, Cambridge, MA, Report No. 6886.

Campione, J.C. and Brown, A.L. (1990) Guided learning and transfer: implications for approaches to assessment. In N. Frederiksen, R. Glaser, A. Lesgold, and M.G. Shafto (Eds.) *Diagnostic monitoring of skill and knowledge acquisition.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 141–172.

Carpenter, T.P., Corbitt, M.K. Kepner, H.S., Lindquist, M.M. and Reys, R.E. (1980) Results of the second NAEP mathematics assessment: secondary school. *Mathematics Teacher, 73*, 329–338.

Carpenter, T.P. and Moser, J.M. (1981) The development of addition and subtraction problem-solving skills. In T.P. Carpenter, J.M. Moser and T. Romberg (Eds.), *Addition and subtraction: developmental perspective.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Carpenter, T.P. and Moser, J.M. (1984) The acquisition of addition and subtraction concepts in grades one through three. *Journal for Research in Mathematics Education, 15*, 179–202.

Carraher, T.N., Carraher, D.W., and Schliemann, A.D. (1985). Mathematics in the streets and in schools. *British Journal of Developmental Psychology, 3*, 21–29.

Carraher, T.N., Carraher, D.W., and Schliemann, A.D. (1987). Written and oral mathematics. *Journal for Research in Mathematics Education, 18*(2), 83–97.

Carraher, T.N., and Schliemann, A.D. (1987). Culture, arithmetic and mathematical models. Manuscript, Universidade Federal de Pernambuco, Recife, Brazil.

Carraher, T.N., Schliemann, A.D. and Carraher, D.W. (1988) Mathematical concepts in everyday life. In G.B. Saxe and M. Gearhart (Eds.) *Children's mathematics.* San Francisco, CA: Jossey–Bass, 71–87.

Charniak, E. (1969) Computer simulation of calculus word problems. In D.E. Walker and C.M. Norton (Eds.) *Proceedings of the International Joint Conference on Artificial Intelligence.* Los Altos, CA: Morgan Kaufman Publishers, Inc., 303–316.

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

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Clement, J. (1982) Algebra word problem solutions: thought processes underlying a common misconception. *Journal for Research in Mathematics Education, 13*(1), 16–30.

Clement, J. (1983). A conceptual model discussed by Galileo and used intuitively by physics students. In D. Gentner and A.L. Stevens (Eds.) *Mental models* (pp. 325–341). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Cohen, P. A (1982) *A calculating people: the spread of numeracy in America.* Chicago, Illinois: The University of Chicago Press.

Cole, M. (1990) Cultural psychology: a one and future discipline? Center for Human Information Processing, Report 131, January.

References

Allen, J.F. (1983). Maintaining knowledge about temporal intervals. *Communications of the Association for Computing Machinery, 26*(11), 832–843.

Allen, J.F. (1984). Towards a general theory of action and time. *Artificial Intelligence, 23*(2), 123–154.

Anderson, J.R. (1987) Methodologies for studying human knowledge. *Behavioral and Brain Sciences, 10*, 467–505.

Anderson, J.R. (1988) The place of cognitive architectures in a rational analysis. *Proceedings of the 10th Annual Cognitive Science Society Meetings.* Hillsdale, NJ: Lawrence Erlbaum, Associates, 1–10.

Anderson, J.R., Boyle, B.J., and Reiser, B.J. (1985) Intelligent tutoring systems. *Science, 228*, 456–461.

Becker, H.S. (1982) *Art worlds.* Berkeley, CA: University of California Press, Ltd.

Bobrow, D.G. (1968). Natural language input for a computer problem–solving system. In M. Minsky (Ed.) *Semantic information processing* (pp. 135–215). Cambridge, MA: MIT Press.

Brachman, R.J. (1979). On the epistemological status of semantic networks. In N.V. Findler (Ed.) *Associative networks: Representation and use of knowledge by computers* (pp. 3–50). New York: Academic Press.

