MATHEMATICAL SENSE, MATHEMATICAL SENSIBILITY:

THE ROLE OF THE SECONDARY GEOMETRY COURSE IN TEACHING

STUDENTS TO BE LIKE MATHEMATICIANS

by

Michael Kevin Weiss

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics and Education)
in The University of Michigan
2009

Doctoral Committee:

Associate Professor Patricio G. Herbst, Chair
Professor Deborah L. Ball
Professor Hyman Bass
Professor Robert E. Megginson

To Fruma, of course.

Acknowledgements

In 2001, I left the University of Michigan Mathematics Department to become a full-time high school mathematics teacher, taking with me an unfinished research problem in ring theory in my backpack and an unfinished doctoral program on my transcript. At the time I thought the journey out was a one-way trip — and so I can honestly repeat the old cliché, "I never thought this day would come." For my journey back into academia, I owe a tremendous debt of gratitude to Prof. Deborah Ball, Prof. Hyman Bass, and Prof. Patricio Herbst, who provided generous funding, waived requirements, and (it seemed to me) moved bureaucratic mountains to make my transition into Mathematics Education a smooth one. I am particularly gratified that all three of them, together with Prof. Bob Megginson (who mentored me in my very first mathematics teaching position, back in 1995), were available and willing to serve on my doctoral committee. I am extremely fortunate to have had such an eminent group of scholars challenging and supporting my work.

Over the course of the last five years I have been privileged to work with a group of remarkable colleagues in the GRIP (Geometry, Reasoning and Instructional Practices) research group. I have learned so much from Gloriana, Wendy, Talli, Chialing, Manu, Takeshi, and Chieu, and all of us have benefited from the mentorship and supervision of Prof. Herbst. It is hard for me to image a better apprenticeship into scholarly work.

In this dissertation I argue that knowledge of the mathematical sensibility is encoded in individuals' narratives of mathematical practice, and my own narratives bears that out. I wish to thank the many teachers I have had, each of whom taught me (through practice) valuable lessons about mathematics: in particular Profs. Tom Storer, Toby Stafford, Karen Smith, Arthur Wasserman, and Carolyn Dean have had lasting influences on me. I also want to thank Rabbi Eric Grossman, my former colleague at the Jewish Academy of Metropolitan Detroit, who helped me to become aware of the prospective function of assessment items.

Part of this dissertation is based on data collected under NSF grants ESI-0353285 and REC-0133619 to Patricio Herbst. All opinions are those of the author and do not necessarily represent the views of the Foundation.

Most of all I wish to thank my family, who have stood behind me and supported me all of these years. To my wife Fruma, and to our children, Isaac, Sarah, Kinneret, Chana and Tova — It is hard to find words to express how much I love you all. None of you have ever complained when I had to do too much work, but none of you ever let me forget that there are more important things in life. This is your milestone as much as it is mine.

Table of Contents

List of Figures

List of Excerpts

Chapter 1

Introduction

Teaching students to "be like" a mathematician

Mathematics instruction is an endeavor with many stakeholders, each of whom may have a different (and valid) claim as to the ultimate purpose of the endeavor. Among these purposes are, for example, to produce citizens who can use mathematics to compete in a global economy that increasingly calls for technological expertise (Committee on Prospering in the Global Economy 2007); to reduce social inequities by cultivating knowledge and skills in traditionally underserved communities (Gutstein 2003, 2006; Moses, 2001); and, not least, to produce a citizenry that is culturally literate, including mathematical literacy (Kline 1953).

In addition to these worthy goals, in recent years another perspective has become part of the discourse of mathematics education: namely, that one of the goals of mathematics education is to enculturate students into a *community of practice* (Lampert 2001; Sfard 1998; Wenger 1997). From this point of view, the goal of mathematics instruction is, at least in part, to cultivate in students some of the practices, values and sensibilities that characterize the work of mathematicians — that is, to teach students what it is like to *be like* a mathematician; or, more succinctly still, to teach a *mathematical sensibility* (Ball & Bass 2003; Bass 2005). This perspective has been put forth by Askew (2008), who argues that "the mathematical behaviours demonstrated by effective teachers are still proxies for something else – a mathematical sensibility – that cannot be reduced to a list of mathematical topics."

This attention to the practices and values that are characteristic of mathematicians is connected to larger questions of the relationship of education to the various academic disciplines to which education is accountable — questions that have influenced and inspired scholars of education since the early twentieth century. Dewey, writing in *The Child and the Curriculum* (1902), likened disciplinary content knowledge to a map:

> The map is not a substitute for personal experience. The map does not take the place of an actual journey. The logically formulated material of a science or branch of learning, of a study, is no substitute for the having of individual experiences… But the map, a summary, an arranged and orderly view of previous experiences, serves as a guide to future experience; it gives direction; it facilitates control; it economizes effort, preventing useless wondering, and pointing out the paths which lead most quickly and most certainly to a desired result. Through the map every new traveler may get for his own journey the benefits of the results of others' explorations without the waste of energy and loss of time involved in their wanderings — wanderings which he himself would be obliged to repeat were it not for just the assistance of the objective and generalized record of their performances. (pp. 26-28)

For Dewey, then, the process of discovery is an essential element of education — but this discovery should be not fully idiosyncratic and self-directed, but rather a guided one that benefits from the accumulated experience of those who have previously made the journey. Taking Dewey's analogy further, one might say that it is desirable for the novice to learn from the more experienced traveller not only the territory under exploration — its landmarks and geographical features — but also the *art of travelling*; that is, how to choose a good road to travel down, how to blaze a trail, how and where to choose a site in which to pitch one's tent, and so forth.

Half a century later Bruner (1960) made a similar point: "An educated man should have a sense of what knowledge is like in some field of inquiry, to know it in its connectedness and with a feeling for *how the knowledge is gained*… I do not mean that each man should be carried to the frontiers of knowledge, but I do mean that it is possible to take him far enough so that he himself can see how far he has come and *by what means*" (p. 618; emphasis added). Bruner thus emphasizes the importance of metaknowledge, that is, an awareness not only of the content of a discipline, but also of the practices by which that content is constructed. Schwab (1961/1974), writing at about the same time, made a similar argument: that instruction may be conceived as "concerned primarily with the imparting of knowledge", or as "concerned primarily with the imparting of arts and skills", but that in either case it is the structures of the individual disciplines themselves that determines what knowledge, what arts, and what skills matter. Thus, "learning mathematics" necessarily requires engagement with the structure, the practices, and the values — that is, the *sensibility* — that characterizes mathematics as a discipline.

To these considerations, Chazan, Callis & Lehman (2008) add an additional argument based on egalitarian ideals. Chazan et al. observe that, to the extent that mathematics instruction departs from the sensibility underlying mathematical practice — a sensibility that values reason, creativity and agency — it replaces that sensibility with another, one rooted ultimately in authority and elitism. Thus a shift in emphasis towards the mathematical sensibility can be appreciated as an embrace of egalitarian values in education (Cusick ,1983).

But despite the compelling nature of these arguments, it remains largely unclear what is really *meant* by the phrase "mathematical sensibility" — its architecture as a theoretical construct — and, perhaps more challenging, what it takes to translate this ambitious goal into actual classroom practice. To illustrate this latter point, consider the following example: One characteristic that seems to define the work of a research mathematician, but that typically does not form part of the experience of mathematics students, is that the professional mathematician chooses for herself what problems she will work on, rather than work always on assignments set by others.[1] Indeed, research mathematicians are held accountable by their colleagues for choosing difficult and worthwhile problems to work on. Could school mathematics help to foster an environment in which students are similarly free to choose, and be held accountable for choosing, their own problems? The desire to shape education in such a way that this becomes part of the educational experience of students is reflected in the document *Principles and Standards of School Mathematics* (NCTM 2000), which ambitiously identifies as "a major goal of school mathematics programs" the creation of "autonomous learners" who "can take control of their learning by defining their goals and monitoring their progress" (p. 20), and asserts that "posing problems, that is, generating new questions in a problem context" is an important skill for teachers to nurture in their

---

[1] It may be objected that this characterization distorts reality by valorizing the academic research mathematician over her counterpart who works in industry, finance, etc. I concede the point; it is certainly true that the latter may have less discretion over her choice of problems on which to work. Nevertheless it is still the case that a typical student in an elementary or secondary mathematics classroom has considerably less autonomy and freedom to choose a research problem than does her counterpart working in industrial and applied mathematics.

students (p. 116). As an illustration of this goal, the *Standards* offer the following vignette:

> Lei wanted to know all the ways to cover the yellow hexagon using pattern blocks. At first she worked with the blocks using fairly undirected trial and error. Gradually she became more methodical and placed the various arrangements in rows. The teacher showed her a pattern-block program on the class computer and how to "glue" the pattern-block designs together on the screen. Lei organized the arrangements by the numbers of blocks used and began predicting which attempts would be transformations of other arrangements even before she completed the hexagons. The next challenge Lei set for herself was to see if she could create a hexagonal figure using only the orange squares.... (p. 116)

The student is presented here as an apprentice mathematical researcher, posing original (to her) problems and investigating them through various means, guided by a more experienced member of the community of practice (i.e., the teacher). And yet this vignette is notable as much for what it leaves out as for what it shows. What preceded Lei's decision to investigate the ways to cover the yellow hexagon using pattern blocks? What was the task that the teacher set for the students, and how did working on that task create a context in which Lei's question could arise? What was the work that the teacher had to do to create a context for investigation that allowed Lei to decide what she "wanted to know"? Is it only coincidence that she happened to "want to know" at precisely the time when the class was supposed to be engaged in mathematical work? What were other students working on while Lei pursued her questions? Did they all "want to know" the same things? Did students "want" to learn the findings of others, and, if not, were they free to ignore them? Were some investigations deemed more valuable than others? According to what criteria? Who makes such an evaluation?

The above discussion is intended to highlight the fact that it is much easier to create a single illustration of an individual student engaged in serious mathematical inquiry than it is to describe the *work of teaching* that makes such inquiry possible in a classroom. And this is only a small part of the challenge. Posing problems to and for oneself is only one of the many things that a mathematician does as part of her practice; some of the other things are represented in Fig. 1.1, below. The network of practices represented in Fig. 1.1 is merely an illustration based on the author's own experiences, and as such should not to be taken too seriously or too literally; nevertheless it suggests some of the complexity that characterizes mathematical practice. It is one thing (already very difficult) to aspire to engage students in the work of conjecturing, investigating, proving, and so forth (the individual nodes of the diagram in Fig. 1.1); it is yet another thing to cultivate the habits of mind that generate new questions from existing knowledge, or a refined conjecture from a partial but incomplete proof, and so on — the arrows in the diagram. These arrows represent the question, "What should I do next?", and this question is, fundamentally, one mathematical vision and values: vision because the practitioner needs to perceive at any given moment what the available possibilities are, and values because the practitioner needs a way of determining which among those possibilities are most worthwhile. To what extent can mathematics instruction be held accountable for the goal of cultivating these habits of mind, this mathematical sensibility, in students?

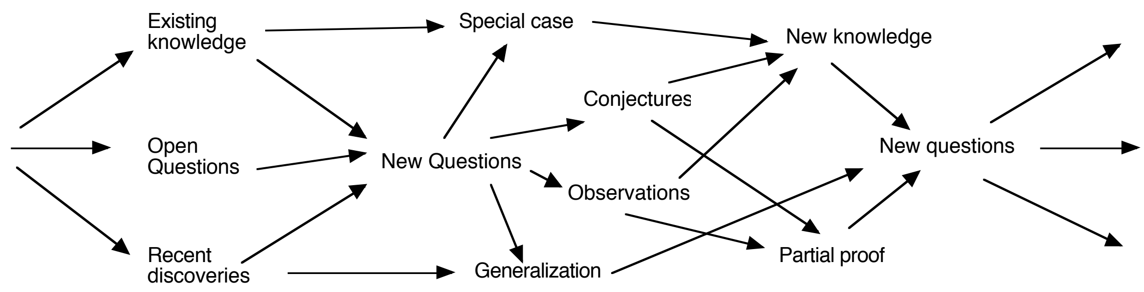


Figure 1.1. Some of the things mathematicians do.

Why Geometry?

For more than a century, a course in Geometry has been part of the standard high school curriculum. During that time, a variety of arguments have been brought forth for why all students should take Geometry; in a survey of those arguments, González & Herbst (2006) identified four modal arguments for the Geometry course. The third of those arguments, which they call the *mathematical argument*, holds that Geometry is uniquely well-suited for providing students with experiences that are close to those of research mathematicians. Indeed, Geometry is the only high school course in which students routinely deal with necessary consequences of abstract properties, and in which students are held accountable for reading, writing, and understanding mathematical proof. To be sure, for years advocates of reform have called for changing that status quo by integrating mathematical proof and reasoning across all of K-12 mathematics (NCTM 1991, 2000), but at the present time Geometry remains students' closest encounter with mathematics as it is practiced by mathematicians. For this reason, it seems reasonable that if any course can serve to cultivate in students a mathematical sensibility, it would be the Geometry course.

And yet despite this, the Geometry course has been criticized for years as a caricature of authentic mathematics, characterized by a proliferation of unnecessary postulates and imprecise definitions, claims that are accepted without proof, and proofs that valorize form over substance (Christofferson 1930; Thurston 1994; Usiskin 1980; Weiss, Herbst & Chen 2009). In light of the fact that the Geometry course has historically struggled to live up to its perceived potential as an opportunity to enculturate students into authentic mathematical habits of mind, I posit that there must be something — some combination of institutional or sociocultural or psychological features — that makes this goal particularly challenging. Taking these considerations in mind, I set out in this dissertation to answer the following research question: *What is the capacity for the high school Geometry course to serve as an opportunity for students to learn the elements of a "mathematical sensibility"?* More specifically, I ask:

1. What, precisely, are the elements of the "mathematical sensibility"?

2. Do Geometry teachers customarily hold their students accountable for learning those elements?

3. If students were to be held accountable for such learning, how could that accountability be represented in the form of assessments?

4. What are the factors that make it difficult to teach a course with the intention of holding students accountable for learning the elements of the mathematical sensibility?

Towards answering these four questions, the dissertation is structured in three main parts. In the first part, I unpack the construct to explain just what could be meant by "mathematical sensibility". In the second part, I document how experienced teachers of

Geometry respond to episodes in which one or more facets of the mathematical sensibility are on display. Finally, in the third part I present a kind of "existence proof" — an analysis of a corpus of assessment items from classroom in which a teacher set out to teach the mathematical sensibility, and to hold students accountable for learning its component parts.

The three parts of this dissertation are, in one sense, separate studies: each draws on its own set of records, theoretical frameworks, and methodologies for analysis. These several frameworks and methodologies will be described in the appropriate chapters. But from another point of view, the three parts of the dissertation are interconnected, both structurally and thematically. The proceeds of the first study, the conceptual analysis of "mathematical sensibility", yields an enumeration of 16 distinct mathematical dispositions which, collectively, comprise my construct. These 16 dispositions are then used in both subsequent studies to code and aggregate the data. Moreover the results of the second study, the analysis of teachers' responses to proposed episodes in which one or more mathematical sensibility is in play, define a kind of baseline or context against which the data for the third study becomes particularly meaningful.

Overview of the dissertation

I turn now to a brief description of the data sources, methods, and theoretical frameworks for the three parts of the dissertation in slightly more detail. These will be further elaborated on in the relevant chapters.

In the first part of the dissertation (Chapter 2), I undertake a conceptual analysis of the notion of a mathematical sensibility. I begin from the position that any practice (such as playing poker, writing computer programs, or doing mathematics) is characterized by a *practical rationality* (Bourdieu 1998; Herbst & Chazan 2003a). This rationality informs the practitioner's judgment about what is appropriate to do at any given moment; it is the "feel for the game" that practitioners draw upon in making timely decisions. Bourdieu describes practical rationality as a network of interacting "dispositions", a technical term referring to shared categories of perception and appreciation[2]. Dispositions are the metaphoric lenses through which a practitioner looks out on practice: they give structure to what the practitioner sees, and what he or she values.

In this dissertation, the phrase "mathematical sensibility" will be taken to mean the practical rationality of research mathematicians. The phrase "elements of a mathematical sensibility" will refer to the mathematicians' dispositions, in the sense above. These dispositions, it must be stressed, are not characteristics of any individual practitioners, but rather belong to the practice itself. That is, the extent to which two individuals can be said to belong to the same practice depends directly on whether they possess *shared* categories of perception and appreciation. Informally, we might say: Do they speak the same language? It also should be stressed that dispositions are, in the main, tacit (Polanyi 1997; Schön 1983); that is, we should not make the assumption that practitioners are consciously aware of the dispositions that shape their world-view. Rather, recent scholarship (Bleakley 2000; Brown 2005; Tsoukas & Hatch 2005) has argued that practitioners' knowledge is stored in the form of narratives of practice. For

---

[2] I discuss other common uses of the word "disposition" (and related terms, such as "values", "beliefs", and "habits of mind") in Chapter 2.

this reason, I seek to reveal the mathematical sensibility through a review of narratives of mathematicians at work.

This review then provides a grounded basis for the second part of the dissertation (Chapter 3), in which I investigate whether the practice of geometry teaching provides a customary place for the elements of a mathematical sensibility. Put another way, I seek to understand whether the categories of perception and appreciation that comprise the practical rationality of *mathematicians* are also found within that of *Geometry teachers.* As a practice, the teaching of Geometry is characterized by its own practical rationality (Herbst & Chazan 2003b), which should not be presumed to be the same as that of research mathematics. Fig. 1.2 shows a schematic representation of the question. To be clear, the analysis in Chapter 2 yields a set of dispositions that are *emic*, i.e. they attempt to describe mathematical practice from a practitioner's own point of view; these dispositions are then deployed in Chapter 3 as *etic* constructs to test the extent to which the dispositions of mathematicians map on to corresponding dispositions within teachers' own practical rationality. (For an overview of the distinction between emic and etic research paradigms, see Guba & Lincoln 1994).

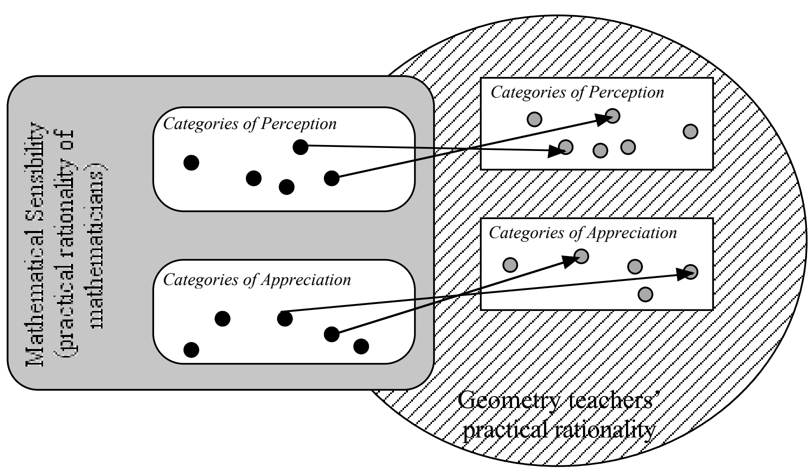


Figure 1.2. Two practical rationalities.