Brown, C.A., Carpenter, T.P., Kouba, V.L., Lindquist, M.M., Silver, E.A. and Swafford, J.O. (1988) Secondary school results for the fourth NAEP mathematics assessment: algebra, geometry, mathematical methods, and attitudes. *Mathematics Teacher, 81*, 337–347.

Brown, J.S., Collins, A. and Duguid, P. (1988) Situated cognition and the culture of learning. Bolt, Beranek and Newman Laboratories Incorporated, Cambridge, MA, Report No. 6886.

Campione, J.C. and Brown, A.L. (1990) Guided learning and transfer: implications for approaches to assessment. In N. Frederiksen, R. Glaser, A. Lesgold, and M.G. Shafto (Eds.) *Diagnostic monitoring of skill and knowledge acquisition.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 141–172.

Gibson, J.J. (1979) *The ecological approach to visual perception.* Boston, MA: Houghton Mifflin.

Gick N.L. and Holyoak K.J. (1980) Analogical problem solving. *Cognitive Psychology, 12*, 306-355.

Gick N.L. and Holyoak K.J. (1983) Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1–38.

Ginsburg, H.P., Kossan, N.E., Schwartz, R., and Swanson, D. (1983) Protocol methods in research on mathematical thinking. In H.P. Ginsburg (Ed.) *The development of mathematical thinking.* New York: Academic Press, 7–47.

Glaser, B.G. and Strauss, A.L. (1967) *The discovery of grounded theory: Strategies for qualitative research.* Hawthorne, New York: Aldine Publishing Company.

Gould, L., and Finzer, W. (1982). A study of TRIP: A computer system for animating time–rate–distance problems. *International Journal of Man–Machine Studies, 17*, 109–126.

Greeno, J.G. (1983). Conceptual entities. In D. Gentner and A.L. Stevens (Eds.) *Mental models* (pp. 227–252). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Greeno, J.G. (1987). Instructional representations based on research about understanding. In A. Schoenfeld (Ed.) *Cognitive science and mathematics education* (pp. 61–88). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Greeno, J.G. (1989) Situations, mental models, and generative knowledge. In D. Klahr and K. Kotovsky (Eds.) *Complex information processing: the impact of Herbert A. Simon.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 285–318.

Greeno, J.G., Brown, J.S., Foss, C., Shalin, V., Bee, N.V., Lewis, M.W., and Vitolo, T.M. (1986). Cognitive principles of problem solving and instruction. Technical Report No. 41, Berkeley Cognitive Science Report Series.

Hall R.P. (1986) Understanding analogical reasoning: viewpoints from psychology and related disciplines. Technical Report 86–10, Department of Information and Computer Science, University of California, Irvine.

Hall, R. (1987). When trains become hoses: Justifying problem comparisons in problem solving. Paper presented at the *Third International Conference on Artificial Intelligence and Education*, University of Pittsburgh and Learning Research and Development Center, Pittsburgh, Pennsylvania.

Hall, R. (1989a) Computational approaches to analogical reasoning: a comparative analysis. *Artificial Intelligence, 39*, 39–120.

Hall R. (1989b) Qualitative diagrams: supporting the construction of algebraic representations in applied problem solving. In D. Bierman, J. Breuker and J. Sandburg (Eds.) *Artificial intelligence and education, Proceedings of the 4th International Conference on AI and Education.* Amsterdam, Netherlands: IOS, pp. 116–122.

Hall, R., Kibler, D., Wenger, E., and Truxaw, C. (1989) Exploring the episodic structure of algebra story problem solving. *Cognition and Instruction, 6*(3), 223–283.

Hayes, J.R. and Simon, H.A. (1977) Psychological differences among problem isomorphs. In N.J. Catellan, D.B. Pisoni and G.R. Potts (Eds.) *Cognitive theory.* Hillsdale, NJ: Lawrence Erlbaum, Associates, 21–41.

Hewitt, C. (1986) Offices are open systems. *ACM Transactions on Office Information Systems*, *4*, 271–287.