Towards this end I analyze records from study groups composed of experienced geometry teachers, and ask to what extent the dispositions identified in the earlier parts of the dissertation are visible in those records, whether they appear to be valued by teachers, and what the attendant costs are for teaching that attempts to make room for such dispositions. This study makes use of data collected within the NSF-funded research project “Thought Experiments in Mathematics Teaching”, or ThEMaT (Herbst & Chazan 2003b, 2006). In the ThEMaT project, teachers are gathered in monthly study groups to view, enact, and discuss representations of teaching in the form of animated cartoons, printed comics, and other media. These representations depict students and teachers in situations that include a blend of real and invented narrative elements: they narrate stories of instruction that are conceivable, but not necessarily probable or desirable. The representations are then used as research probes; teachers’ respond to the representations of teaching by narrating their own alternative stories, or by articulating the circumstances under which the story would be viable, desirable, etc. The records of these conversations yield a corpus of data that can be mined for analysis of teachers’ own practical rationality.

This analysis of the study group data helps to establish a baseline against which the third part of the dissertation is set (Chapters 4-5). In that part, I analyze a corpus of examination questions, written by a secondary mathematics teacher and used in three Honors Geometry classes over a four-year period, and attempt to describe how (some of) these tasks may be understood as embodying an opportunity to learn not only mathematical content, but also some of the elements of the mathematical sensibility that are characteristic of research mathematicians. The same set of dispositions generated in the first part of the dissertation is used as a set of codes for the items in this third part. As part of the analysis, I study how the tasks in the corpus cohere across the timescale of the year by explicitly revisiting past content and connecting new problems to heuristics that have been successful in the past. Additionally, I study the evolution of certain tasks (including the abandonment of some and their replacement by others) and changes in the large-scale structure of assessments over the three-year history of the corpus to illustrate how the nature and goals of assessment adapted over time. This provides indirect evidence as to the factors that obstructed or resisted the goal of holding students accountable for learning the mathematical sensibility.

An example

At this point a concrete example may help to make my point. One of the dispositions that emerge from the analysis of mathematicians' narratives is a propensity for wondering whether under specified conditions there exists a certain kind of mathematical "thing" (a number with certain properties, a set with certain structure, etc.). An essential part of the experience of working on such questions is that the answer is genuinely in doubt; the goal is not merely to prove a claim that is already known to be true, or to provide a counterexample for a claim that is already known to be false, but to *settle* the question. Shiryaev (2000) recounts how on several occasions the mathematician A.N. Kolmogorov produced counterexamples to claims that were widely believed to be true, and even produced counterexamples to his own conjectures, which served to advance his long-term research goals even as they frustrated his own short-term aims. (Details may be found in Chapter 2.) Work of this nature is common in mathematics; the whole of Lakatos's (1976) *Proofs and Refutations* foregrounds the role that such existence questions and counterexamples can play in advancing mathematical knowledge.

Data from the ThEMaT study groups supports the contention that existence questions of this sort play little role in most Geometry teachers' practice. For example, in a discussion of a ThEMaT animation in which students work for an extended period with a diagram of a figure that could not exist (i.e., a figure that has been overlaid with a system of markings that are mutually inconsistent with each other and with the class's usual convention for handling geometric diagrams), teachers expressed consistently that questions of the sort "Is this possible?" are uncommon in their classroom. Teachers agreed that, if a proposal to consider such an impossible geometric object were to come from the mouth of a student (as may unavoidably happen from time to time), it would be preferable to handle such an event by asking "What's wrong with this figure?" (signaling immediately that the figure is defective). Moreover, they report that they would only be inclined to call the possibility of a figure into question if the figure were, in fact, impossible; that they would typically pose such a question in terms of diagrammatic objects, rather than in terms of concepts (Weiss & Herbst 2007); and that they would prefer to avoid such questions altogether, rather than seek out opportunities to ask them.

In marked contrast to this, questions of existence can be seen to play a significant role in the corpus of examination questions that I examine in chapters 4 and 5. An analysis of the assessments from the 2003-2004 school year shows that each test contains one or more questions that ask whether something is possible, or, alternatively, ask under what conditions something would be possible. Some examples include:

- If an angle is bisected by a ray, can either of the two sub-angles ever be obtuse? Justify your answer. (Chapter 1 Test, #7)
- Suppose *S* is a set of points that are non-collinear. What is the minimum possible number of points in *S*? (Chapter 1 Test, #9a)
- Is it possible, using only a straightedge and compass, to construct a 45° angle? Describe how in words, and demonstrate below. (Chapter 1 Test, #12).
- Make up your own example of a sentence which is false, but has a true contrapositive. If this is not possible, explain why. (Chapter 2 Test, #6).
- If two lines intersect, can they be non-coplanar? If yes, what are lines like this called? (Chapter 3 Test, #7).
- If two lines are coplanar, can they be non-intersecting? If yes, what are lines like this called? (Chapter 3 Test, #8).
- Can a triangle be both isosceles and obtuse? If so, draw an example and indicate the measure of the three angles. If not, explain why. (Chapter 4 Test, #1)
- Determine whether it is possible or not to draw a triangle with sides of the given lengths. In each case, write *P* if it is possible, or *N* if it is not possible. If not possible, state clearly why. [3 proposed triplets of lengths follow.](Chapter 5 Test, #4).
- Is it possible for a regular polygon to have every angle equal to 155°? If so, how many sides must it have? If not, why not? (Chapter 6 Test, #4)

More examples could be offered; in fact, out of 545 points possible on these six tests, fully 81 points (approximately 15%) were constituted of questions of this type. Furthermore, roughly half of these questions describe a mathematical possibility, while the remainder describe objects or configurations that are impossible. In this set of items, then, the question "Is it possible?" is a manifestly nontrivial one, in that the mere posing of the problem does not automatically signal the correct answer.

It is also worth noting that, while all of these questions call for a proof of one kind or another, they are written using a kind of language that is different from those commonly associated with proof exercises. The latter are typically posed in reference to objects represented as diagrams, while the examples above primarily use verbal descriptions that denote abstract classes of objects. These two registers normally play different functions in the geometry classroom: it has been shown (Weiss & Herbst 2007) that the diagrammatic register is commonly associated with proof exercises, while the conceptual register is reserved for the teaching of theorems. The use of the conceptual register in the context of problems that call for proof is thus itself a departure from what is normal in Geometry.

For purposes of comparison, the "chapter tests" printed at the end of every unit of the course textbook (Larson, Boswell and Stiff 2001) were coded using the same procedures; of the 229 items in the textbook corpus, *none* asked questions of existence.

There is thus strong empirical evidence to support the claim that, in the classroom in which the assessment items were used, questions of existence played a significant role, corresponding to the role such questions play in mathematical practice; and there is further evidence that this constituted a significant deviation from the norm, as articulated by the ThEMaT study group participants and as represented by the chapter tests printed in the textbook. In the chapters below, I provide additional examples of such deviations, corresponding to the other dispositions identified in Chapter 2.

Discussion

The goal of this dissertation, then, is to stage an assault on the principal research question by attacking it from three directions: a conceptual analysis of the "mathematical sensibility", an empirical study documenting Geometry teachers' customary practice vis-à-vis teaching the sensibility, and a second empirical study examining both how teaching can hold students accountable for the sensibility and what happens when this is tried. Before undertaking this inquiry, it is appropriate to briefly anticipate what they might yield, and how those proceeds connect with the larger questions touched on at the beginning of this chapter.

A study focused exclusively on the high school Geometry class may seem to have only limited relevance for those who are concerned with the larger-scale problems of K-12 mathematics education. But, as I have noted above, the Geometry class is traditionally students' closest point of contact with many of the activities of mathematics, and as such it seems reasonable to expect that it would be the most likely context in which students could find opportunities to learn the mathematical sensibility. Put another way, any challenges to teaching of the mathematical sensibility in a Geometry class are likely to be all the more difficult to overcome in other contexts. Educators who call for a renewed emphasis on mathematical reasoning and practices in the elementary grades may find that the issues raised in this dissertation only scratch the surface of the issues that will be faced in those settings.

On the other hand, it may be valuable to consider the (generally accepted) perspective that educational experiences in the formative years can have a profound influence on children's later intellectual development. If so, then the challenges faced in trying to teach a mathematical sensibility to high-school Geometry students might, in part, be attributed to a failure to teach the sensibility at an earlier stage: put simply, one might wonder whether by the time students reach Geometry, it might be *too late* to cultivate in them a mathematical sensibility. If this is so, then the work reported in this dissertation has an additional value: by articulating in some detail what the mathematical sensibility consists of, it can help to start a conversation about what it would mean to make those elements part of classroom practice in a developmentally appropriate way at the elementary school level.

Finally, it seems appropriate to consider these matters from the perspective of teacher education. To what extent can teacher education programs cultivate a mathematical sensibility *in the teachers themselves*? Readers who are concerned with improvements in teacher education may benefit from considering the high school Geometry course as an analog of the "Geometry for teachers" course taken by most preservice teachers. The illustration of how the former course can take on the responsibility for teaching students a mathematical sensibility may serve as a model for how a similar goal can be undertaken in the latter course; and the documentation of the challenges of doing so in the high school course may help teacher educators to anticipate and plan for the difficulties that will likely face themselves.

Chapter 2

A mathematical sensibility: Analyzing mathematical narratives to uncover mathematicians' categories of perception and appreciation

Mathematical practices and mathematical narratives

In subsequent chapters of this dissertation I will undertake an exploration of the question: *Can a secondary geometry course provide a context in which students can be taught what it is like to be a mathematician*? This chapter lays the groundwork for that exploration by answering a necessary preliminary question: What, after all, does it mean to "be like a mathematician"? In this chapter I attempt to answer this question with some precision through a detailed analysis of a number of *narratives of mathematicians' practice.* Before undertaking this, however, it is necessary to clarify a few key terms and providing some theoretical grounding for the method of analysis. Thus I now take up three basic questions:

(1) What is meant here by *practice*, and why should educators care about the practice of mathematicians?

(2) What kind of *narratives* are meant, and why should one expect them to be of use in answering the question of what it is like to "be a mathematician"?

(3) Who are the mathematicians whose narratives will be considered, and by what criteria is that sampling of mathematicians to be determined?

The relevance of the first question stems from the perspective that "learning mathematics" can (and ought to) be understood not only as the acquisition of content knowledge, but also as the enculturation into a collection of practices that are characteristic of the community of mathematicians (Lampert 2001; Sfard 1998; Wenger 1997). That is, "learning mathematics" can be understood to include (among other things) "learning to be like a mathematician". Of course "being like a mathematician" requires extensive subject matter knowledge — mathematicians know things about numbers and shapes and logic, and anyone who is to be like a mathematician certainly needs to learn those things as well. But there is more to knowledge of mathematics, or indeed any practice, than can be expressed in the form of declaratory statements. Additionally there is the tacit knowledge that is encoded in, and expressed through, one's actions (Polanyi 1997; Schön 1983). "Learning mathematics" in this sense includes a behavioral component, in which the learner comes to behave like the members of the community of practice. In much the same way that "learning chess" includes not only coming to know the rules of the game but also the development of a sense of what distinguishes a good position from a poor one, when it is appropriate to sacrifice a piece in order to gain advantage on the board, anticipating what one's opponent is likely to do, and so forth (Chase & Simon 1973; de Groot 1978), "learning mathematics" includes a sense of what kinds of moves are timely and appropriate in a given circumstance.

It would, however, be over-stating things to suggest that the practice of mathematics includes only what one does; it also includes what one perceives, and how one values it. To stay with the example of a chess player, for example: when players watch a game in which they themselves are not playing, they nevertheless make judgments about how the game is going, what ought to or is likely to happen next, which player is stronger than the other, and so forth. That is to say, a practice consists not only of actions but also contains (and is from one point of view defined by) a set of lenses through which the practitioner looks out on the practice. One who is not enculturated into the practice — a novice player, say — is likely to lack some of those lenses, and thus be unable to see what the practitioner sees.

In the view of the French sociologist Bourdieu (1998), a practice is characterized by a habitus or practical reason composed of a set of shared categories of perception and appreciation, which he refers to as “dispositions”. The word “disposition” has a wide range of uses in education research (Diez & Raths 2007; Dottin 2009). The word is frequently used to refer to a propensity to act in a certain way under certain circumstances, as “tendencies or inclinations to act in particular ways” (Feiman-Nemser & Remillard; Schwab 1976). Others use the word “disposition” as a virtual synonym for “values” (e.g. Misco & Shiveley 2007) or “habits of mind” (e.g. Bass 2005). Here I use the term in the more technical sense of Bourdieu’s reflective sociology. In this usage, “dispositions” are an observer’s reconstruction of the tacit regulatory mechanisms of practice. As such they are “neither individual commitments nor institutional requirements; they are *like* requirements in so far as they create a sense of intersubjective normality but they are implicit; they are *like* commitments in that they also accommodate personal preferences, though they are also transposable among people who do similar work” (Herbst & Chazan 2003a).

Taking this notion as a starting point, I seek to describe part of the network of dispositions that characterizes the community of mathematicians. This construct, the practical rationality of mathematicians, stands in this dissertation as my proxy for what Ball and Bass (2003) have referred to as a "mathematical sensibility". This sensibility makes it natural for mathematicians to ask certain kinds of questions rather than others, to work toward answering those questions using certain methods rather than others, and to value certain kinds of results over others. The "categories of perception" are the metaphorical lenses through which a mathematician views mathematical practice; the "categories of appreciation" are the metaphoric scales with which the mathematician appraises his work and that of others.

To be still more precise, I here take on the problem of describing the dispositions that characterize a particular *zone* of mathematical practice: namely, that part of mathematical practice concerned with the perception and appreciation of *results*, where a "result" is understood as either an implication of the form "*P* implies *Q*", or as a "non-implication" of the form "*P* does not imply *Q*". I also include in this zone the posing of mathematical problems, which could be considered partial anticipations of results: they may take the form "Does *P* imply *Q*?", or "What *P*'s imply *Q*?", or "What *Q*'s are implied by *P*?"

I note that in demarcating this particular zone of mathematical practice, I deliberately exclude other significant zones of practice. I bracket almost entirely the vast zones of mathematical problem-solving and proving, of defining and symbolization, and of discourse and conviction. Much scholarship has attended to these areas of mathematical activity. For my present purposes, they will be treated as a kind of black box. I am interested in how the inputs and outputs of mathematical work — respectively, the problem posed and the result obtained — are perceived and appreciated by mathematicians; the work that mediates between the two belongs to other studies.

In stressing the *practices* of mathematicians (as opposed to the results eventually produced by those practices) this study stands in contrast to the view of mathematics contained, for example, in Wu (1999). In that work, "the characteristic features of mathematics" are said to be "precise definitions as starting point, logical progression from topic to topic, and most importantly, explanations that accompany each step" (p. 5). Wu recognizes that this work may be preceded by a more informal stage in which one encounters mathematical concepts in a less orderly fashion, but he likens this to the scientist's "data collecting phase", and is not to be regarded as mathematics *per se*. Here, I wish to argue to the contrary that the work of navigating around in an ill-defined conceptual space, generating questions in the context of provisional definitions, in fact lies at the heart of what it means to be a mathematician. The mathematics described by Wu is not the practice of mathematics, but rather the result of a practice that erases its own tracks (Lakatos 1976). This, paradoxically, is itself a characteristic of the mathematical sensibility, which publicly values product over practice, but which requires practice to generate product.

The goal of this chapter is to identify some of the elements of this mathematical sensibility. Towards this goal, it is important to consider what sources of information might contribute to that understanding. There is of course no shortage of scholarly works on the history of mathematics, the philosophy of mathematics, and so forth. One might also survey mathematicians' own research publications, monographs, textbooks, and so forth with the intention of trying to extract from them some sense of how mathematical knowledge grows and changes over time; this is essentially the method employed to great effect by Lakatos (1976). Here I take a different approach. In recent years many scholars have argued forcefully that the wisdom of a practice — that is, the network of explicit and tacit knowledge that practitioners draw upon in the course of their activity — is not primarily stored in the form of declarative, general principles (as one might find in a philosophy of mathematics or in a mathematics textbook), but rather in the form of *narratives of practice* (Bleakley 2000; Tsoukas & Hatch 2005). John Seely Brown (2005) articulated the distinction between learning disciplinary content and disciplinary practices thus:

> Learning has to do, not only with *learning about* something — we all know how to *learn about* something by reading books and so on — but also with, how do you *learn to be*? There's an immense difference between learning *about* and learning *to be*... How can you *be* a physicist? How can you *be* a doctor? How do you enculturate someone into the profession? There's a massive amount of tacit practices and sensibilities and lenses that we use to see and make sense of the world and act effectively in the world. .... You can never talk someone rationally through a change in religion. You design or craft experiences. You go to the gut. That's what stories can do. They may be able to help us unlearn. (pp. 56-7, emphasis in original.)

If Brown is right, then our best hope of understanding what it means to *be* a mathematician may be through close examination not of mathematics publications (textbooks, monographs, and the like), nor works on mathematical philosophy, but rather through narratives about mathematicians and their work.

Such narratives come in many forms. On the one hand there are first-person memoirs written by mathematicians, and third-person biographies written about them. There are also fictional accounts, as in the theatrical play *Proof,* the motion picture *Good Will Hunting*, and the TV series *Numb3rs* (all three of which, it should be noted, employed research mathematicians as script consultants). Straddling these categories are semi-fictionalized biographies (such as the film version of *A Beautiful Mind,* loosely based on Sylvia Nasar's non-fictional account of the life of John Nash), as well as historical reconstructions narrated in a quasifictional form (as in Lakatos 1976). Each of these could be regarded as a "mathematical narrative" in one sense or another. To make progress towards defining a coherent sample of narratives for analysis, some criteria for inclusion are required.

The first criterion I follow is perhaps obvious, but it deserves emphasis. I am interested here in *narratives about mathematicians and their work*, not simply narratives *by* mathematicians or narratives that happen to have mathematical content. Thus I will not consider "mathematical fantasies" such as Edwin Abbot's *Flatland* or its many sequels, or the dialogue portions of Hofstadter's (1979) *Gödel, Escher, Bach*; nor will I consider works of fiction that happen to have been written by mathematicians, such as Lewis Carroll's *Alice* books or Martin Gardner's short stories and novels. For complementary reasons I will likewise exclude from consideration other work that does focus on mathematicians' practices, but in a way that disconnects it from the biography; e.g., Polya's (1957) *How to Solve It* or Rothstein's (1995) *Emblems of mind: The inner life of music and mathematics.*

The second criterion I employ to define the sample is in a sense the reverse of the first one: I am explicitly seeking narratives about mathematicians' *work*, not the individuals themselves. That is, I am not particularly interested in the biographical details of the subject's personal life, except insofar as those details may help to illuminate the values and practices that undergirded his or her mathematical activity. For this reason a biography that does not provide at least some descriptive detail about the subject's mathematical work is of little use here. On the other hand, I am after more than just a collection of major mathematical results, as one might find in a *festschrift* survey of a scholar's research output; rather the goal is to try to see, to the extent possible, how those results came about, how the mathematician worked on them. (Or, to be even sharper: the target is the after-the-fact *reconstruction* of how a mathematician poses, works on, and values problems.) For this reason I also exclude any work that does not include at least some element of biographical detail.

A third criterion I adopt is to consider only *modern* mathematicians in the current sample — where by "modern" I mean mathematicians who did the bulk of their work in or after the 20th century. This criterion is based in part on the pragmatic need to bound the extent of the literature reviewed, but also rests on theoretical considerations. Mathematics has been a human activity for thousands of years, and it would be naive to assume that the practices and values that characterize that activity have remained unchanged over such a great span of time. Indeed it is well-known that the history of mathematics shows evidence of profound shifts in mathematicians' notions of what constitutes a well-formed definition, proof, and so forth (see Grabiner 1974 for an exemplification of this point). Thus, while there may be much that Euclid, Descartes, and Erdös have in common, any characterization of the practice of mathematics that attempts to be all-encompassing is likely to be too general to be of much value for the present study.

This brings us to perhaps the most difficult question of all, namely: Just what is meant by "mathematician", after all? The umbrella term "mathematician" is quite broad, encompassing a variety of different professional and amateur practices. Definitions range from the nearly tautological (as in Merriam-Webster's "a specialist or expert in mathematics") to the thoroughly whimsical (as in Alfréd Rényi's much-quoted, but undocumented, quip "A mathematician is a machine for turning coffee into theorems"). For the present purpose it is helpful to narrow the focus by distinguishing between those who might be said to be *users* of mathematics (a broad field encompassing most scientists, engineers, accountants, actuaries, etc.) and those who can be considered *producers* of mathematics, and to restrict attention to the latter category. Of course this distinction is somewhat artificial, as any significantly novel use of mathematics may arguably be considered as a contribution of something new to the field. Still, as a heuristic it seems reasonable for the present study to understand the term "mathematician" as referring to an individual whose primary occupation is the generation of new mathematical knowledge. In other words, I am primarily concerned in this chapter with mapping out the dispositions that characterize the practice of *research mathematicians*[3].

What remains, after all of the above considerations are taken into account, are narratives of two types: (a) *memoirs*, as in, e.g., Hardy (1941/1992), Wiener (1956), and

---

[3] This criterion should not be misunderstood as a denigration of "applied mathematics" as in any way subordinate or inferior to "pure mathematics". On the contrary, narratives of both applied and pure mathematicians may be included in the sample, with the proviso that in the case of applied mathematics we are interested in the ways novel mathematical methods are generated and refined, and *not* primarily in how those methods are applied to solve disciplinary problems in areas other than mathematics. Of course the two cannot be entirely disentangled from one another, but my focus is on the mathematics, not its application.

Poincaré (1913/2001); (b) and *biographies,* as in Bell (1965), Kanigel (1991), and Parker (2005). A full list of the narratives that made up the sample for this analysis can be found in Figure 2.1, below. No claim is here made that Figure 2.1 represents an exhaustive survey of the literature. The goal in the present chapter is not to produce a rigorous, evidence-based profile of mathematicians' practical rationality, but rather to develop and flesh out my conceptualization of the construct "mathematical sensibility" by using these narratives as a source of inspiration and a context for testing out concepts as they develop.

MEMOIRS

Author	**Title**
Hardy, G.H.	*A mathematician's apology*
Ulam, S.	*Adventures of a mathematician*
Wiener, N.	*I am a mathematician.*
Davis, P.	*The education of a mathematician.*

BIOGRAPHIES

Subject	**Title and author**
Zariski, O.	*The unreal life of Oscar Zariski* (C. Parikh)
Kolmogorov, A.N.	*Kolmogorov in perspective* (History of Mathematics series, American Mathematical Society and London Mathematical Society)
Moore, R.L.	*R.L. Moore: Mathematician and teacher* (J. Parker.)

Figure 2.1. The sample of mathematical narratives.

Emergent themes: An overview of the mathematical narratives

In subsequent sections of this chapter, I will analyze in close detail the narratives listed in Fig. 2.1, seeking common themes. From this analysis, certain recurring elements of the mathematical sensibility emerge. Here I give an overview and brief discussion of three of those themes, which will be exemplified later in the subsequent detailed analysis.

The first of the emergent themes is perhaps the most basic. Again and again, the narratives portray mathematicians as individuals who not only work to solve problems, but also are characterized by a propensity to *pose* mathematical questions to and for themselves (that is, to wonder what is true); in other words, mathematicians are *curious*. This may seem self-evident. And yet a major question of the next two chapters is whether schooling can provide students with an opportunity to wonder as mathematicians do (cf. my critique of the "Lei" vignette in Chapter 1).

Beyond the observation that mathematicians pose problems, the narratives provide evidence of certain standard *ways* in which they pose problems. I describe in detail a model of mathematical problem-posing in which new questions, in the form of possible relationships between contingent possibilities, are generated from already-known relationships by an iterative process of modifications to the hypothesis and conclusion of a conditional statement. Several variations of this process will be discussed, each based on a close reading of a passage from the mathematical narratives. While I argue that these processes are typical of the way mathematicians work, they are not necessarily part of mathematicians' explicit mathematical knowledge; they are rather part of the practitioner's tacit knowledge, what Schön (1983) called "knowing-in-action". In what follows I will explicate these processes in detail. I will also provide examples of how they could conceivably be used to generate problems in the context of a secondary course in Geometry.

But not all of the problems mathematicians pose are viewed as equally deserving of effort being expended on them. Some are never pursued; and even among those that are solved, not all results are equally valued. Mathematicians employ logical criteria for judging the relative quality of mathematical claims (what makes some claims stronger or weaker than others). In addition to these logical criteria, mathematicians also possess (and for the most part share) aesthetic criteria (Sinclair 2002) that regard certain problems and results as more "beautiful", "elegant", or "deep" than others. Among these criteria are *simplicity* and *unexpectedness*. As will be shown below, another mathematical value is an orientation towards *theory-building*. By this I mean that mathematicians place a high value on work that clarifies or reconfigures the organization of a large collection of mathematical propositions by foregrounding the interrelationships among them. This disposition, I show below, is instantiated in narratives that value the organization of a mathematical theory (the classification of certain propositions as axioms and others as theorems, the definitions of key concepts, the sequence in which topics are covered and theorems proved, etc.) as always provisional, tentative, and subject to revision. All of these categories of perception and appreciation will be based on examples from the mathematical narratives.

It is worth noting here that the themes above stand in approximate correspondence to Bourdieu's (1998) "categories of perception and appreciation". The analysis of the processes by which mathematicians generate new questions aims at a description of how mathematicians *perceive* new problems as emerging from the milieu composed both of that which is already known to be true and of the other problems that have already emerged: the categories of perception are the actions that a mathematician sees as available to do next. The categories of appreciation, on the other hand, are those values with which mathematicians judge the relative worth of those different possible courses of action, as well as judge the worthiness of expending effort on new results and new questions. Understanding how mathematicians pose problems, and how they appraise both problems that have been posed and solutions to those problems, is thus tantamount to mapping out the practical rationality that mathematicians bring to bear when engaged in posing and selecting problems to work on.

These three themes — that mathematicians are problem-posers, that they employ certain generative moves to produce new questions, and that they employ multiple categories of appreciation to value both problems and solutions — structure my analysis below. I certainly do not claim that there is nothing more to a mathematical sensibility other than what fits within these three broad themes. On the other hand it seems that any account of what it means to "be like a mathematician" ought to at the very least include all of these themes, and consequently it is worth investigating the question of whether a secondary course in Geometry can provide a context in which these dispositions in particular can be cultivated. That question will occupy my attention in subsequent chapters of this dissertation.

Encounters with the unknown: Mathematicians as problem-posers

I begin by recounting two anecdotes culled from the collection of mathematical narratives I analyzed. The first anecdote is taken from Stanislaw Ulam's (1976) account of a 1937 road trip from Princeton to Duke University accompanied by John von Neumann. In this recount, Ulam intertwines his encounters with American culture, reflections on ancient and modern warfare, personal details of Von Neumann's married life, and — not least — musings on various mathematical and metamathematical issues:

> On the way back from the meeting I posed a mathematical problem about the relation between the topology and the purely algebraic properties of a structure like an abstract group: when is it possible to introduce in an abstract group a topology such that the group will become a continuous topological group and be separable?.... The group, of course, has to be of power continuum at most — obviously a necessary condition. It was one of the first questions which concern the relation between purely algebraic and purely geometric or topological notions, to see how they can influence or determine each other.
>
> We both thought about ways to do it. Suddenly, while we were in a motel I found a combinatorial trick showing that it could not be done. It was, if I say so myself, rather ingenious. I explained it to Johnny. As we drove Johnny later simplified the proof in the sense that he found an example of a continuum group which is even Abelian (commutative) and yet unable to assume a separable topology.... Johnny, who liked verbal games and to play on words, asked me what to call such a group. I said, "nonseparabilizable." It is a difficult word to pronounce; on and off during the car ride we played at repeating it. (p. 103)

The second anecdote comes from Wiener's (1956) account of his first years as a young faculty member at M.I.T. Wiener describes how he sought mathematical inspiration from the landscape:

> It was at M.I.T. too that my ever-growing interest in the physical aspects of mathematics began to take definite shape. The school buildings overlook the River Charles and command a never changing skyline of much beauty. The moods of the waters of the river were always delightful to watch. To me, as a mathematician and a physicist, they had another meaning as well. How could one bring to a mathematical regularity the study of the mass of ever shifting ripples and waves, for was not the highest destiny of mathematics the discovery of order among disorder? At one time the waves ran high, flecked with patches of foam, while at another they were barely noticeable ripples. Sometimes the lengths of the waves were to be measured in inches, and again they might be many yards long. What descriptive language could I use that would portray these clearly visible facts without involving me in the inextricable complexity of a complete description of the water surface? This problem of the waves was clearly one for averaging and statistics, and in this way was closely related to the Lebesgue integral, which I was studying at the time... (p. 33)

In many ways these two anecdotes are quite different. Ulam's account is of mathematics emerging in a discourse that is essentially a *social* one: the friendly give and take of ideas with Von Neumann provides the context for mathematical inquiry. Wiener's account, on the other hand, is one of a solitary thinker pondering the natural world around him. Of course one may well regard cognition of this sort as a kind of internal discourse (Sfard 1998), but the fact remains that Ulam's discourse comes across as qualitatively very different from that of Wiener.

Moreover the accounts differ to the extent that the authors attach importance to the experiential world as an object of inquiry for mathematics. Wiener is inspired by the natural world around him, and ascribes great value to the problem of finding a mathematical language for the description of complex physical phenomena; in stark contrast Ulam's musings on the algebraic and topological properties of groups seem to be entirely disconnected from the world of material experience.

Nevertheless the two accounts have one key feature in common, a characteristic that is present in all of the mathematical narratives I have analyzed: That is, they present the mathematician as someone who *wonders mathematically.* By this I mean that the mathematician works on mathematical problems not out of obligation or coercion but out of an intrinsic curiosity to know what the truth is. Of course, the mathematician may also be motivated by other concerns — e.g., by a desire for professional advancement and prestige — but at least part of what drives the mathematicians' inquiry is the desire to know what is true. The questions a mathematician wonders about are typically ones for which no answer is currently known, and which may resist attempts at a solution for months, years or even centuries.

Where does this curiosity come from? More precisely, to what do practitioners of mathematics attribute their own curiosity? In asking this question I am fully aware that mathematicians, like all of us, are prone to self-deception and selective memory. However, to the extent that mathematicians may report, in their narratives, stories of how their interest in mathematics was kindled, it seems worthwhile to attend closely to those recounts. It may be facile to take at face value a story that a mathematician tells about a childhood episode; we certainly have no basis for attributing any kind of causal link between childhood experiences and subsequent mathematical expertise in adulthood. But what such narratives *can* offer us is insight into what adult mathematicians value about their practice. Therefore I examine these stories not to study the embryology of the child mathematician, but rather to understand the sensibility of the adult mathematician reflecting back on formative mathematical experiences.

Ulam recounts how, as a child, he was "intrigued by things which were not well understood" . He enumerates among his early interests the question of the shortening of Encke's comet (an astronomical phenomenon that could not be explained by the physics of the day), Einstein's theory of special relativity (which, to an 11-year old Ulam, was fascinating precisely because none of the adults in his life understood it), Euler's *Algebra* (in which the symbolic notation "looked like magic signs" and gave him "a mysterious feeling"), and the question of the existence of odd perfect numbers (an unsolved problem in Ulam's adolescence, and today). These examples include both "things which were not well understood" *by him* and "things which were not well understood" *in general*. Ulam reports that these encounters with the unknown provoked in him a deep desire to understand, and by the age of 11 he had begun to record his mathematical discoveries in a notebook. In contrast to Ulam's own exploration of these mysteries, the mathematics taught in school seemed to him to be "dry" and consisted of memorizing "formal procedures".

Ulam's description of encountering the mysterious in mathematics finds parallels in other narratives. For example, Davis (2000) narrates how a childhood encounter with number games ("Take a number, any number. Double it, add three, reverse the digits. Tell me what you get. Ah then, the number you first picked was such and such") sparked a curiosity that led him not only to figure out why the games worked, but also to invent new variations of his own. Like Ulam, Davis narrates how, as a child of 11 or 12, he used to record his mathematical discoveries in a notebook that had been given to him by an older cousin: "Things such that the sum of two odd numbers is always even. The product of three consecutive numbers is always divisible by six. By the time I went to college, Cousin H.'s lab book contained what is known in mathematics as the Newton formula for polynomial interpolation, a substantial but by no means remarkable achievement" (pp. 7-8). A similar anecdote is told about the Russian mathematician A.N. Kolmogorov (Shiryaev, 2000, p.4), who wrote of noticing "at the age of five or six" the pattern $1 = 1^2$, $1+3 = 2^2$, $1+3+5 = 3^2$, $1+3+5+7=4^2$, and so on. The discovery was printed in a school newsletter, together with arithmetic problems that Kolmogorov, like Davis, devised on his own initiative.

Although Ulam, Davis, and Kolmogorov write of nurturing their curiosity outside the formal boundaries of schooling, there is also evidence from the mathematical narratives that vibrant and creative instruction can play a critical role in cultivating this curiosity. In particular John Parker's (2005) biography of R.L. Moore suggests that instruction can spark mathematical curiosity in those who do not fit the "child prodigy" mold. Moore, of course, is as well-known for his innovative teaching methods as he is for his impressive body of mathematical research. In fact many of Moore's unconventional teaching methods seem to have been intended precisely to stimulate mathematical curiosity in those who might not otherwise have pursued careers in research mathematics. Parker documents, for example, how many of Moore's eventual doctoral students originally intended to study other fields, but "were 'caught' by happening to take one of his advanced courses as an elective to round out their main objectives," which included actuarial mathematics (R.L. Wilder), chemistry (G.T. Whyburn and F. Burton Jones), medicine (John Worrell), and so on. Referring to this years later, Moore's student R.L. Wilder remarked, "If any proof were needed that the *capability* of doing creative work in mathematics is not the rare genetic accident that it is commonly considered, Moore certainly gave it during his career as a teacher" (quoted in Parker, p. 131).

It is perhaps no coincidence that this uncommon success at stimulating mathematical curiosity occurred in an educational setting that was, by design, equally uncommon. Moore would ask his students to prove plausible-seeming claims that were subsequently revealed to be false; he would assign his classes sets of theorems to be proved, with unsolved problems concealed among them. In forbidding his students from reading textbooks, and in forcing them to develop the body of a theory on their own, Moore attempted to lead his students to confront the same qualities of mystery in mathematics that Ulam and Davis discovered in childhood.

I close this section with one final illustration of how an encounter with the unknown can stimulate one's curiosity in mathematics. In Davis's (2000) reflections on his high school years, he recounts how an appendix to one of his schoolbooks contained a "proof" that every triangle is isosceles. Realizing the result was nonsense, Davis set out to painstakingly check the details of the proof, ultimately determining that the error lay not in the proof itself but in a mistaken diagram on which the proof relied. Davis compares his experience with that of the philosopher Thomas Hobbes, as reported in John Aubrey's *Brief Lives*:

> He [Hobbes] was 40 yeares old before he looked on Geometry; which happened accidentally. Being in a Gentleman's Library, Euclid's "Elements" lay open, and 'twas the 47 *El. Libri I*. By G—, sayd he (he would now and then sweare an emphaticall Oath by way of emphasis) this is impossible! So he reads the Demonstration of it, which referred him back to such a Proposition; which proposition he read, That referred him back to another which he also read. *Et sic deinceps*. That at last he was demonstratively convinced of that trueth. This made him in love with Geometry.

What unites Hobbes's experience with Davis's, and finds resonance in the teaching methods pioneered by Moore, was the role of *doubt* in stimulating mathematical curiosity. Davis, like Hobbes before him, faced a claim that he disbelieved; it was the desire to *disprove* that led them down unexplored pathways of logical reasoning. Moore's students learned to doubt what they were told and to distrust any argument that was not thoroughly rigorous:

> He [Moore] often told a class to prove something that he knew was not true, for example to prove that if a point set is closed so is its projection onto the x-axis. Moore asked: Isn't this much better than to tell them to prove that the projection onto the x-axis of a closed and bounded point set is closed? Why should any teacher want to follow the latter procedure and therefore deprive a student of the opportunity to discover independently that one of these propositions is true and the other one is false? (Parker, 2005, p. 262)

What this approach to teaching takes for granted — and what, in a sense, is the essence of mathematical curiosity — is that discovering the truth is “an opportunity” to be embraced, not a burden to be borne. Encountering such opportunities requires the discoverer to work precisely in regions where the truth is unclear and in doubt.

Moves for generating new questions

In the pages above I have argued that a mathematical sensibility is characterized by a propensity to pose questions to, and for, oneself — that is, by the tendency to *wonder about what is true.* In this section I attempt to sharpen the discussion by describing the kinds of questions mathematicians ask, and characterizing some of the ways in which old questions (and their answers) give rise to new ones.

As a point of entry I return to the excerpt from Ulam (1976), cited above, and step slowly through his account of his exchange of ideas with von Neumann. As we will see, Ulam and von Neumann consider not just one question but a whole series of them, each produced by a modification of the one preceding it. Based on these modifications I will describe a series of “generative moves” that, I argue, are typical of the ways in which mathematicians pose new problems.

Ulam begins by posing to his driving companion the question, “When is it possible to introduce in an abstract group a topology such that the group will become a continuous topological group and be separable?” In other words Ulam seeks after conditions on a group G that are sufficient to ensure that G is “separabilizable”, i.e. that it may be endowed with a separable topology[4]. In a sense Ulam is attempting to "solve" for the unknown predicate P in the implication

$$P(G) \rightarrow \text{Separabilizable}(G) \tag{1}$$

Mathematics has been described as "the science that draws necessary conclusions" (Peirce, 1882). From this point of view, what matters most to a mathematician is whether two contingent propositions stand in a relationship of logical implication together. That is to say, a mathematician considers a certain category of mathematical objects, and certain properties (say P and Q) that may or may not be held by each of those objects, and asks: If a certain object x is one for which P holds, does Q necessarily hold for x as well?

Ulam's initial question exemplifies this point of view. His account does not tell us what he was thinking about, or what he and von Neumann were discussing, before he posed the question, so we unfortunately have no way of knowing why Ulam was wondering about the topological properties that abstract groups may be endowed with. That is, we do not know where the conclusion (the "Q") comes from in Ulam's question. But we do know what happens next. Before proceeding to work further on this problem directly, Ulam (in his narrative, immediately) observes that "the group, of course, has to be of power continuum[5] at most — obviously, a necessary condition." In other words Ulam asserts as obvious (and hence not requiring a detailed proof) the related implication

---

[4] A topological space is called *separable* if it contains a countable dense subset (Munkres, 1975, p. 191). For example, the real plane is separable, because the set of points with rational coordinates comprise a countable, dense subset. In general an abstract group may be given many different topological structures; the question here is, "Can a particular abstract group be endowed with a separable topology?"