Hinsley, D., Hayes, R., and Simon, H. (1977). From words to equations: Meaning and representation in algebra word problems. In P. Carpenter and M. Just (Eds.) *Cognitive processes in comprehension* (pp. 89–106). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Hirstein, J.J., Weinzweig, A.I., Fey, J.T. and Travers, K.J. (1982) Elementary algebra in the United States: 1955–1980. In *Proceedings of Comparative Studies of Mathematics Curricula — Change and Stability 1960–1980.*

Hudson, T. (1983) Correspondences and numerical differences between disjoint sets. *Child Development*, *54*, 84–90.

Hutchins, E. (in press) The technology of team navigation. In J. Balegher, R. Kraut, and C. Egido (Eds.) *Intellectual teamwork: social and technical bases of collaborative work.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson, W.A.L. (1988) *Mental models as mediators of language comprehension and domain knowledge in word problem solving.* Ph.D. Dissertation, University of Pittsburgh, Pittsburgh, PA.

Johnson-Laird, P.N. (1980) Mental models in cognitive science. *Cognitive Science*, *4*, 71–115.

Johnson–Laird, P.N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge: Harvard University Press.

Johnson-Laird, P.N. and Bara, B.G. (1984) Syllogistic inference. *Cognition*, *16*, 1-61.

Kaput, J.J. and Sims-Knight, J.E. (1983) Errors in translations to algebraic equations: roots and implications. *Focus on Learning Problems in Mathematics*, *5*(3,4), 63–78.

Kilpatrick, J. (1967). Analyzing the solution of word problems in mathematics: An exploratory study. (Ph.D. Dissertation, Stanford University.) University Microfilms, 68–6442.

Kilpatrick, J. (1985) A retrospective account of the past 25 years of research on teaching mathematical problem solving. In E. Silver (Ed.) *Teaching and learning mathematical problem solving: multiple research perspectives*, Hillsdale, New Jersey: Lawrence Erlbaum, Publishers, 1–16.

Kintsch, W. (1986). Learning from text. *Cognition and Instruction, 3*(2), 87–108.

Kintsch, W. (1988) The role of knowledge in discourse comprehension: a construction–integration model. *Psychological Review*, *95*(2), 163–182.

Kintsch, W., and Greeno, J.G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*(1), 109–129.

Kitcher, P. (1983) *The nature of mathematical knowledge.* New York: Oxford University Press.

Kolman, B., and Shapiro, A. (1981). *College algebra and trigonometry.* New York: Academic Press.

Kotovsky, K., Hayes, J.R. and Simon H.A. (1985) Why are some problems hard? Evidence from tower of Hanoi. *Cognitive Psychology*, *17*, 248–294.

Kouba, V.L., Brown, C.A., Carpenter, T.P., Lingquist, M.M., Silver, E.A. and Swafford, J.O. (1988) Results of the fourth NAEP assessment of mathematics: number operations, and word problems. *Arithmetic Teacher*, *35*, 14–19.

Lampert, M. (1986) Knowing, doing, and teaching multiplication. *Cognition and Instruction*, *3*(4), 305–342.

Larkin, J.H. and Simon, H.A. (1987) Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*, 65–99.

Latour, B. (1986) Visualization and cognition: thinking with eyes and hands. *Knowledge and Society: Studies in the Sociology of Culture Past and Present*, *6*, 1–40.

Lave, J. (1986). The values of quantification. In J. Law (Ed.) *Power, action and belief* (pp. 88–111). London: Routledge and Kegan.

Lave, J. (1988a). *Cognition in practice: Mind, mathematics and culture in everyday life.* Cambridge: Cambridge University Press.

Lave, J. (1988b) Word problems: a microcosm of theories of learning. Paper presented to the American Educational Research Association Meetings, April.

Lave, J. (1988c) The culture of acquisition and the practice of understanding. Institute for Research on Learning, Palo Alto, CA, Report No. IRL88-0007.