[5] A set is said to be *of power continuum* if it has the same cardinality as the set of real numbers.

$$\text{Separabilizable}(G) \rightarrow \text{PowerContinuum}(G) \tag{2}$$

Now, where did (2) come from? We note that Ulam has taken the *conclusion* of his original question (1) and swapped it into the position of the *hypothesis* of a new implication. In other words he has temporarily replaced question (1) with the *converse problem*: to find a property P' such that

$$\text{Separabilizable}(G) \rightarrow P'(G) \tag{3}$$

Moreover, since he knows already something of the theory of separable topological spaces, he is able to immediately find a property P' that holds under the hypothesis that G is separabilizable. Thus, en route to finding a *sufficient* condition for separabilizability, he has identified a *necessary* condition.

In his next paragraph Ulam reports having found "a combinatorial trick showing that it could not be done". But what is the "it" to which Ulam refers? What was "it" that he found could not be done? Does he mean to say that the problem (1) has no solutions — in other words that there are no conditions sufficient to ensure that G is separabilizable? Surely not; on the contrary, all of the classical continuous groups have separable topologies, so one could simply take as a property $P(G)$ that G is isomorphic (as a group) to one of those. So Ulam cannot mean to say that there is no property P that will work as a solution to (1). Rather, it appears from his next few sentences that the "it" to which he refers is a proof of the converse of the implication (2):

$$\text{PowerContinuum}(G) \rightarrow \text{Separabilizable}(G) \tag{4}$$

In other words, Ulam is seeking to determine whether the necessary condition he has already identified is also sufficient, as well. An affirmative answer to this question would constitute not only a solution to the original question (1); it would in addition be a "best possible" solution, insofar as no weaker hypothesis could possible exist. On the basis of the narrative so far, we can identify one of the generative "moves" that produces new questions from old ones: *consideration of the converse*. By applying this move twice in rapid succession, Ulam has replaced the open-ended question (1) with the candidate solution (4).

What Ulam discovered with his "combinatorial trick" was that he could produce an example of a group of power continuum that could not be endowed with a separable topology; in other words, he found a counterexample to establish that

$$\text{PowerContinuum}(G) \not\to \text{Separabilizable}(G) \tag{5}$$

Ulam does not recount the details of his combinatorial trick, but in a sense it does not matter for the present purpose. Here the goal is to describe not how mathematicians *solve* problems, but rather how they *pose* them, and in particular in the way one problem can generate others. So far we have seen Ulam pose a problem that seeks a sufficient condition for a particular conclusion; replace that problem by its converse; find a (to him, obvious) solution to that converse problem; and then ask whether the solution to the converse problem works as a solution in the original. Having resolved this latter question negatively, what could happen next?

Ulam goes on to recount that after he explained his "combinatorial trick" to von Neumann, "Johnny... found an example of a continuum group which is even Abelian (commutative) and yet unable to assume a separable topology". Ulam describes this as a simplification of the proof, but it might be more precise to say that von Neumann has asked (and answered) a new question generated from the last. What von Neumann has done is to strengthen the hypothesis of the failed implication (5) to see whether so doing produces, at last, a sufficient condition to answer Ulam's original question (1). That is, von Neumann considers the additional assumption that G is abelian, and produces an example to show that

$$\text{PowerContinuum}(G) + \text{Abelian}(G) \nrightarrow \text{Separabilizable}(G) \tag{6}$$

It is important to note that von Neumann's proposed (and rejected) condition — that G is an abelian continuum group — does not come from the same generative move as Ulam's earlier attempt. Von Neumann did not come up with this proposed condition by considering a converse and finding a necessary condition; on the contrary, most separable continuous groups are not abelian. Rather, von Neumann is putting into play a second and distinct generative move: *strengthen the hypothesis* of an implication that is known to be false, and see whether it continues to be false.

Here Ulam's narrative stops — but the questions need not. One could continue to ask whether one could further strengthen the hypotheses by adding additional conditions on G that might be sufficient to prove that G is separabilizable. It may be that some such strengthening would eventually succeed. If this were to happen, one natural next move would then be to see if the sufficient condition could be *weakened* at all to find a "better" one. If this subsequent refinement were to prove unsuccessful, one might then (re-)consider the converse, i.e. examine whether the (discovered) sufficient condition is also a necessary one.

Alternatively, one could abandon (at least temporarily) the attempt to find a sufficient condition for (1), and instead try to weaken the conclusion of (6). That is, one could say: All right, it has now been established that not every abelian continuum group can be endowed with a separable topology. Are there any different, weaker topological properties that *could* be proven from the assumptions that *G* is abelian and of at most power continuum? The process of strengthening and weakening hypotheses and conclusions, occasionally interchanging them, can continue in this fashion until one either finds a theorem that one can prove, or a large collection of counterexample that persuades (by weight of evidence) that the algebraic properties of an abstract group do not directly determine its topological properties.

The analysis so far suggests that one of the ways in which mathematicians generate new questions from existing ones is by an iterative series of adjustments made to the hypothesis and conclusion of a "candidate" proposition. These adjustments can be regarded as some of the "moves" in the "game" of mathematics. Among the moves discussed so far are:

(1) *Considering the converse.* In this move, the hypothesis and conclusion of a proposition under consideration are interchanged with one another to generate a new proposition.

(2) *Strengthening a hypothesis*. In this move, the hypothesis of the proposition under consideration is amended, while the conclusion is kept constant. The goal here is to see whether a hypothesis that has so far proved insufficient for a desired conclusion can be bolstered into sufficiency, or alternatively whether it is still possible to produce a counterexample.

In addition to these two, there are five obvious additional moves of a similar type. I now briefly list them; following that I exemplify each of them from the mathematical narratives and discuss them in some more detail. The moves are:

(3) *Weakening a hypothesis.*

(4) *Weakening a conclusion.*

(5) *Strengthening a conclusion.*

(6) *Generalization.*

(7) *Specialization.*

The move *weakening a hypothesis* is in some respects the opposite of (2). This move can be brought into play when a hypothesis has already been established as sufficient; in weakening the hypothesis, one asks whether the result (i.e. the implication $P \rightarrow Q$) can be improved upon (logically strengthened) by taking a weaker hypothesis. As in *strengthening a hypothesis,* in this move the hypothesis of a conditional statement is modified while the conclusion is left untouched (contrast with the related move *generalizing*, below).

For an example of this move, we turn to Davis (2000). Davis recalls (pp. 58-66) how we he was fascinated for years by what is known as Pappus's Theorem. This theorem asserts that given any two lines, and any choice of three points on each of the two lines, if certain segments are drawn joining the points to one another, then the segments will intersect in three points which will always be collinear. Later, Davis learned of a generalization, known as Pascal's Theorem. This latter describes essentially the same phenomenon occurring when the six points are chosen on a conic section. Davis concludes his summary of the two theorems by noting that "Pascal implies Pappus. Why? Simply because the initial two straight lines of Pappus can be considered a degenerate conic section" (p. 63). We could describe Pascal's theorem as being "generated" from Pappus's by saying that the latter's hypothesis of six points on two lines is replaced by the weaker Pascal hypothesis of six points on a conic section — with the same result being proved under the weaker conditions. I do not claim that *historically* Pascal's theorem was produced this way. But Pascal's theorem *could* be generated from Pappus's theorem by applying a kind of standard move (namely, "weaken the hypothesis"). More importantly, the narrative shows that this way of regarding Pascal's theorem has salience for the mathematician, i.e. Davis. Notice, however, that Davis's ability to recognize the Pascal hypothesis as a weakening of the Pappus hypothesis itself relies on the fact that he already knows that "two straight lines... can be considered a degenerate conic section". That is, one's capacity to weaken or strengthen a hypothesis (or to recognize a weakened hypothesis as such) requires that one already knows a network of results related to that hypothesis.

To illustrate how this mechanism might play out in a more elementary context, consider the following theorem of Euclidean geometry: *If ABCD is a rectangle, then its dual*[6] *is a rhombus*. How might one attempt to weaken the hypothesis "*ABCD* is a rectangle"? One could begin by generating a list of (already known to be true) propositions of the form "If *ABCD* is a rectangle, then..." For example, "If *ABCD* is a rectangle, then it is a parallelogram" would be one such proposition on that list. One next takes, in turn, the conclusion of each of those propositions (e.g., "*ABCD* is a parallelogram"), and proposes it as a candidate hypothesis that might (or might not) imply that a dual quadrilateral is a rhombus. In this fashion, one can generate a list of questions like the following:

---

[6] The *dual* of any polygon is formed by joining the midpoints of successive sides so as to create a new polygon.

If ABCD is a parallelogram, will its dual necessarily be a rhombus?

If ABCD contains a right angle, will its dual necessarily be a rhombus?

If ABCD has reflection symmetry about a line that is not a diagonal, will its dual necessarily be a rhombus?

If ABCD has two congruent diagonals, will its dual be a rhombus?

And so forth. In each of these questions we consider a property that is strictly weaker than "*ABCD* is a rectangle" and ask whether that weaker property is strong enough to force the necessary conclusion about the dual of *ABCD*. The questions generated by this process are not trivial: note that of the four examples listed above, the first two must be answered "No" but the second two can be answered "Yes". Moreover, when this process yields a true proposition, a reconsideration of that proposition's hypothesis could lead to the definition of a new mathematical entity, or a new definition for an existing one. That is to say, the predicate "*ABCD has reflection symmetry along a line that is not a diagonal*" (the third example from the above list) describes both isosceles trapezoids and rectangles (and nothing else); recognizing this fact might suggest that an exclusionary definition of "trapezoid" be replaced with a more inclusive definition that subsumes "rectangle" within it as a special case. On the other hand, the predicate "*ABCD has two congruent diagonals*" describes a still larger class of quadrilaterals, including but not limited to the isosceles trapezoids; this class has no conventional name, but the fact that this property has been identified it as a sufficient condition for a desired conclusion might warrant giving it one. Thus the generative move of weakening a hypothesis not only generates new questions; it can lead to the recognition of entirely new classes of objects.

The move *weakening a conclusion* was hinted at above, when I suggested that Ulam and von Neumann might have continued their inquiry by investigating whether it might be possible to prove some weaker topological property about abelian groups of power continuum. This move functions in more or less the same way as does *strengthening a hypothesis*: That is, when a certain property has been shown to be *not necessary* under particular conditions, we might attempt to find some weaker property that is necessary under the same conditions. On the other hand if we have succeeded in proving that a particular property *is* necessary under conditions, we might see if we can go farther by *strengthening the conclusion.*

From a formal point of view, "weakening a hypothesis" and "strengthening a conclusion" can be regarded as the same thing; likewise "strengthening a hypothesis" and "weakening a conclusion" are, at a purely formal level, equivalent. This is because every implication of the form $P \Rightarrow Q$ is logically equivalent to its contrapositive; and "weakening the hypothesis" of one such implication is identical to "strengthening the conclusion" of its contrapositive (and vice versa). Thus these two moves may be regarded as one and the same. This is particularly clear when a proposition is stated in an "unparsed" or "succinct" form — one in which the roles of *P* and *Q* are partially masked. For example, Hardy (1940) discusses at some length Pythagoras's proof of the irrationality of $\sqrt{2}$. Now, the statement "$\sqrt{2}$ is irrational" does not, on its face, have the form of an implication $P \Rightarrow Q$. It can be translated into this form, however, in either of the two following ways:

"If *x* is rational, then $x^2 \neq 2$"

"If $x^2 = 2$, then *x* is not rational"

The relevance of this example for the present discussion is that, almost immediately on completing his narration of the proof, Hardy makes use of one of the moves under discussion:

> We should observe first that Pythagoras's argument is capable of far-reaching extension, and can be applied, with little change of principle, to very wide classes of 'irrationals'. We can prove very similarly (as Theodorus seems to have done) that $\sqrt{3}$, $\sqrt{5}$, $\sqrt{7}$, $\sqrt{11}$, $\sqrt{13}$, $\sqrt{17}$ are irrational...

Now this list of irrational numbers seems not to have been arbitrarily chosen. Only a few pages earlier, Hardy had presented Euclid's proof of the existence of infinitely many primes, and illustrated the definition of "prime" with the sequence "2, 3, 5, 7, 11, 13, 17, 19, 23, 29..." A reader who has been following Hardy's arguments closely is likely to note the correspondence between these two lists of numbers. So what Hardy seems to be saying in the paragraph quoted above is, in essence: *The square root of any prime number is irrational.* Comparing this with the two parsed statements above, one could regard this latter statement either as a weakening of the hypothesis in the first formulation (in that "x^2 is prime" is weaker than "$x^2 = 2$"), or a strengthening of the conclusion in the second formulation (in that "x^2 is not prime" is stronger than "$x^2 \neq 2$". The distinction is immaterial. What matters is that this movement (from "$\sqrt{2}$ is irrational" to "the square root of any prime is irrational") is an illustration of how mathematicians generate new questions (which may be answered affirmatively, as in this case, or negatively) by weakening or strengthening the hypothesis or conclusion of a proposition whose truth status has been already established.

In the examples above, one side of an implication $P \Rightarrow Q$ was modified (strengthened or weakened) while the other side was kept fixed. Such moves produce new implications, or new counterexamples, that are "better" than the original ones. A related, but somewhat different, move consists of *simultaneously* weakening both sides of an implication. That is to say, one might assume less, but also prove less. The resulting implication cannot be regarded as "better" than the original one, but it can be appreciated as a *generalization* of the first result. On the other hand one might proceed in the opposite direction, simultaneously *strengthening* both sides of an implication (proving more, but at the cost of stronger assumptions); such a move can be regarded as a *specialization*.

Generalization and *specialization* form another pair of moves that a mathematician may make use of when maneuvering in a complex problem space. Shiryaev's (2000) retrospective of Kolmogorov's life and work illustrates how Kolmogorov and his contemporaries used these twin moves to generate research questions. In a seminal 1928 paper, Kolmogorov found necessary and sufficient conditions for what he called the "generalized law of large numbers". Two years later, he found (stronger) sufficient conditions for a stronger conclusion, the "strong law of large numbers". This result was itself a strengthening of Borel's original (1909) formulation of the strong law of large numbers, which required a sequence of independent Bernoulli random variables. This law was proved again in 1927 by Khinchin, "who gave a sufficient condition for it that is applicable also to *dependent* variables" (Shiryaev, p. 15). Khinchin's result can be seen as an application of the move "weakening the hypotheses", and Kolmogorov's 1930 result is a further application of the same move. Along similar lines, Kolmogorov produced generalizations of, specializations of, and strengthenings of the "law of the iterated logarithm".

In fact much of Kolmogorov's early work seems to have been motivated by the desire to generalize established results. In a 1930 paper, he set out to produce a generalized, axiomatic theory of integration. Writing in that paper he wrote that his goal was:

> ... to clarify the logical nature of integration. And while, besides a unification of different approaches, a generalization emerged of the concept of integral, the point here apparently is that a generalization of a concept is often useful for comprehending its essence... It is not impossible that all these generalizations may be of interest also for applications, though I see the merits of the more general approach first and foremost in the simplicity and clearness introduced by the new concepts. (Quoted in Shiryaev, pp. 9-10)

Part of the significance of this last passage is that Kolmogorov does more than argue that generalization has utility as a means for generating new questions; additionally he situates that utility in the context of a *set of values* that determines what kinds of questions are important ones. In particular Kolmogorov places a high value here on work that illuminates a common "essence" that may be shared by seemingly disparate mathematical ideas. In the next section, I will turn in more detail to an analysis of the values that mathematicians draw upon in appraising the relative value of a question and/or an answer.

To illustrate how the twin moves of generalization and specialization could work in the context of geometry, consider again the problem of the angle bisectors of a quadrilateral, described earlier. From the proposition "*If ABCD is a parallelogram, then its angle bisectors form a rectangle"* one could try to specialize the proposition by considering a stronger hypothesis and seeing what additional conclusions can be proved — e.g., "*If ABCD is a rectangle, then its angle bisectors form a square*". Or one could try to generalize it, by considering a weaker hypothesis and seeing what can be proven — e.g., one might find that "*If ABCD is a trapezoid, then its angle bisectors form a quadrilateral with two right angles opposite each other*", or the still more general "*If ABCD is any quadrilateral, then its angle bisectors form a quadrilateral in which opposite angles are supplementary*".

Dialectic in mathematics

The above example illustrates the way in which theorems and counterexamples can take shape through a series of modifications to a hypothesis and conclusion. This indeed is one of the principle themes of Lakatos (1976), which characterizes the process of mathematical discovery as a dialectic between that which is assumed (e.g. the definitions and axioms of a theory) and that which is to be proved, and identifies *proofs* and *refutations* as opposing forces that drive the dialectic forward.

The notion of *dialectic* is almost as old as philosophy itself. At its root, dialectic means "the argumentative usage of language" (Popper 1940), in which theories are put forth and critiqued. The dialogues of Plato are dialectics, in which the question-and-answer *dialogic* structure reflects the underlying back-and-forth structure of the argument. Hegel elevated the role of the dialectic, basing on it a dynamic model of history, and indeed of nature itself. Hegel considered developmental processes in which each stage of development is generated as an attempt to resolve internal contradictions (or oppositions) present in the previous stage. Thus in Lakatos, for example, the presence of a proof of a theorem, and the simultaneous proffer of a counterexample to that same theorem, forces a reconsideration of the proof, which can lead to the recognition of certain unexamined assumptions, and eventually to a reformulation of the theorem and its proof to take account of the counterexample. This resolution of the contradiction is likely, however, to contain within it further contradictions, which in turn generate additional refinements to the theorem, and so forth. Processes that are dialectical in this manner are "self-moving" (Larvor 1999).

The so-called "Hegelian dialectic" (named after, but not actually used by Hegel himself) is a particular formulation of this dynamic in which the dialectic unfolds through the three distinct stages of *thesis*, *antithesis,* and *synthesis*. Here, the *thesis* and *antithesis* are two opposing poles of an argument, or two opposing forces in history, and the *synthesis* is that which is generated by the confrontation of thesis with antithesis. This triadic formulation is often understood to refer to distinct *chronological* stages: that is, the thesis is put forward first, the antithesis is put forward second as a response to the thesis, and the synthesis comes last as a resolution of the contradiction. In the more general notion of dialectic, the thesis and synthesis need not be temporally separated in this fashion: rather, any time two opposing forces or arguments are in play, they can be said to constitute a *dialectical pair* that may have generative force as their contradictions drive the process forward.

Curiously, Hegel himself regarded mathematics as un-dialectical, and hence "inert and lifeless" (Hegel, 1807/1977, p. 26). According to Larvor (1999), Hegel understood "mathematics" as a synonym for the formal product of mathematician's work. In this form, mathematics proceeds linearly from definitions and first principles to theorems. Hegel's critique was that, despite the fact that each stage of a mathematical argument may be *deduced* from the previous stages, there is nevertheless nothing in the previous stages that *compels* what the next stages ought to be. At any point in a proof, there are many possible next lines; mathematicians make their selection of what should follow based on considerations that are external to mathematics itself; therefore "according to Hegel, a proof tells us more about the ingenuity of the mathematician than it does about the meaning of the theorem" (Larvor 1999, p. 1).

Lakatos, in revealing the heuristic mathematics that precedes that formal product, reveals the fundamentally dialectic (and Hegelian) nature of mathematical work. In his reconstruction of the proof of Euler's Theorem, Lakatos shows how a series of proposed proofs and refutations led to a reconsideration of the basic assumptions of the theory of polyhedra, ultimately leading to an entirely new conceptualization of a polyhedron not as a solid but as a closed surface, and to a collection of modifications of Euler's Theorem that both reduce the scope of the claim and extend it to other cases. Thus the proofs and refutations mediate between a conclusion and a set of assumptions that are both continually in flux. Lakatos writes, "Mathematics, this product of human activity, 'alienates itself' from the human activity which has been producing it. It becomes a living, growing organism, that *acquires a certain autonomy* from the activity which has produced it; it develops its own autonomous laws of growth, its own dialectic.... Any mathematician, if he has talent, spark, genius, communicates with, feels the sweep of, and obeys this dialectic of ideas" (p. 146).

If Lakatos is right in claiming that "any mathematician... feels the sweep of" the dialectical driving force that pushes mathematics forward, then it is reasonable to suppose that any attempt to map out the elements of the mathematical sensibility is likely to find expression in the form of dialectical pairs. Note the deliberate use of the plural, here: As

Kiss (2006, p. 309) observes, Lakatos made no claims that the dialectic of proof and refutation was the *sole* motive force in mathematics; on the contrary, Kiss argues, Lakatos recognized translation and interpretation among other motive forces. Nor did Lakatos insist on fitting all mathematical dialectics into the rigid triadic thesis-antithesis-synthesis pattern:

> [Lakatos's] approach to dialectic does not mean to accept a small set of schemes given by Hegel. Dialectic is meant much more in a classical sense, where the dialogue and the debate are the basic sources of knowledge. Hegel is a good source for Lakatos, because he emphasizes the positive role of contradiction in the development of knowledge. (Kiss, p. 309)

One of my chief goals in the remainder of this chapter will be to describe, on the basis of the mathematical narratives, a number of mathematical dispositions that, I argue, are organized in dialectical pairs in the Hegelian sense. Each of these pairs points to two mathematical values that coexist and push against one another in a kind of dynamic tension. It is important to reiterate, however, that I do not claim that mathematicians are *consciously aware* of these dialectical pairs, or that they deliberately "activate" certain generative moves when seeking ideas for what to do next. Rather, my goal is something akin to Lakatos's "rational reconstruction" of a mathematical argument. Lakatos (1978, p. 103) distinguishes between *heuristic* and *methodology*: the former refers to the rules of discovery, while the latter is a *post hoc* synthetic description of the apparent mechanism by which results were discovered. It is the latter at which I aim. In claiming that these pairs are among the categories of perception and appreciation of mathematicians, I am trying to show that mathematicians themselves perceive these categories as useful descriptors of their work.

In this context it is worth revisiting one of the generative moves described above; namely, *Considering the converse problem.* From one point of view this move seems not to fit into a dialectical framework: what is its partner? But there is another way of regarding this move: namely, one can interpret *Considering the converse* as its *own* dialectical partner. More explicitly, this move is "self-dual" in the sense that applying it twice in succession returns one to the problem originally under consideration. Significantly, though, such a double reversal does not leave one at a standstill; on the contrary, considering the converse and then reconsidering the preconverse can be generative in that one ends up with more than one began with. Ulam's narrative is an example of this: he begins with a problem ("Find a condition P that implies separabilizability"), considers the converse problem ("Find a condition P' that *is implied by* separabilizability"), solves that problem ("separabilizability implies 'power continuum'"), and then considers the converse problem of that solution ("Does 'power continuum' imply separabilizability?"). Because of its generative quality, I regard *Considering the converse* and *Reconsidering the preconverse* as a dialectical pair of generative moves.

In summary, to this point I have identified in the mathematical narratives five generative moves that mathematicians use in determining what could be done next. The five moves are organized in dialectical pairs (one of the moves is its own dialectical partner), in that the two members of each pair point in (so to speak) opposite directions. These generative moves, shown in Fig. 2.2, stand as a first approximation at mapping the categories of perception of mathematical practice in the context of problem-posing: they provide a language for describing what a mathematician sees as visible possibilities. In the subsequent section I turn to the narratives again to build an inventory of categories of appreciation. Together these two inventories will comprise a description of mathematicians' practical rationality, or, as I have been referring to it, the mathematical sensibility, as it applies to posing and valuing mathematical problems and results.

(strengthen/weaken) hypothesis	↔	(strengthen/weaken) conclusion
generalize	↔	specialize
consider converse	↔	consider converse

Figure 2.2. Categories of perception (generative moves).

Categories of mathematical appreciation

Overview

It is important to note that while the generative moves described in the previous section can be instrumental for generating many new questions for a mathematician to investigate, not all of those questions will be deemed to be of equal value. Indeed a mathematician may be aware of many problems that could potentially be investigated, but only choose to work on a small subset of those problems. We may describe the problem space as *filtered*, in the sense that some problems are given a higher value than others. To some extent this filtration is conditioned on pragmatism: at any given moment, some problems seem to be well within reach, while others seem more distant. The choice of what problem to work on may also be guided by a strategic sense that solving certain "nearby" problems may prove instrumental in a future attack on other "more distant" problems.

But quite apart from pragmatic concerns about which problems are tractable and which are not, there is an additional set of criteria that mathematicians bring to bear in determining which problems are *important* and which are not. We can see this most clearly in the case of a mathematician reading a research report, or learning that a colleague has successfully solved a problem: some such accomplishments are esteemed very highly, others less so. In this section, then, I turn again to the mathematical narratives to understand how mathematicians decide what is important, and what is not. How do mathematicians choose which questions to pursue and which to ignore? In other words, what are the values that a mathematician brings to bear in judging a question (whether one of his own posing, or an extant open problem) as worthy of pursuit? And are there other *categories of appreciation*, besides "importance", that are shared among practicing mathematicians?

In asking those questions one must be prepared to find that not all mathematicians care about the same things, and to some extent "what is important" may vary across mathematical sub-specialties. A seminal result in logic and foundations of mathematics (e.g., the independence of the axiom of choice from the Zermelo-Frankel axioms of set theory) may be of only mild interest (or none at all) to a topologist. Despite this, there may well be categories of appreciation that cut across content areas –values that are more or less *universal* among mathematicians. Andrew Wiles's proof of Fermat's Last Theorem, for example, was widely hailed as important, even among mathematicians far removed from the technical details of elliptic curves and modular forms (Singh 1998). What was it about this accomplishment that made it seem so significant?

In his celebrated *A Mathematician's Apology*, Hardy (1940/1992) identified two such categories of appreciation, namely *beauty* and *seriousness.* Regarding *beauty*, he wrote:

> The mathematician's patterns, like the painter's or the poet's, must be *beautiful*; the ideas, like the colours or the words, must fit together in a harmonious way. Beauty is the first test: there is no permanent place in the world for ugly mathematics.... It would be difficult now to find an educated man quite insensitive to the aesthetic appeal of mathematics It may be very hard to *define* mathematical beauty, but that is just as true of beauty of any kind — we may not know quite what we mean by a beautiful poem, but that does not prevent us from recognizing one when we read it.

Hardy freely admits to being unable to define mathematical beauty, and instead shifts towards a deeper analysis of the related notion of "mathematical seriousness". He identifies two criteria that contribute to a judgment of the seriousness of a piece of mathematical work: *generality* and *depth*, but concedes that "neither quality is easy to define at all precisely". Moreover he hastens to point out that "some measure of generality must be present in any high-class theorem, but *too much* tends inevitably to insipidity" (emphasis in original). In regards to *depth*, Hardy likewise struggles to make explicit what he means:

> This is still more difficult to define. It has *something* to do with *difficulty*; the 'deeper' ideas are usually the harder to grasp: but it is not at all the same.... It seems that mathematical ideas are arranged somehow in strata, the ideas in each stratum being linked by a complex of relations both among themselves and with those above and below. The lower the stratum, the deeper (and in general the more difficult) the idea.... It may happen that [an idea in one stratum] can be comprehended completely, that we can recognize and prove, for example, some property of the integers, without any knowledge of the contents of lower strata.... But there are also many theorems about integers which we cannot appreciate properly, and still less prove, without digging deeper and considering what happens below.... This notion of 'depth' is an elusive one even for a mathematician who can recognize it, and I can hardly suppose that I could say anything more about it here which would be of much help to other readers.

Sinclair (2002), in her doctoral thesis on the mathematical aesthetic, surveys a number of prior attempts to go beyond Hardy's theory of the aesthetic, but finds little in common among them. Of particular note is the work of Le Lionnais (1948 / 1986, cited in Sinclair, pp 26-28). The latter, according to Sinclair, identifies two distinct categories of mathematical conceptions of beauty, *classicism* and *romanticism*. The former finds beauty in "equilibrium, harmony and order"; the latter, in "lack of balance, form obliteration and pathology". The classicist aesthetic values highly structured patterns, such as magic squares or Pascal's triangle, and methods of proof that contain a regular structure, such as proofs by induction. The romanticist aesthetic, on the contrary, highly regards "pathological" phenomena such as asymptotes and imaginary numbers, and "romantic methods, such as *reductio ad absurdum* proofs, [which] are characterised by indirectness." Although neither Le Lionnais nor Sinclair use the language of dialectic, we may see in these opposing categories an example of the kind of dialectical pair discussed above. Le Lionnais suggests that mathematicians are either romantics or classicalists, and that being one or the other is a matter of personal preference, and so it may be; on the other hand we might say that both of these approaches are characteristic of the *practice* of mathematics, and that in any given instance, the individual practitioner may base a judgment on an appeal to one or the other of these factors. Much the same approach is taken by Larvor (2001):

> First, the dialectical philosopher of mathematics adopts what I have just called the 'inside-phenomenological stance'. Do not let the word 'phenomenological' mislead you. This is *not* a study of what it feels like to do mathematics. The phenomenologist takes up a point of view and studies its logical constitution as it were 'from the inside'. But we are not concerned with the individual mathematician. We are interested in the 'point of view' belonging to mathematics itself. This way of speaking is of course analogical. Mathematics has no subjectivity in the proper sense.

> It feels neither joy nor pain. Nevertheless, the analogy is not mysterious. We can say that the theory of projectile motion is 'blind' to the ethical difference between a distress flare and an assassin's bullet (though theor*ists* are not). There is no mystery in the remark that analysis was 'conflicted' over the rival versions of the early calculus (though individual mathematicians clearly supported one over the other).... The point of the inside-phenomenological stance is to insist that changes in the body of mathematics normally take place for mathematical reasons. (pp. 214-215).

Sinclair's work is also significant in that she pays particular attention to the *generative* and *motivational* role of the mathematical aesthetic. By this she means that an aesthetic sensibility not only serves for making after-the-fact evaluations of completed mathematical work; additionally it can both guide work that is in progress, and play a crucial function in the posing of new mathematical problems. In this respect, Sinclair's "aesthetics" corresponds closely to the notion of practical rationality as I have described it above. The main difference between the theoretical constructs rests in the role they ascribe to beauty. A theory of aesthetics takes beauty as the object of inquiry, and attempts to elucidate the component qualities that combine to produce a judgment that something is or is not beautiful. A theory of practical rationality, on the other hand, seeks to identify the full set of lenses with which the practitioner looks upon the world: while beauty may be *one* of those lenses, it is not assumed to be an all-encompassing or even dominant one.

Corfield (2001) takes on the question of how mathematicians describe the relative value of one mathematical concept (the notion of *groupoid*). He proposes five broad categories of mathematical values:

1. when a development allows new calculations to be performed in an existing problem domain, possibly leading to the solution of old conjectures;

2. when a development forges a connection between already existing domains, allowing the transfer of results and techniques between them;
3. when a development provides a new way of organising results within existing domains, leading perhaps to a clarification or even a redrafting of domain boundaries;
4. when a development opens up the prospect of new, conceptually motivated domains;
5. when a development reasonably directly leads to successful applications outside of mathematics. (p. 508-509).

Corfield stresses that these categories are neither exclusive nor comprehensive. In fact each of these five categories corresponds roughly to one of the categories of appreciation I outline below. Here I simply note in passing that Corfield's list of categories seems to possess, at least in part, an unrecognized dialectical structure: the first two categories are in some sense directed opposite one another, as are the next two.

In the pages that follow I return to the corpus of mathematical narratives to find evidence of multiple categories of appreciation that seem to be part of the mathematical sensibility. As has been suggested above, these categories appear to be organized in dialectical pairs. That is to say, instead of identifying a particular quality Q (beauty, depth, etc.) and appraising a mathematical work by saying it either "has Q" or "lacks Q", in some cases it seems more appropriate to identify *pairs* of qualities (Q_1, Q_2) that are in a sense opposed to one another, and show how the mathematical narratives characterize mathematical work as being "more Q_1" or "more Q_2". As Popper (1940) points out, the opposing pairs in a dialectic are in a relationship that is more accurately described by the metaphor of *polarity* than by the language of *contradiction*. It is not contradictory to assert (as I do below) that, for example, "Mathematicians appreciate simplicity" and at the same time "Mathematicians appreciate complexity". Rather, the fact that both simplicity and complexity are simultaneously held in regard by mathematicians creates an inner tension that provides a motive force generating new mathematics.

In particular I identify the following categories of appreciation (the paired categories will be treated together):

(1) Utility / Abstraction

(2) Surprise / Confirmation

(3) Theory-building / Problem-solving

(4) Simplicity / Complexity

(5) Formalism / Platonism

I will devote most of my attention to the first three of these pairs, which appear to be the most prominently on display in the narratives; these three will also be the ones that I use as an organizing framework in subsequent chapters of this dissertation. The latter two pairs will be discussed in less detail.

Utility / Abstraction

It is generally accepted among the public at large that mathematics is a useful discipline. Recent reports, such as the National Academies' *Rising above the gathering storm* (Committee on Prospering in the Global Economy 2007), have argued forcefully that the economic and technological challenges of the future demand a mathematically sophisticated populace. The very first paragraphs of the NCTM's *Principles and Standards* describe the increasing role of mathematics in everyday life:

> We live in a mathematical world. Whenever we decide on a purchase, choose an insurance or health plan, or use a spreadsheet, we rely

> on mathematical understanding. The World Wide Web, CD-ROMs, and other media disseminate vast quantities of quantitative information. The level of mathematical thinking and problem solving needed in the workplace has increased dramatically.
>
> In such a world, those who understand and can do mathematics will have opportunities that others do not. Mathematical competence opens doors to productive futures. A lack of mathematical competence closes those doors.

Mathematics that is not directly relevant to "real life" will be referred to here as "abstract" mathematics. That is, "abstract" mathematics is understood to be disconnected from any kind of experiential context or warrant. The distinction between "utilitarian" and "abstract" mathematics is not a sharp one; mathematics that is quite abstract (e.g. the classification of quadratic equations in two variables) may become quite useful when applied to a real-world context (e.g. the study of gravitational orbits in a two-body system). "Utility" and "abstractness" are not presented here as mutually exclusive labels, nor as two ends of a continuum. Rather, they are here used to refer to *alternative categories of appreciation*, categories which are opposed to one another and thus comprise a dialectical pair. That is to say, it is not an individual "piece" of mathematics that is abstract or utilitarian; rather, "abstract" and "utilitarian" are distinct ways of appreciating mathematical work and of generating new mathematics. The fact that these two opposing categories may both point to the same mathematical "thing" is an illustration of the workings of the dialectical triad: the thesis and antithesis find their resolution in the production of a synthesis.

Before proceeding it is important to clarify that by "useful" here is meant *utility in a domain other than pure mathematics.* Of course even the most abstract and rarified mathematical ideas and results may be instrumental in solving some other mathematical problem; that is a different kind of usefulness. Here we are concerned with what Hardy calls ‘practical utility’:

> It is undeniable that a good deal of elementary mathematics — and I use the word ‘elementary’ in the sense in which professional mathematicians use it, in which it includes, for example, a fair working knowledge of the differential and integral calculus — has considerable practical utility. These parts of mathematics are, on the whole, rather dull; they are just the parts which have least aesthetic value. The ‘real’ mathematics of the ‘real’ mathematicians, the mathematics of Fermat and Euler and Gauss and Abel and Riemann, is almost wholly ‘useless’ (and this is as true of ‘applied’ as of ‘pure’ mathematics)...
>
> But here I must deal with a misconception. It is sometimes suggested that pure mathematicians glory in the uselessness of their work, and make it a boast that it has no practical applications. The imputation is usually based on an incautious saying attributed to Gauss, to the effect that, if mathematics is the queen of the sciences, then the theory of numbers is, because of its supreme uselessness, the queen of mathematics — I have never been able to find an exact quotation. I am sure that Gauss’s saying (if indeed it be his) has been rather crudely misinterpreted. If the theory of numbers could be employed for any practical and obviously honourable purpose... then surely neither Gauss nor any other mathematician would have been so foolish as to decry or regret such applications. (pp. 119-121)

In light of the widespread agreement on the usefulness of mathematics, it is notable that the mathematical narratives are frequently critical of the position that the primary function of education is to teach useful mathematics. Davis (2000) for example, writes derisively about the commonplace perception that only “useful” mathematics belongs in school:

> In those days, it was considered part of a liberal education to know the quadratic formula. I’m not sure how the formula fares now. Mathematical educators are apt to ask: Does your average dentist have occasion in his practice to use the quadratic formula? Does your average insurance salesman? Will the solution to the quadratic equation bring peace on earth? You’re not sure? Then to hell with it. (p. 17)

The importance of this critique rests not in what it has to say about education, but in what it tells us implicitly about the values Davis attaches to *his own work*. That is, his critique of "mathematical educators" serves to establish a contrast between "us" (mathematicians) and "them" (mathematical educators). In caricaturing the other as preoccupied with utility to the exclusion of all other values, Davis indirectly communicates the extent to which his own practice can attach value to mathematics that has no overt utilitarian application.

The above examples may give the impression that mathematicians look with unanimous disinterest on the utilitarian character of mathematics. But this would be a misrepresentation. Indeed, the mathematical narratives provide ample evidence that mathematicians also place a high value on mathematics that finds its origins and its application in real-world contexts. The passage from Wiener (1956), quoted already at the beginning of this chapter, is but one example:

> The moods of the waters of the river were always delightful to watch. To me, as a mathematician and a physicist, they had another meaning as well. How could one bring to a mathematical regularity the study of the mass of ever shifting ripples and waves, for was not the highest destiny of mathematics the discovery of order among disorder? At one time the waves ran high, flecked with patches of foam, while at another they were barely noticeable ripples. Sometimes the lengths of the waves were to be measured in inches, and again they might be many yards long. What descriptive language could I use that would portray these clearly visible facts without involving me in the inextricable complexity of a complete description of the water surface?