Lave, J., Murtaugh, M. and de la Rocha, O. (1986) The dialectic of arithmetic in grocery shopping. In B. Rogoff and J. Lave (Eds.), *Everyday cognition: Its development in social context.* Cambridge, MA: Harvard University Press, 67–94.

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

Levin, J.A., Reil, M.M., Cohen, M., Goeller, M., Boruta, M., and Miyake, N. (1986) Reflexibility in problem solving: the social context of expertise. *Proceedings of the Eight Conference of the Cognitive Science Society.* Hillsdale, NJ: Lawrence Erlbaum Associates, 346–352.

Livingston, E. (1986) *The ethnomethodological foundations of mathematics.* London: Routledge and Paul.

Lucas, J.F. (1980). An exploratory study on the diagnostic teaching of heuristic problem-solving strategies in calculus. In J.G. Harvey and T.A. Romberg (Eds.) *Problem-solving studies in mathematics* (pp. 67-91). University of Wisconsin — Madison: Wisconsin Research and Development Center for Individualized Schooling, Monograph Series.

Lynch, M. (1988) The externalized retina: selection and mathematization in the visual documentation of objects in the life sciences. *Human Studies*, *11*, 201–234.

Mayer, R.E. (1981). Frequency norms and structural analysis of algebra story problems into families, categories, and templates. *Instructional Science, 10*, 135–175.

Mayer, R.E., Larkin, J., and Kadane, J. (1984). A cognitive analysis of mathematical problem-solving ability. In R. Sternberg (Ed.) *Advances in the psychology of human intelligence*, *2* (pp. 231–273). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

McCloskey, M. (1983). Naive theories of motion. In D. Gentner and A.L. Stevens (Eds.) *Mental models* (pp. 325–341). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

National Commission on Excellence in Education (1983) *A nation at risk: the imperative for educational reform.* Washington, DC: U.S. Government Printing Office.

National Council of Teachers of Mathematics (1989) *Curriculum and evaluation standards for school mathematics.* Reston, VA: The National Council of Teachers of Mathematics, Inc.

Nesher, P, Greeno, J.G. and Riley, M.S. (1982) The development of semantic categories for addition and subtraction. *Educational Studies in Mathematics*, *13*, 373–394.

Nesher, P. and Teubal, E. (1974) Verbal cues as an interfering factor in verbal problem solving. *Educational Studies in Mathematics*, *6*, 41–51.

Newell, A. (1969). Heuristic programming: Ill–structured problems. In J.S. Aronofsky (Ed.), *Progress in Operations Research, 3*, 360–414.

Newell, A. (1982) The knowledge level. *Artificial Intelligence*, *18*, 87–127.

Newell, A. (1989) Putting it all together. In D. Klahr and K. Kotovsky (Eds.) *Complex information processing: the impact of Herb Simon.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 399–440

Newell, A. and Simon, H.A. (1972) *Human problem solving.* Englewood Cliffs, NJ: Prentice–Hall.

Newell, A. and Simon, H.A. (1976) Computer science as empirical inquiry: symbols and search. *Communications of the ACM*, *19*(3), 113–126.

Newman, D., Griffin, P. and Cole, M. (1989) *The construction zone.* Cambridge, UK: Cambridge University Press.

Nisbett, R.E., Fong, G.T., Lehman, D.R., and Cheng, P.W. (1987). Teaching reasoning. *Science, 30*, 625–631.

Norman, D.A. (1988) *The psychology of everyday things.* New York: Basic Books.

Norman, D.A. and Hutchins, E.L. (1988) Computation via direct manipulation. Final Report, Office of Naval Research, Contract N00014–85–C–0133.

Novak, G.S. (1976) *Computer understanding of physics problems stated in natural language.* Technical Report NL–30, Department of Computer Sciences, The University of Texas at Austin, Texas.

Novick, L.R. (1987). Problem similarity, analogical transfer, and individual differences. University of California at Los Angeles, Unpublished manuscript.

Ohlsson, S. (1987). Sense and reference in the design of interactive illustrations for rational numbers. In R. Lawler and M. Yazdani (Eds.), *Artificial intelligence and education.* Norwood, New Jersey: Ablex, 307–344.