Nor is Wiener unique in this regard. Ulam (1976) finds similar inspiration from pondering the meaning of the word "billowing":

> 'Billowing' is a motion of smoke, for example, in which puffs are emitted from puffs. It is almost as common in nature as wave motion. Such a word may give rise to a whole theory of transformations and

> hydrodynamics. I once tried to write an essay on the mathematics of three-dimensional space that would imitate it. (p. 105)

Wiener and Ulam begin from different places — Wiener gazes at the Charles River, Ulam ponders the meaning of words — but both are inspired mathematically by the challenge of modeling complex real-world phenomena with mathematics. Thus Wiener's use of the Lebesgue integral[7] to study idealized Brownian motion[8] — in which he showed that such motions corresponds to the class of continuous but nowhere-differentiable curves (Wiener, p. 39) — can be appreciated both as utilitarian *as well as* abstract. Much the same could be said of Ulam's invention of Monte Carlo methods[9] (Ulam, pp. 196-201) for solving physical problems.

Another example comes from the work of Kolmogorov. Arnol'd (2000, p. 89-90) recounts how he had constructed for himself a hypothetical narrative of how Kolmogorov came to his groundbreaking research on invariant tori[10]. Arnol'd supposed that Kolmogorov's interest on invariant tori had emerged from his studies of turbulence. When, in 1984, he had the opportunity to ask Kolmogorov whether his supposition was correct, the latter replied that in fact his research had been motivated by problems in celestial mechanics. It is interesting that Arnol'd was wrong about the particulars of Kolmogorov's motivation, but right in ascribing it to a problem that emerges from

---

[7] The *Lebesgue integral* is an extension of the familiar notion of the Riemann integral: it allows for the integration of a broader class of functions.

[8] *Brownian motion* is the random movement of particles suspended in a liquid or gas. Although the motion of an individual particle undergoing Brownian motion is unpredictable, certain statistical measures of an ensemble of such particle may be calculated.

[9] *Monte Carlo methods* refers to a class of computational methods in which a large number of random or semi-random processes are simulated numerically.

[10] *Invariant tori* arise in the solution of many integrable systems of differential equations in classical mechanics.

attempting to mathematically model complex empirical phenomena. This example shows that the category of "utility" is shared by both of these mathematicians.

Surprise / Confirmation

"Utility" and "abstraction" are not the only categories with which mathematicians appraise pieces of mathematical work. Another dialectical pair of categories of appreciation that emerges from the mathematical narratives is "surprise" / "confirmation".

A mathematical result is surprising if it establishes something that was not expected to be true. Often surprises come in the form of counterexamples. A classic example is Weierstrass's 1871 discovery of a function that is continuous but nowhere-differentiable. Weierstrass's function was not the first example of a such a function, but it was the first to achieve wide recognition, and it called into question work done by prior mathematicians who had naively assumed that a continuous function would always be differentiable except, perhaps, at isolated points. Peano's 1890 example of a continuous, space-filling curve similarly came as a surprise to mathematicians of the time.

The opposing category of appreciation for "surprise" is "confirmation". A mathematical result may be appreciated for confirming something that was widely believed to be true, but not yet proven. Thomas Hales's 1998 proof of the Kepler conjecture — that the face-centered cubic lattice has maximal density among all sphere packings — falls into this category: the finding was seen as significant in no small part because it put to rest a long-standing conjecture. A proof of Goldbach's conjecture, if one were to be discovered, would similarly be hailed as an important confirmatory result.

(And, of course, if a counterexample to Goldbach's conjecture were found, it would be hailed as a surprising counterexample.)

It should be noted that being "surprising" is a time-dependent quality. At the time Weierstrass and Peano presented their counterexamples, they were quite unexpected; the same could not be said of successive counterexamples that showed the same thing. If these other counterexamples are also valued, it must be with other categories of appraisal (such as "simplicity"; see below).

Consider, for example, the history of Fermat's Last Theorem (Singh 1998). At the time Fermat made his infamous margin note, the result he claimed to have proven may have been quite surprising. Why, after all, would anyone have expected that the equation $a^n + b^n = c^n$ have no whole-number solutions for $n > 2$? On its face it seems quite an unlikely claim. Of course, by the time Andrew Wiles came to work on the problem, it was generally taken for granted that it was true. Wiles' result was not monumental for proving something surprising, but rather for confirming what generations of mathematicians had come to believe was true but were unable to prove themselves.

Shiryaev's (2000) biography of Kolmogorov notes that the latter made his initial mark by finding counterexamples to refute claims made by established mathematicians:

> During his first year [as a university student] (1920-1921) he attended N.N. Luzin's lectures on the theory of analytic functions.... At one of the lectures devoted to a proof of Cauchy's theorem Luzin used the following assertion: "Let a square be partitioned into finitely many squares. Then for any constant *C* there is a number *C′* such that for every curve of length at most *C* the sum of the perimeters of the squares touching the curve does not exceed *C′*." Luzin posed this as a problem for his listeners to prove. "I was able to show that this assertion is actually erroneous," recalled [Kolmogorov]. "[Luzin] at once saw the idea of the example disproving the supposition. It was decided that I should report the counterexample at the student mathematical circle." (pp. 6-7)

Only two years later (1922) Kolmogorov constructed another counterexample, one that became “his most famous result in the area of trigonometric series: the construction of an example of a Fourier-Lebesgue series diverging almost everywhere” (Shiryaev, p. 8), a counterexample which he improved on in 1926 when he presented an example of an integrable function whose Fourier series diverges *everywhere*.[11] Shiryaev remarks that “both these examples were completely unexpected for specialists and made an enormous impression”. So deep was the impression, in fact, that Arnol’d (2000, p. 90) reports a conversation with the French mathematician M.R. Fréchet in 1965 (more than four decades later) in which the latter said, “... Kolmogorov, isn’t he the young fellow who constructed an integrable function with almost everywhere divergent Fourier series?” As Arnol’d remarks, “All the subsequent achievements of Andrei Nikolaevich [Kolmogorov] — in probability theory, topology, functional analysis, the theory of turbulence, the theory of dynamical systems — were of less value in the eyes of Fréchet.”

Parker’s (2005) biography of R.L. Moore narrates mathematical episodes in the life not only of its principal subject, but also of many of Moore’s students. In describing the work of two of them, Edwin Moise and R. H. Bing, Parker discusses how Moise’s dissertation analyzed the properties of the “pseudo-arc”. According to Parker, “Moise was convinced at the time that the pseudo-arc wasn’t homogeneous. Bing became very interested, and... eventually [produced a] characterization of the pseudo-arc as a homogeneous indecomposable, chainable continuum. This result contradicts most

[11] I note in passing that this "improvement" is itself an exemplification of generative moves described above.

people's intuition about the pseudo-arc, including, at the time, Moise's, and directly contradicted a published, but erroneous, 'proof' to the contrary."

Moore appears to have made a particular effort to create opportunities in his classroom for students to experience this kind of surprise by finding counterexamples for theorems that were suspected of being true. Parker cites F. Burton Jones, one of Moore's former students, to this effect:

> "Quite frequently when a flaw would appear in a proof everyone would spend some time (possibly in class) trying to get an example to show that it couldn't be 'patched up', i.e., a counterexample to the argument (even though the theorem might be correct). This kind of experience is seldom encountered in courses or in any place outside of one's own research work. Yet this kind of activity is vitally necessary for the research worker." (Parker, p. 152)

This practice can be regarded as an attempt to cultivate through pedagogy one of the mathematical dispositions. As such it touches on and presages the substance of the subsequent chapters of this dissertation, which ask whether the high school Geometry class can serve as a site for cultivating in students the mathematical sensibility.

Theory-building / Problem-solving

In this section I discuss two qualities that form another dialectical pair of categories of appreciation: *theory-building* and *problem-solving*. By *theory-building* I refer to an orientation towards mathematics that places a high value on the *organization* of a body of mathematical work. Theory-building comes to the fore when, for example, a mathematician takes a collection of already established results and arranges (or rearranges) them into a theory, perhaps proposing new and more economical definitions and postulates. From one point of view theory-building focuses less on proving "new" mathematics than on illuminating the relationships among existing pieces of mathematical knowledge. From another point of view, of course, establishing such a relationship can be regarded as a "new" and important result in its own right. The following examples will help to clarify this point.

John Parker's (2005) biography of R.L. Moore describes the young Moore's early work on the foundations of geometry. In his studies of Hilbert's then-recent *Grundlagen der Geometrie* (1899), Moore proved that one of Hilbert's axioms (the fourth 'betweenness axiom') was derivable from the remaining axioms, and thus could be regarded as a theorem rather than an axiom. According to Parker, Moore was so excited at this finding that "he dashed over the campus late in the evening", and, seeing a light shining in the window of his mentor G.B. Halsted, went at once to show the latter his proof (Parker, p. 35).

What I wish to point out about this episode is that from one point of view it could be said that Moore's discovery did not add any "new knowledge" to the field of geometry. The property that Moore proved was already an established component of the theory. But to view the episode in that way is to miss the point entirely. A mathematical theory is more than a collection of properties; it is the organization of those properties into a web of contingent implications that makes a theory. What Moore had done — and indeed what Hilbert had done before him — was to contribute to an ongoing project (spearheaded by Hilbert) of fundamentally restructuring an existing theory by placing it on a new set of foundations. Indeed for the next decade much of Moore's mathematical output concerned the independence and interdependence of various proposed axiomatizations of geometry.

In fact, as Parker goes on to explain, the fact that the fourth betweenness axiom followed from the others was already known, and had been published the previous January by his namesake, Professor E.H. Moore of the University of Chicago. However, whereas the elder Moore's proof was lengthy and awkward (Halsted referred to it as an "obscure and bungling adumbration"), the young Moore's result was elegant and simple. Here again we see that the significance of a mathematical discovery may lie somewhere other than in its propositional content.

This example is particularly important in the light of the analysis I will present in the subsequent chapters of this dissertation. Its importance resides in the fact that Moore was working on questions concerning the organization of the theory of Euclidean geometry — that is, the *same* theory that (in a simplified form) is the content of the high school Geometry curriculum.

Parker reports that Moore made his most significant contributions in the emerging theory of point-set topology, or *analysis situs* as it was then known. Here again Moore's focus was primarily on the organization of the theory itself. Moore's seminal papers in the field explored the consequences of various proposed sets of axioms for topology and verified the independence of the axioms in those sets. Moore's goal appears in part to have been the development of a theory that contained just enough axioms, and no more, to produce all of point-set topology. But of course "point-set topology" was itself a construct undergoing constant revision and reinvention at the time. Moore's work, and that of his contemporaries, was rooted in an approach to axioms and definitions that regarded them always as *provisional* and subject to revision.

Moore was, of course, not the only mathematician who placed a high value on the structural aspects of mathematical work. Parikh's (1991) biography of Oscar Zariski deals extensively with Zariski's efforts to reorganize the projective geometry of the so-called "Italian school" along more rigorous, algebraic lines. So radical was Zariski's restructuring of the field that a talk he gave at the University of Moscow in 1935 evoked open hostility from an earlier generation of mathematicians: "Finikov and some other geometers of the old school who'd not been very much exposed to algebraic geometry rose up to complain: 'Is this algebraic geometry? What is the matter? We've never seen such geometry!'... Pontrjagin, Sobolev, and the other younger men tried to defend me, but it was no use" (p. 79). In this work, Zariski was pushing the frontiers of algebraic geometry, discovering new theorems and defining new concepts, while simultaneously creating a new structure for the existing field: creating new definitions for existing concepts, discovering new proofs of existing theorems, and — perhaps most significantly — generalizing the theory to encompass fields of arbitrary prime characteristic. Thus Parikh reports,

> By combining the algebraic notion of a regular local ring with the geometric notion of a simple point, he was eventually able to define both normal varieties and simple point in the case of characteristic *p*. (p. 90)

The idea that producing a *definition* may be rightfully regarded as a significant mathematical accomplishment is a good example of the theory-building point of view, as well as the value of generalization. It stands in marked contrast with the position of Wu (1999), quoted above, which regards precise definition as the *starting* point of mathematics, and denigrates all work leading up to that moment as a kind of inductive "pre-mathematics" akin to the scientist's "data collecting phase".

The dialectical partner of *theory-building* is *problem-solving*. Where *theory-building* places a premium on the development (and re-development) of concepts and the organization of results about those concepts, *problem-solving* places high value on the development of a large and flexible toolkit of heuristics. In recognizing problem-solving as the counterpart to theory-building, I am following the mathematician and Fields medalist W.T. Gowers, who wrote about the two frames in his essay on the "two cultures of mathematics" (Gowers 2000). In that essay, Gowers wrote:

> The "two cultures" I wish to discuss will be familiar to all professional mathematicians. Loosely speaking, I mean the distinction between mathematicians who regard their central aim as being to solve problems, and those who are more concerned with building and understanding theories. This difference has been remarked on by many people, and I do not claim any credit for noticing it. As with most categorizations, it involves a certain oversimplification, but not so much as to make it useless..... When I say that mathematicians can be classified into theory-builders and problem-solvers, I am talking about their *priorities*, rather than making the ridiculous claim that they are exclusively devoted to only one sort of mathematical activity.....
>
> It is equally obvious that different branches of mathematics require different aptitudes. In some, such as algebraic number theory, or the cluster of subjects now known simply as Geometry, it seems (to an outsider at least — I have no authority for what I am saying) to be important for many reasons to build up a considerable expertise and knowledge of the work of other mathematicians are doing, as progress is often the result of clever combinations of a wide range of existing results. Moreover, if one selects a problem, works on it in isolation for a few years and finally solves it, there is a danger, unless the problem is very famous, that it will no longer be regarded as all that significant.
>
> At the other end of the spectrum is, for example, graph theory, where the basic object, a graph, can be immediately comprehended. One will not get anywhere in graph theory by sitting in an armchair and trying to understand graphs better. Neither is it particularly necessary to read much of the literature before tackling a problem: it is of course helpful to be aware of some of the most important techniques, but the interesting problems tend to be open precisely because the established techniques cannot easily be applied. (pp. 1-3)

Gowers claims that there is a well-entrenched bias among many mathematicians for valuing theory-building more highly than problem-solving, and argues at length that this bias is unwarranted. He argues that, contrary to appearances, combinatorics (his exemplar of the kind of mathematics that is attractive to those with a preference for problem-solving over theory-building) contains just as rich a structure as (for example) algebraic geometry; the difference, he claims, is that "the important ideas of combinatorics do not usually appear in the form of precisely stated theorems, but more often as general principles of wide applicability". By "principles" Gowers means (as his examples show) *problem-solving heuristics*, of which he says "they play the organizing role in combinatorics that deep theorems of great generality play in more theoretical subjects."

The problem-solving disposition — which might also be referred to as an appreciation of "method" — finds little expression in the corpus of narratives analyzed for this study. This may be a matter of selection bias in the construction of the corpus; it may be that problem-solving, precisely because of its technical and heuristic nature, does not lend itself to story-telling as well as theory-building does. Problem-solving finds its greatest expression in works such as Polya's (1957) *How to solve it* — works, it will be recalled, that were specifically excluded from the corpus because of their non-narrative form.

Nevertheless it would be incorrect to claim that problem-solving is *entirely* absent from the corpus of narratives. One prominent example of it is in Ulam's (1976) account of the development of Monte Carlo methods, already referred to above. These methods do not belong to a "theory" of anything; rather they constitute a toolkit for solving a variety of problems in diverse fields.

Other categories of appreciation

Utility / *abstraction*, *surprise* / *confirmation*, and *theory-building / problem-solving* are the main categories of appreciation in evidence in the mathematical narratives analyzed. They should not, however, be taken as an exhaustive list of the dispositions that constitute a mathematical sensibility. In addition to these, there are others that are present to a lesser extent in the narratives, and I here discuss a few of them.

In undertaking my analysis of the narratives I anticipated that I would find evidence for *formalism* as a category of appreciation. By formalism I mean an orientation towards the syntactical, rather than semantic, aspects of mathematics. Formalism is perhaps best exemplified by Russell & Whitehead's opus *Principia Mathematica*, and by Hilbert's famous declaration that "one must be able to say at all times — instead of points, lines, and planes — tables, chairs, and beer mugs." Formalism is more than an extreme variant of abstraction: for the formalist, mathematical objects and expressions have *no intrinsic meaning* other than that which can be encoded in the formal language of mathematical symbolism.

The counterpart to formalism will be referred to as *Platonism*. From the Platonist perspective, one regards mathematical objects as having some reality all their own; the purpose of formal definitions is to try to *capture* and *reflect* that reality. To a Platonist, for example, the set of natural numbers is a real thing, with properties of its own, and the set if Peano axioms is merely an attempt to formally reproduce those properties. From such a point of view, Gödel's incompleteness theorems serve as a caution that no formal system can fully capture the reality of the mathematical universe. From the formalist point of view, on the other hand, these theorems teach us precisely the opposite: that there *is* no single "true" set of natural numbers, but only a multiplicity of incompletely specified models.

Formalism and Platonism have been much discussed in the philosophy of mathematics. My use of the terms here is intended not to contrast two different views of mathematical reality, but rather to foreground their role in shaping the professional vision of mathematicians. That is to say, I am interested in the extent to which a mathematician will ascribe value (positive or negative) to a problem or a mathematical result on grounds that it is formalist or Platonist.

To my surprise I found little sympathy for the formalist point of view in the narratives I examined. On the contrary, Davis (2000) devotes an entire chapter to the topic "How I was turned off formalism" (pp. 50-53). Davis recounts how in the fall of 1942 he studied mathematical logic under W.V.O. Quine. At the time, Davis was "madly in love with the logical notation in Whitehead and Russell and followed, more or less, by Quine... I expressed a goodly fraction of my thesis in the notation of the *Principia*. Even as I was doing it, I realized that this notation was not eliciting any new substance from the basic questions I tackled. But I did it anyway. It was a useful exercise and led me to the conclusion that the relationship between form and substance in mathematics is an exceedingly complicated matter. I don't think that anyone has yet written deeply on this dialectical split." Unfortunately, Quine gave Davis a grade of B in his course, "and that," says Davis, "is why ever since I have disliked the philosophy of formalism."

Davis is not the sole example in the mathematical narratives of an explicit recognition of the opposition between formalism and Platonism. Hardy (1940) writes:

> There is no sort of agreement about the nature of mathematical reality among either mathematicians or philosophers. Some hold that it is 'mental' and that in some sense we construct it, others that it is outside and independent of us.... I should not wish to argue any of these questions here even if I were competent to do so, but I will state my own position dogmatically in order to avoid minor misapprehensions. I believe that mathematical reality lies outside us, that our function is to discover or *observe* it, and that the theorems which we prove, and which we describe grandiloquently as our 'creations', are simply our notes of our observations.

Hardy's Platonism does not seem to be exactly the same as Davis's preference for substance over form, but they are alike in two respects: they both reject formalism, and they both profess themselves unqualified (and to a degree uninterested) in any detailed or systematic discussion of philosophy of mathematics.

Another dialectical pair of categories of appreciation in the narratives is *simplicity* and *complexity*. The value of simplicity can be illustrated in the following anecdote, also from Davis. He recounts how, as a young student of algebra, he was perplexed by why the left-hand side of a quadratic equation "should always add up to zero." Eventually, after the passage of some months, Davis realized the answer to his own question: "One simply moved all the terms to the left side of the equation, and zero was what was left on the right hand side." Some time later, Davis had the opportunity to discuss his confusion, and his eventual resolution of it, with the mathematician L.L. Silverman, who told him:

> Mathematicians have a tendency to throw everything onto the left-hand side... They think it makes things neat. You know who else did it recently?... Einstein. Einstein wrote down $G + cT = 0$ and that, he said, sums up the universe.