Paige, J.M., and Simon, H.A. (1966). Cognitive processes in solving algebra word problems. In B. Kleinmuntz (Ed.) *Problem solving: Research, method, and theory* (pp. 51–119). New York: John Wiley.

Pea, R.D. (1987) Socializing the knowledge transfer problem. *International Journal of Educational Research 11*(6), 639–663.

Piaget, J. (1970) *The child's conception of movement and speed.* New York: Basic Books.

Piaget, J. (1977) *Epistemology and psychology of functions.* Dordrecht, Holland: D. Reidel Publishing Company.

Pirolli, P.L., and Anderson, J.R. (1985). The role of learning from examples in the acquisition of recursive programming skills. *Canadian Journal of Psychology, 39*(2), 240–272.

Polya, G. (1945). *How to solve it.* Princeton, New Jersey: Princeton University Press.

Quintero, A.H. (1981). The role of semantic understanding in solving multiplication word problems: Two step R x E problems. In Schwartz, J.L. *The role of semantic understanding in solving multiplication and division word problems.* Final report to the National Institute of Education, NIE–G–80–0144.

Quintero, A.H., and Schwartz, J.L. (1981). The development of the concept of ratio in children. In Schwartz, J.L. *The role of semantic understanding in solving multiplication and division word problems.* Final report to the National Institute of Education, NIE–G–80–0144.

Palmer, S.E. (1978) Fundamental aspects of cognitive representation. In E. Rosch and B.B. Lloyd (Eds.) *Cognition and categorization.* Hillsdale, New Jersey: Lawrence Erlbaum and Associates, 259–303.

Reed, E. and Jones, R. (1982) *Reasons for realism: selected essays of James J. Gibson.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Reed, S.K. (1984). Estimating answers to algebra word problems. *Journal of Experimental Psychology: Learning Memory, and Cognition, 10*(4)., 778–790.

Reed, S.K. (1987). A structure–mapping model for word problems. *Journal of Experimental Psychology; Learning, Memory, and Cognition, 13*(1), 124–139.

Reed, S.K., Dempster, A., and Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 106–125.

Reed, S.K. and Ettinger, M. (1987) Usefulness of tables for solving word problems. *Cognition and Instruction, 4*(1), 43–59.

Resnick, L.B. (1987a) *Education and learning to think.* Washington, DC: National Academy Press.

Resnick, L.B. (1987b) Learning in school and out. *Educational Researcher, 16*(9), 13–20.

Richards, D.D and Siegler, R.S. (1979) The development of time, speed, and distance concepts. *Developmental Psychology, 15*, 288–298.

Riley, M.S. and Greeno, J.G. (1988) Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction*, *5*(1), 49–101.

Riley, M.S., Greeno, J.G., and Heller, J.I. (1983). Development of children's problem-solving ability in arithmetic. In H.P. Ginsburg (Ed.) *The development of mathematical thinking* (pp. 153–196). New York: Academic Press.

Rorty, R. (1979) *Philosophy and the mirror of nature.* Princeton, NJ: Princeton University Press.

Roschelle, J. and Behrend, S. (1988) Paper presented at the American Educational Research Association annual meetings, San Francisco, CA.

Rumelhart, D.E., McClelland, J.L., and the PDP Research Group (1986) *Parallel distributed processing: explorations in the microstructure of cognition*, Volumes 1 and 2. Cambridge, MA: The MIT Press.

Schoenfeld, A.H. (1985). *Mathematical problem solving.* New York: Academic Press, Inc.

Schwartz, J.L. (1982). The semantic calculator. *Classroom Computer News*, *2*, 22–24.

Seifert, C. and Norman, D. (1987) Levels of research. *Behavioral and Brain Sciences*, *10*, 490–492.

Shalin, V., and Bee, N.V. (1985). Analysis of the semantic structure of a domain of word problems. Technical Report No. APS-20, University of Pittsburgh Learning Research and Development Center.