This brief anecdote illustrates well the aesthetic value of *simplicity.* Although mathematicians often work in areas of great complexity, there is a profound appreciation for notation that allows complex notions to be expressed in a simple form, and for broad syntheses that, in the words of Wiener, find “order among disorder”. It is notable in this connection that Einstein once claimed as one of his most significant discoveries in mathematics a notational convention that allowed complicated tensor equation to be written more economically[12].

It is important to note that notational simplicity — such as moving all the terms to one side of the equation — has more than merely *aesthetic* value. As Davis notes, doing so allows for seemingly distinct problem types to be brought together under a single conceptual umbrella. Davis notes that mathematicians who worked before the invention of negative numbers were forced to regard equations such as $3x^2 + 10x = 30$ and $3x^2 + 30x = 10$ as belonging to essentially different classes of problems, requiring different methods of solution; “they could not have written down $ax^2 + bx + c = 0$, and think of it as the generic case.” The value of “simplicity” thus has a strong connection with the production scheme “generalization” (see Kolmogorov’s remarks on his generalized, axiomatic theory of integration, quoted above).

The antithetical partner of “simplicity” is “complexity”. The collection of mathematical narratives used for the present study did not include any clear illustrations of this value, but I would be remiss if I did not note that some pieces of mathematical work are highly regarded precisely because of the enormous complexity of the

---

[12] Einstein introduced his summation convention in 1916. According to Pais (1982, p. 216) his later claim (in a letter to a friend) that the convention was a "great discovery in mathematics" was meant in jest, but I see no reason not to take it seriously.

undertaking. One recent example in this regard is the successful mapping of the space of irreducible representations of the exceptional simple group E_8, completed in March 2007. This accomplishment, the result of four years of collaboration by a group of 18 mathematicians and computer scientists, involved staggeringly complicated computations that many believed were simply out of reach. Certainly this result captured a great deal of attention in large part because of its great complexity. Similar observations could be made regarding the computer-assisted proofs of the Four-Color Conjecture and Kepler's Conjecture (both of which also could be appreciated for the value of "confirmation").

Summary

In this chapter I have attempted to partially map out the network of dispositions — categories of perception and appreciation — that constitute the mathematical sensibility, as it is brought to bear on the valuing of results (implications of the form $P \to Q$) and on the posing of questions (which can be regarded as partial anticipations of future results). Following the view of scholars who argue that knowledge of a practice is recorded and exemplified in *narratives of practice*, I have turned to a collection of narratives by and about mathematicians to understand the particular features of the mathematicians' worldview.

In the course of this analysis I have described the mathematician as, primarily, one who *wonders* about what is true — more specifically, about the implications and interrelationships among complex sets of contingent properties of abstract objects. This wondering is not, however, an aimless meander around a shapeless mathematical terrain; on the contrary, the narratives show that mathematicians structure the problem space under investigation through the use of certain generative moves that generate new questions from old ones, and that in their appraisal of certain problems as being more or less important than others they are guided by a set of dispositions that are organized in dialectical pairs. I have characterized the generative moves as *categories of perception* in the sense that they can be used to describe what kinds of questions a mathematician sees as a possible line of inquiry. But I have also noted that mathematicians do not pursue every question that occurs to them; on the contrary, there are *categories of appreciation* with which the mathematician organizes the landscape of possibilities, deciding which are worthy of further inquiry and which are not. Like the moves, these categories of value are organized in dialectical pairs.

These categories of perception and appreciation are displayed in Figure 2.3. These eight dialectical pairs may be taken as a map of the mathematicians' practical rationality with respect to posing and valuing mathematical results — the vision and values that make up that portion of the landscape of the mathematical sensibility.

Categories of perception (generative moves):

(strengthen/weaken) hypothesis	⟷	(strengthen/weaken) conclusion
generalize	⟷	specialize
consider converse	⟷	consider converse

Categories of appreciation

utility	⟷	abstraction
surprise	⟷	confirmation
theory-building	⟷	problem-solving
simplicity	⟷	complexity
formalism	⟷	Platonism

Figure 2.3. The mathematical sensibility with respect to posing problems and valuing results.

In the next chapter, I gather evidence from a corpus of study groups among groups of experienced teachers to show the role that these dispositions customarily play in the teaching of secondary geometry. In the chapters after that, I use these dispositions to code a corpus of examination questions from one geometry teacher, and explore the question of whether the geometry classroom can serve as a site for teaching students the elements of a mathematical sensibility.

Chapter 3

The mathematical sensibility in the practice of experienced geometry teachers

Introduction

In the previous chapter I analyzed a collection of mathematical narratives to identify some of the dispositions that are characteristic of the mathematical sensibility. In the present chapter, I use that conceptual analysis as an organizing framework for an empirical study of teachers' practical rationality. This study attempts to answer the question: What role does geometry instruction customarily provide for the teaching and cultivation of the mathematical sensibility?

As was discussed in the previous chapter, the French sociologist Bourdieu (1998) describes a practice as being characterized by a collection of shared dispositions (categories of perception and appreciation). In the last chapter I attempted to describe the dispositions that are characteristic of the practice of research mathematicians. Here I undertake the analogous question for a different practice: that of the *teaching* of mathematics, and in particular secondary geometry.

It is important to stress that here I wish to focus on the study of *teaching*, rather than the study of individual teachers or individual teaching performances. There is, of course, much about what teachers do that is personal and idiosyncratic; but there are also large components of teaching practice that seem to be shared in common among practitioners. These stable elements of instruction, replicated year after year, in classrooms across the country, have been approached by scholars in various ways: via cross-cultural comparisons (Stigler & Hiebert 1999), studies that probe teachers' practical rationality (Herbst & Chazan 2003a, 2006), and analyses of classroom discourse (Leinhardt & Ohlsson 1990; Leinhardt & Steele 2005; Lemke 2002). In the case of the present work, such study provides a vital counterpoint to the conceptual analysis of the mathematical sensibility, on the one hand, and the corpus of examination questions from a single classroom, on the other: it makes it possible to ask to what extent the values and practices described in detail through the analysis of tasks created by one individual teacher are like, or unlike, the values and practices of the profession writ large, and in turn to what extent those values and practices correspond to those of the discipline of mathematics.

My data source for this study comes from an archive of focus groups and study groups collected by the research group GRIP (Geometry, Reasoning, and Instructional Practice), particularly those that are part of the research project ThEMaT (Thought Experiments in Mathematics Teaching), directed by Patricio Herbst of the University of Michigan and co-directed by Dan Chazan of the University of Maryland. Building on the ethnomethodological tradition, and consistent with the theoretical position that knowledge of a practice is stored and recorded in the form of narratives (see Chapter 2), ThEMaT uses "breaching experiments" to probe the norms of a social practice (Herbst & Chazan 2006). In a breaching experiment (Mehan & Wood 1975), individuals and groups who are hypothesized to share a (perhaps tacit) set of norms around certain social practices are placed into an artificially-constructed situation in which many of those hypothesized norms are present but some are violated or subverted. Through their behavior, participants both mark and *repair* those breaches; that is, they indicate how conditions could be altered so as to render the situation normal once more.

The ThEMaT breaching episodes are narratives of teaching practice, represented with a variety of rich-media tools, including video, animations, comic books, and written vignettes. The episodes are designed to blend together elements that are hypothesized to be normative of a situation, together with elements that are hypothesized to be breaches. These representations of teaching are then used to prompt discussions among experienced practitioners about planning, teaching, and assessing; about what it means to be a successful teacher, and what one can expect from one's students; and about managing the tensions that arise as a result of the (normally unrecognized) competing imperatives that teachers must negotiate. The practitioners respond to the represented teaching episode by noticing some things while ignoring others; by endorsing or indicting what they see there; by specifying the conditions under which such events would be possible or impossible, desirable or undesirable, etc.; and by narrating alternatives. These alternatives comprise a web of interconnected possibilities — the "thought experiments" of the project's title.

In keeping with the goal of understanding the rationality of the *practice*, as opposed to the judgments of individual practitioners, Project ThEMaT brought together groups of experienced teachers to view and react to the representations of teaching collectively. By doing so it is hypothesized that teachers will feel some obligation to not only express their own opinions, but also to defer to their perceptions of what is customary within the practice at large, as embodied by their peers. This is not to say that teachers will completely elide their own individuality, but rather that they will do so in a manner that simultaneously acknowledges the norms of the practice. The word "norm" is meant to denote not a regulation, the violation of which carries a sense of transgression, but is rather used in the sense of a central tendency around and against which teachers position their idiosyncrasies (Herbst, Nachlieli, & Chazan, submitted). Thus, for example, a statement such as "Most teachers wouldn't accept that as an answer, but I'm not so picky" provides confirmation of a norm while simultaneously critiquing it and alienating the speaker from it. By gathering teachers in study groups composed of their own experienced colleagues, as opposed to conducting individual interviews, it is expected that teachers will feel some obligation to at the very least *mark* the norms, even if they do so by distancing themselves from them.

During its first year, 17 teachers participated in the ThEMaT project. These teachers were divided into two groups, each of which met once every month during the school year for an after-school study group. In addition, all teachers met together for daytime study group meetings at the beginning and end of the school year. Some of the animations were designed to raise issues related to the teaching of theorems; others were intended to focus on the work of engaging students in proving. In a typical study group session, teachers would watch a new animation or review a previously screened one, and discuss the mathematical and pedagogical issues these animations raised to them. In the second year of the project, 23 teachers (including 11 returning participants from year 1) met for another set of study groups. In all, the ThEMaT data corpus from the first two years of the project consists of more than forty study group meetings over a two-year period, each approximately three hours in length, yielding more than 120 hours' worth of videorecords and more than a thousand pages' worth of transcript data, with 29 participating teachers.

In addition to the ThEMaT study groups, the GRIP data archives also include records from a smaller number of non-ThEMaT "focus groups". These focus groups were similar in structure to the study groups, with two principal differences. First, each group meeting was self-contained, in that participants did not return on a monthly basis to meet again with the same people. Second, the focus groups primarily made use of video records, rather than the ThEMaT animations that were at the heart of the study groups. Despite these differences, the GRIP focus groups have much in common with the ThEMaT study groups. In both, participants were confronted with representations of teaching — representations that included deliberate breaches of hypothesized norms — and asked to discuss the extent to which the episodes depicted were like or unlike their own classroom experiences.

As part of the process of preparing this corpus for analysis, sessions are parsed into intervals on the order of 2-8 minutes in length, on the basis of changes in patterns of activity, indicated by markers such as changes in speaker, turn length, attendant focus, use of material resources, and the like (Herbst, 2006-2009). Intervals are then tagged with markers for various theoretical foci, allowing subsequent retrieval and analysis. All of this data has been entered into a complex, multi-relational database, allowing rapid cross-referencing between transcript, intervals, video, information about participants, session agendas and animations, and artifacts, as well as searching and aggregating data both within specific sessions as well as across the entire corpus. The result is a rich and extensive archive of information, consisting of more than 66,000 turns of transcript, associated to 2945 intervals, that can be mined for various analytical purposes.

In this chapter, I draw upon these records to develop a sense of what is usual and normal (in the sense of the word "norm" as used above) in teaching. Using the search and cross-referencing capabilities of the GRIP-ThEMaT database, I identify intervals and turns of transcript that appear to provide evidence concerning the various mathematical dispositions that were identified in the previous chapter. In these intervals, teachers respond to animated vignettes that depict moves that might be either appreciated or indicted on grounds related to the elements of the mathematical sensibility described in Chapter 2, and look for evidence as to whether the elements of that sensibility are part of the teachers' practical rationality. In the sections that follow, I describe my method for locating this evidence within the database, and show how the evidence sheds light on the question of whether the practice of teaching secondary geometry provides a customary role for teaching the elements of a mathematical sensibility.

Methods for searching the data corpus

Searching through a data corpus on the scale of the ThEMaT archive for discussions of the mathematical sensibility poses a number of challenges. To begin with, with the exception of only a few rare cases, neither the session agenda nor the representation of teaching used as a prompt in the study group were specifically targeted at provoking a discussion of one or more of the mathematical dispositions. When discussion of those dispositions did occur, it often did so in the form of a digression. For this reason it was not reasonable to narrow the scope of the search by focusing on only a single session, or on a set of intervals that made reference to a specific animation. Rather, it proved necessary to search across the entire corpus. But it must also be recalled that sessions were parsed into intervals not on the basis of changes in thematic content but rather on the basis of changes in participation structure. For that reason, the metadata defining any particular interval contains little clue as to what participants were actually talking about.

For these reasons, I began to look for relevant data by searching the text of the transcripts themselves. I adopted an iterative search-and-retrieve algorithm:

- I began by generating, for each of the dispositions, an initial set of keywords that I thought were likely to occur in discussions of the disposition.
- A cross-corpus search for those keywords in the database generated a found set of transcript turns.
- Each of those turns was read individually to see how the keyword was used in it; "false positives" were discarded.
- The remaining turns were referred back to the intervals to which they belonged; the full transcripts of these intervals were then read in their entirety.
- If an interval found in this fashion was deemed to have some relevance for the disposition under investigation, the interval was tagged after that disposition.

- In some cases, the transcript for a tagged interval would suggest additional words or phrases that could be used as keywords. These keywords were then added to the list of search terms, and the process was repeated.

The above procedure yielded a set of intervals for each of the dispositions under investigation. Each of these intervals included one or more turns of speech in which participants invoked, disclaimed, or otherwise acknowledged a mathematical disposition as a warrant for pursuing a particular course of action.

Students as problem-posers

"Healthy confusion" as productive for problem-posing

The examination of mathematical narratives in the last chapter provided evidence that mathematicians view themselves as wonderers; that is, they place great value on questions that seek to determine what is true. While mathematicians certainly engage in other work as well — for example, seeking a new proof of a result already known to be true, or a simpler counterexample for a claim already known to be false — one dominant motif throughout those narratives was the motivating power of an encounter with the unknown. And the excerpts presented from Parker's (2005) biography of R.L. Moore offer a glimpse of how such an encounter with the unknown can be made a deliberate element of mathematical pedagogy. As was shown in those excerpts, Moore would ask his students to prove plausible-seeming but false claims; he would assign homework sets in which unsolved problems were concealed among straightforward exercises. These unconventional methods had in concert the effect of creating a learning environment in which students could not rely on their instructor to serve as the final arbiter of truth, and thus had to rely on themselves and each other to determine what was and was not true.

If, as these accounts suggest, the encounter with an unknown plays a significant role in supporting problem-posing, then one might wonder to what extent Geometry teaching provides a role for students to have such an encounter. Do teachers create contexts in which mathematical truth is deliberately made unclear, and use those contexts as opportunities to cultivate a problem-posing orientation in their students? If so, how are such contexts managed? These questions recall the challenges I raised in Chapter 1 concerning the "Lei" vignette found in *Principles and Standards* (NCTM 2000), and they lie at the heart of the problem of teaching students the mathematical sensibility.

Throughout the study group corpus we find several occasions in which teachers discussed the role of doubt and uncertainty in stimulating student inquiry. In the very first study group meeting in Year 1, teachers viewed and discussed "The Square"[13] — an animation in which a class collectively discusses the question, "What can be said about the angle bisectors of a quadrilateral?" In that animation, Alpha[14] states that he considered the case of a square, and found that in that case "the diagonals bisect each other". After a significant amount of discussion, Alpha re-shapes his claim as "In a square, the angle bisectors meet at a point, because they are the diagonals." Following the statement of this claim, a different student, Lambda, presents an argument that each diagonal of the square bisects both of the angles through which it passes. It soon

---

[13] This animation, and all others referenced in this chapter, may be viewed at the ThEMaT Researcher's Hub. Accounts may be requested at http://grip.umich.edu/themat.

[14] Students in the ThEMaT animations are named after letters of the Greek alphabet.

becomes clear, however, that Lambda's argument has been widely misunderstood by both his classmates, and (it appears) the teacher as well. The animation ends with a sense of general confusion, as one student asks, "What are we doing?"

In discussing this animation, one teacher made reference to the potential benefit of what she called "healthy confusion":

104	Lucille[15]	Well I was thinking that sometimes having a little bit of healthy confusion and debate actually serves a purpose, because if what's going on is not particularly clear, then people think "Oh okay, we're not sure if this right," and everyone starts thinking about, "is it right, and if it is, how do we know it is, and if it's not, why isn't it?" If you're always saving them by giving them the answer — eventually you have to do that because that's our job — you kind of rob them of that reflection of what they know, and um... Yeah, it's hard to know when and when not to, but sometimes I think it's almost better to allow the uncertainty to go for a little while because then more people start thinking about it than if you let the kid who always knows it fix it, or the teacher who always knows it fix it. And I think sometimes you start to see, oh there are way more possibilities of looking at it than you or I had thought about. So I know it's a hard call, as the bell rings at a certain time.

Excerpt 1. From ThEMaT081905, interval 6.

Notice that Lucille at once endorses the potential benefits of "healthy confusion" as motivating student curiosity, and recognizes that the teacher can quash that curiosity by stepping in and resolving all questions. At the same time, however, she signals that such confusion cannot be allowed to last indefinitely: "Eventually you have to do that [give them the answer] because *that's our job*" (emphasis added). This characterization of the teacher's responsibility to be the arbiter of truth stands in marked contrast to the approach of R.L. Moore. Lucille's ambivalence towards the idea of letting students be confused over extended periods of time was echoed and expanded on by other teachers in the interval immediately following:

[15] All teacher and institution names in this chapter are pseudonyms.

110	Penelope	I agree that a little healthy confusion is a necessary part of the classroom in order to get students into the process of discovery. Because sometimes the teacher may not always be the one to say, "Hey that's not exactly the correct path we're going to go down." Some of your classmates may say, "Hey, you know, I don't think so, I see it this way." So a little healthy confusion can spark a dialogue between the students, and I believe in seeing the teacher as a facilitator, just as someone to kind of say, "Hey this is where we're going today and I want you to devise a plan for getting there." Let students be individual in their own thought processes in getting there.
111	Greg	I think the healthy confusion is really good, but I think there's a fine line between letting it go too long also, because then the kids start confusing the kids who were starting to get it. And then you have chaos. So, I think healthy confusion is good, controlled and only to a certain extent though.
112	Megan	You know, the timing, too. When she said that I thought, what a great phrase. I think healthy confusion is very good if it occurs at the beginning of the period. [Laughter from group] If it occurs in the last five minutes, it's a very bad thing! [laughter] I just, personally, I don't like kids to leave thinking, "What just happened there? Do we even know what we were doing there? What a waste of time, I didn't come to any conclusion." So the timing of healthy confusion I think is very important.
Excerpt 2. From ThEMaT081905, interval 7.		

Megan's worry about ending the class with students confused is consistent with Lucille's observation that "the bell rings at a certain time" (104). Among the teachers in the study group, there appears to be an emergent consensus that, while confusion can play a productive role in motivating inquiry and "sparking dialogue", such confusion must be resolved by the end of the class period.