Scribner, S. (1984) Studying working intelligence. In B. Rogoff and J. Lave (Eds.), *Everyday cognitions: its development in social context.* Cambridge, MA: Harvard University Press, 7–40.

Silver, E.A. (1979). Student perceptions of relatedness among mathematical verbal problems. *Journal for Research in Mathematics Education, 10*, 195–210.

Silver, E.A. (1981). Recall of mathematical problem information: solving related problems. *Journal for research in mathematics education, 12*(1), 54–64.

Simon, H.A. (1973). The structure of ill-structured problems. *Artificial Intelligence, 4*, 181–201.

Simon, H.A. and Hayes, J.R. (1976) The understanding process: problem isomorphs. *Cognitive Psychology, 8*, 165–190.

Singley, M.K. (1986). Developing models of skill acquisition in the context of intelligent tutoring systems. Doctoral dissertation, Department of Psychology, Carnegie-Mellon University.

Smith (1987) The correspondence continuum. Center for the Study of Language and Information, Report No. CSLI-87-71.

Star, S.L. (1983) Simplification in scientific work: An example from neuroscience research. *Social studies of science*, *13*, 206–228.

Star, S.L. (1988) Introduction: the sociology of science and technology. *Social Problems*, *35*(3), 197–205.

Star, S.L. (1989a) *Regions of the mind: brain research and the quest for scientific certainty.* Stanford, CA: Stanford University Press.

Star, S.L. (1989b) The structure of ill–structured solutions: boundary objects and heterogeneous distributed problem solving. In L. Gasser and M.N. Huhns (Eds.) *Distributed artificial intelligence*, Volume II. San Mateo, CA: Morgan Kaufmann Publishers, Inc., pp. 37–54.

Star, S.L. (1989c) Human beings as material for computer science, or what computers science can't do. Paper presented to the American Philosophical Association, Berkeley, California, March 23.

Stigler, J.W., Fuson, K.C., Ham, M. and Kim M.S. (1986) An analysis of addition and subtraction word problems in American and Soviet elementary mathematics textbooks. *Cognition and Instruction*, *3*(3), 153–171.

Strauss, A.L. (1987) *Qualitative analysis for social scientists.* Cambridge: Cambridge University Press.

Suchman, L.A. (1987) *Plans and situated actions: the problem of human–machine communication.* Cambridge, UK: Cambridge University Press.

Suchman, L.A. (1988) Representing practice in cognitive science. *Human Studies*, *11*, 305–325.

Swetz, F. (1987) *Capitalism and arithmetic.* La Salle, Illinois: Open Court Publishing Company.

Thompson, P.W. (1988). Artificial intelligence, advanced technology, and learning and teaching algebra. To appear in C. Kieran and S. Wagner (Eds.), *Research issues in the learning and teaching of algebra.* Reston, VA: National Council of Teachers of Mathematics.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.

Ullman, S. (1980) Against direct perception. *The Behavioral and Brain Sciences*, *3*, 373–415.

van Dijk, T.A. and Kintsch, W. (1983) *Strategies of discourse comprehension.* New York: Academic Press.

Vergnaud, G. (1981) A classification of cognitive tasks and operations of thought involved in addition and subtraction problems. In T.P. Carpenter, J.M. Moser and T. Romberg (Eds.), *Addition and subtraction: developmental perspective.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Vergnaud, G. (1983) Multiplicative structures. In R. Lesh and M. Landau (Eds.) *Acquistion of mathematics concepts and processes.* New York: Academic Press, 127–174.

Weaver, C.A. and Kintsch, W. (1988) The conceptual structure of word algebra problems. Technical Report 88–11, University of Colorado, Institute of Cognitive Science.

Wilkening, F. (1981). Integrating velocity, time, and distance information: A developmental study. *Cognitive Psychology, 13*, 231–247.

Wollman, W. (1983) Determining the sources of error in a translation from sentence to equation. *Journal for Research in Mathematics Education, 14*(3), 169–181.