Once named by Lucille in the first study group meeting, the notion of "healthy confusion" was recalled periodically in subsequent study groups. For example, in a meeting one month later, teachers viewed an animation titled "The Tangent Circle". In that animation, a teacher asks her class to draw a circle tangent to two given lines at two specified points in a pre-printed diagram. As posed, the problem is impossible because the given points are not equidistant from the intersection of the two lines; when one student argues that the points should be moved, her proposal is ridiculed by her classmates, who regard it as an illegitimate change of the problem. In their subsequent discussion of this episode, teachers described various ways of restructuring the task so as to reduce the frustration they perceived among the students in the animation and to channel it in a more "healthy" direction:

636	Denise	I mean, so if you know that your students were starting to get frustrated, and I know my students, I can tell when they're starting to get frustrated, and if I can't tell they're sure enough gonna let me know, I need to start going in a different direction, and maybe it's time to stop exploration and give them more directives, because I don't want them to shut down, because then they can't learn anything. Am I making sense?
637		[participants make affirmative sounds]
638	Moderator	But then, do you use this frustration for them to, to get to a point that they wouldn't?
639	Denise	Yeah, sometimes. Sometimes, like I said, if it, the problem he said, you give them the one point, and then you give them the two points after that, like I think like most of my students that'd spark curiosity, that's frustration, but it's good frustration, like, [inaudible]
640	Cynthia	What did we say? Healthy confusion.
641		[laughter]
642	Denise	Yeah, healthy confusion.
643	Researcher	And, and is this one a case of constructive frustration, or is it not?
644	Denise	I, I don't think so.
645	Megan	She thinks that's *confusion*.
646	Denise	That's just confusion, yeah.
647	Researcher	That's just confusion, okay.
648	Denise	If she would have stopped there at some point and did something else, that might have been healthy confusion, healthy confusion.
Excerpt 3. From ITH092805, interval 26.		

For these teachers, the distinction between "healthy confusion" and "just confusion" hinges on timing and on the teacher's responsibility to step in at the right moment. According to Denise, the teacher ought to have sensed that the students were getting frustrated and seized the moment to "stop exploration and give them more directives". In failing to do so, the teacher allowed a potentially healthy confusion to degenerate into mere chaos.

In that same study group meeting, participants discussed the possibility of leaving questions unresolved over the course of multiple lessons:

511	Moderator	If you did that two or three classes in a row, do you think they would cha – who would break first?
512	Karen	Me.
513		[laughter]
514	Researcher	But why?
515	Karen	Because they really don't want to learn the stuff.
S516	James	[spells out] S-T-M-P.
517	Researcher	Because of the STMP[16]?
Excerpt 4. From ITH092805, interval 45.		

Karen and James are here seen to propose separate explanations for why leaving questions unresolved over extended periods of time is not a viable way of engaging students' interests. For Karen, the problem resides with the students themselves: she sees them as essentially uninterested in "the stuff" of the course. Karen's comments echo those made in an earlier focus group by Laura and Rick:

127	Laura	I guess one of the things, I'm sorry, one of the things that I um struggle with when it comes to proof is, we as math teachers, we're curious. I mean we're sitting here, examining this quadrilateral, and trying to think of everything that we can possibly think of that might be true or can be proven to be true, but our students many times don't have that same curiosity. They, their mindset is....
128	Rick	Just tell me.
129	Laura	Who cares. Yeah, just tell me. And that's what I'm curious about, how do you develop that curiosity in your students, how do you get them to *want*.
Excerpt 5. From ABP-081704-1.		

Similar comments are found throughout the data corpus, as when for example Denise says of her students, "If I say it, they believe it. They don't question anything" (ITH092805, interval 17, turn 253). At still other moments, teachers speak as if mathematical curiosity is an innate quality that is held by some students and not by others:

[16] The STMP (State Test of Mathematical Proficiency) is a pseudonym for a standardized test taken by all students in the state.

576	Cynthia	Some classes are more inquisitive, I think. My first hour is very inquisitive, they ask questions, and they are like, Oh I wonder what would happen if, like that, but my fourth hour class.... I have to humor them a little bit, for them to pay attention. I, serious, I...
577	Denise	But it's the end of the, it's closer to the end of the day.
578	Cynthia	Right, and they're ready to leave, and so to keep their attention, I have to throw out "What do you call a dead parrot? A polygon." You know, and they stay with me that way. But first hour I don't have to do that because they're already with me no matter what, because I've got a lot of freshmen in there, they're very talented, and the other kids just kind of go with, and just depends on the personality, like you said.
579	Moderator	So you provide different tasks to those kids?
580	Cynthia	No, I don't give different tasks, I just approach things differently some times.
581	Tina	I, like my fourth hour, they have the Garfield effect I call it, you know Garfield the cat, he eats and he wants to sleep? They come back from lunch [laughter] and I'm doing everything I can to keep them awake, and it's the hardest time. Seriously, they eat and they want to take a nap. And then my sixth hour, like you said at the end of the day, they're ready to bounce out the door, they, the last thing they want to talk about is geometry, and then my class in the morning, you know like you said they're asking questions, they're with ya, they're...
Excerpt 6. From ITH092805, interval 34.		

Note here how the various speakers adduce different explanations for why some classes seem to be more curious and inquisitive than others. Cynthia attributes these variations to intrinsic differences among students: 9th grade students taking Geometry are more "talented" than are 10th graders taking the same course. Denise and Tina, in contrast, ascribe these differences to the time at which the class meets (right after lunch, at the end of the day, etc.) But in all cases the teachers seem to regard the extent to which their students wonder about things as a fact of life, one that they must simply accept as it is; none of them speak of trying to cultivate curiosity or teach inquiry. Laura's question quoted above ("How do you get them to *want*?") seems to be a rhetorical one, as neither she nor any of her colleagues offers an answer.

In contrast to the view just described, which ascribes the lack of viability of leaving questions unresolved over the course of multiple lessons to intrinsic qualities of students, James's observation regarding the STMP exam (excerpt 4) points to the institutional realities of schooling (and in particular, the need to prepare students for standardized testing) as the reason why such a practice is unworkable. It is not quite clear from the transcript of excerpt 4 whether we should understand James's turn 516 as an alternative to Karen's turn 515, or an expansion of it. That is, he may be suggesting that one effect of standardized testing is to deaden or suppress students' curiosity; on the other hand he may be pointing to standardized testing as a constraint on the *teacher's* ability to leave questions unresolved over an extended period of time.

Despite their aversion to the prospect of allowing doubt to survive beyond the boundaries of the class period, teachers were not entirely unsympathetic to the notion that such doubt can play a productive role in motivating student inquiry. In one study group meeting (ESP101105), teachers were asked to respond to the following prompt:

> Suppose in a classroom students proposed the following three conjectures:
> 1. When you join the midpoints of the neighboring sides of any quadrilateral, the result is a parallelogram.
> 2. The parallelogram formed in this fashion has exactly half the area of the original quadrilateral...
> 3. ... and half the perimeter of the original quadrilateral.
>
> What would the teacher need to have done in order for these three conjectures to emerge?
> What would the teacher do next to engage the students in planning a proof, and proving these conjectures? [17]

[17] As an aside, I note that although the first two of these conjectures are true, the third is false. This was not mentioned by the teachers in the study group, and the moderators did not bring it to their attention. The fact that the teachers themselves gave no indication of wondering about the truth of these conjectures is obviously relevant for the present discussion.

In their discussion, teachers considered the possibility that students might use the measurement tools in dynamic geometry software (such as the Geometer's Sketchpad) to discover the second and third properties. In this context one of the teachers cautioned that such software can be so forceful at persuading students of the truth of a proposition that it can have the effect of inhibiting their capacity to inquire further:

53	Esther	They might, they might think they've proved it because it worked on Sketchpad. And they might be less inclined to think they have some reason to prove something works because...
54	Carl	They don't understand that Sketch-, that Sketchpad is not proof.
55	Esther	Yeah.
56	Carl	They don't --
57	Esther	-- Maybe when they do it by hand somebody's works and somebody else's doesn't so now they're not really sure if it's right or not. So, can you give some other evidence about why something works? Whereas in Sketchpad they say well, it's that, you know, this plus this is this, I mean, there it is. They don't see that they need as much evidence, I think sometimes when they see --
58	Moderator	-- Is that something that happens? Where, um, because some students have found a relationship and other students have not that they feel the need to do a proof?
59	Esther	Well, I would like to think that...
60	Melissa	[Laughing] Feel the need to prove!
61		[General laughter]
62	Moderator	Or is it you?
63	Esther	Yeah, yeah. I don't know that they care, but sometimes they, I have some kids that would say well, why did this one work and why did this one not work, or is there some [rea--], does it work sometimes and not other times, or did somebody do something wrong in the picture? I mean I've had that happen, I don't think --
64	Moderator	-- And then that, that can serve then as a launch for the proof?
65	Esther	I wouldn't say that it happens on a regular basis, but I've had that happen before.
Excerpt 7. From ESP101105, interval 6.		

The salience of this excerpt for us is located in Esther's recognition that uncertainty (here engendered by the fact that some students' "by hand" drawings "work" but others' do not) can motivate students to look deeper into a problem. Esther here describes the possibility that a student's activity might be driven not by a teacher's instructions, but by a desire to understand a puzzling lack of consistency. That, she says, is why using dynamic geometry software such as Sketchpad can undermine the imperative to produce proof: it is precisely because the software *removes uncertainty* that students would see no need for a proof (53). This recognition is tempered, however, by the laughter Melissa and her colleagues share at the notion that a student might "feel the need to prove" something (as if it the very notion were absurd), in response to which Esther hedges her earlier comment with the concessionary "I wouldn't say that it happens on a regular basis".

Honesty and Deception

It is one thing to tolerate doubt and uncertainty to remain over the course of several lessons; it is another thing else to *actively court it* by asking students to prove untrue claims, or by declining to intervene when students make errors of reasoning. When asked to consider scenarios in which a teacher was depicted making such a move, the teachers in the ThEMaT study groups were unanimous in their opposition. This was particularly evident in the teachers' responses to the ThEMaT animation "The Parallelogram", a story in which the teacher makes extensive and repeated use of this teaching move.

"The Parallelogram" begins with a teacher reminding his students that "When we studied triangles, we saw that the angle bisectors would always meet at a point," and asking the question, "What happens with a parallelogram?" After 8 minutes, Alpha volunteers to present his work. Alpha draws a picture of a parallelogram with its diagonals (see Figure 3.1) and announces his intention to prove that the angle bisectors of a parallelogram meet at a point. In the animation, the teacher gives no indication that Alpha's claim is incorrect (which it is), or that his diagram indicates some confusion between diagonals and angle bisectors. Instead, he simply asks the class what the "Givens" would be for Alpha's proof. Theta, referring to Alpha's diagram, states that $\overline{AC}$ and $\overline{BD}$ are angle bisectors. The class begins constructing an argument using that as a premise — although it is somewhat unclear (from an observer's perspective) what exactly they are trying to prove. The class begins identifying congruent angles and segments, presumably for the purpose of eventually identifying some congruent triangles.

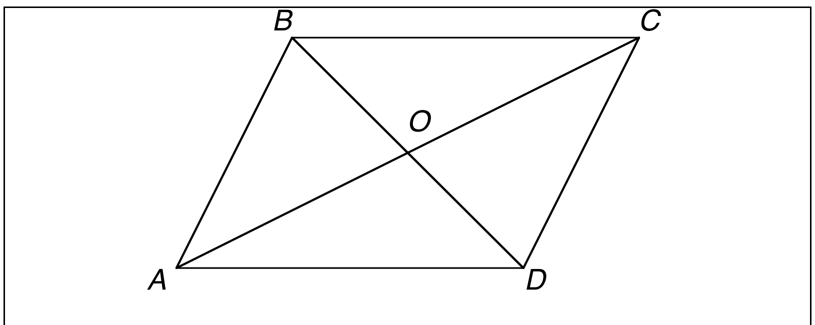


Figure 3.1. Alpha's diagram.

At one point in the animation Theta raises an objection: "But how do you know that the angle bisectors are the diagonals? I mean, don't you have to prove that?" To this Alpha responds, "No — that's kind of my given, see?" In response to this exchange the teacher comments, "So, Alpha, you really are using that the angle bisectors are the diagonals as a given." Alpha agrees, and the discussion resumes. Eventually Theta notices that triangle *ABC* is isosceles, and consequently that the figure is a rhombus. But to this Epsilon objects that "You would only have an isosceles triangle if you start from a rhombus, and *that* is a parallelogram." Alpha, puzzled, says that they seem to have proved that their parallelogram is a rhombus, to which he adds, “This is crazy.” Iota asks, “Are you then saying that all parallelograms are rhombi?” The teacher responds, “Is that what you think we are saying, Iota?” General confusion spreads throughout the class, until finally Xi and Epsilon identify Alpha’s initial assumption (that the diagonals are angle bisectors) as problematic and false for non-rhombic parallelograms. Having reached this resolution, the teacher asks, “Epsilon, can you predict what the angle bisectors of such a parallelogram *would* look like?” Here the animation ends.

In the study group meeting (ESP110105) teachers reacted strongly to the teacher’s decision to allow incorrect statements to go uncorrected, and in particular to the way he encouraged students to pursue a line of inquiry that was founded on an incorrect presumption. When the moderator asked participants to propose topics for discussion, Carl and Karen both raised this concern, echoed shortly after by James.

368	Carl	I think the whole concept of going with false information from the get-go. Assuming that diagonals are the same as the angle bisectors and going with it.
369	Moderator	Okay. Other things?
370	Karen	I’m just, you know, wondering if kids could stand the agony of it not working out as long as it went.
371	Researcher	Not working out?
372	Karen	Yeah, like there’s no resolution, it’s like, it’s like a plot that’s, you know, we’re just getting tangled and more and more confusion.

386	James	It really bothers me that the teacher doesn’t come in when they’re sitting there talking about the angle bisectors and the diagonals interchangeably, and doesn’t try and stop the discussion there and get that taken care of.
387	Moderator	Do you think – should the teacher do that like right away, as soon as it comes up?
388	James	Yeah, I mean, as soon as that’s put up there and they’re startin’ – well, at least, at the very least, when all of a sudden he’s got a given up there that it says it’s an angle bisector, and he says well that’s my given, that it’s an angle bisector and a diagonal. I mean...

397	Moderator	So then your response to Karen’s question of how long should this go on is, it shouldn’t even start, basically. As soon as the picture’s put up there-
398	James	I’m like Carl, they just spent 15 minutes, and now you’re at the dilemma of

		do we go back, or do we just, you know, chuck the whole project and throw it out?

401	Moderator	... You're saying that the teacher would not endure that?
402	James	I wouldn't.
403		[laughter]
404	James	I'll let them make mistakes and find conflicts, but I mean that was just – I mean to actually come and say, that's part of my given – I couldn't handle it. [laughs] I'm sorry.
Excerpt 8. From ESP110105, interval 36.		

As the discussion of this episode continued into the subsequent interval, teachers were in general agreement that the teacher should not have allowed Alpha's misconception to continue without correction, and went so far as to propose a method for preventing it from ever surfacing in the first place: to exaggerate the proportions of the parallelogram so that "you weren't gonna connect *A* and *C* and call it an angle bisector" (Carl, interval 37, turn 431). Esther notes that prior to calling on Alpha, there were several minutes during which the teacher circulated around the room while students worked independently; based on this observation, Esther proposes that the teacher would have had the opportunity to call on a different student, one that had been observed to have a correct diagram, to point out the special nature of Alpha's drawing. If no such student could be found, then Lynne suggests "I would add my own drawing to the board and say, 'This is weird. This is the drawing I was thinking of.... What's the difference here? Talk to your partner for three minutes" (interval 39, turns 450, 452).

In the face of the widespread agreement that Alpha's mistaken assumption would have been better off avoided (or at least corrected immediately), teachers in the study group were hard-pressed to ascribe any intentionality to the animated teacher's decision to call on Alpha. Only one of the six teachers present, Karen, had any sympathy for the episode as time well-spent, a position to which Lucille responds with a lengthy caution:

555	Karen	What if her objective... was to get them to start to talk about, to have a discussion about what was going on and discover that they had made a mistake, you know, like just have a discussion and discovery, you know, like, to start to get some sense of how you can backtrack and follow down the path and then back up and follow down another path. And that there's nothing wrong with it there. They did the right thing.
556	Moderator	Umm...
557	Researcher	But, is it–I mean, what do the others think about that? I mean, I know that you could do it, but sometimes, sometimes we hear well uh, there's not enough time for that, and, you know, you always have to make compromises among all the things that you could do. Which of those are worth doing, because your resources are limited. Is this one that would be worth doing for the sake of uh, having an exploration?
558	Lucille	I think, [inaudible] you know, I think there's, there's a safety issue, like if you know that your class functions in a way where you can explore, and not be ridiculed for going down the wrong path, and if you felt confident that in the end wherever that end was, that clarity would come into play, then kids might be okay with this. But for example, if you never do this, because you're not comfortable with it as a teacher, and all of a sudden you do it, they're gonna start to go, hey well wait a minute, because they've learned your style and your approach too. And they might wonder well what's going on here, don't you really know? You know, so, um I don't know, it does seem like a really good exercise in the development of thinking through why things do work. But I think you [gesturing to Esther] said at one of our meetings a while back, what about the one kid or the two kids that never really catch on and all this conversation's going on, and after a while they may just tune out because they don't really know what the discussion is about because they're still just learning the basics, and so, then you're having really good conversations with a small number of kids, and the other people we're hoping are listening, and they might–maybe they are, but a good chance they're not, because they're young, and high schoolers are like that.
Excerpt 9. ESP110105, interval 45.		

Within Lucille's long turn 558 we can identify three separate sets of grounds on which to indict the decision to call forward a student with an incorrect idea. First, one risks doing harm to students' self-esteem: this is signaled by her use of emotionally charged language ("safety issue", "not be ridiculed"). Second, the teacher runs the risk of appearing ignorant or incompetent ("Don't you really know?"). Third, the teacher ends up teaching only to a small number of bright and attentive students, ignoring the majority of her class who "tune out" and "never really catch on".

The first of these points — that teachers need to protect students from public error in order to preserve their emotional well-being — was raised in other study groups. When teachers first discussed the animation "The Tangent Circle" (in which the teacher gives students a problem that cannot be solved as posed) one teacher[18] asked, "Do you think it's like, that it's bad in general to let kid, to sort of lead them in the wrong, to give them a problem that's really not solvable? Do you think it erodes your trust, or...? 'Cause I struggle with that a little bit you know." When the same question was posed to another study group one year later, participants had mixed feelings about the risks entailed, and identified the conditions under which such a move would be safe:

519	Moderator	... In the past when we've shown this movie to – or these scenes to other groups of teachers um, we've had people comment that if you do this it can erode the trust that students have in their teachers – like it can actually sort of damage that relationship because they feel like you're playing gotcha with them all the time because you're always trying to trick them or something. Does anyone have any – share that feeling at all, that there's sort of a risk of giving them problems like this?
520	Catlynn	I think you can't do it all the time, yeah and I think you – it would erode them if they...
521	Cadie	But they also have to realize that mathematics is not always trusting, I mean you're not going to be able to get to that answer.
522	Jillian	And what if they knew that up front, what if they knew like you know you build this trust with your class up early in the beginning and you give them a few easy ones and there's humor so they can laugh about it and maybe for the low level, we were talking about some of the lower learners; maybe where it would be a game or a gimmick where it's a gotcha problem, but it's part of how we operate. And it's not meant to be like me one-upping you, it's just sort of how we can all do it. Or can you come up with a problem, your team for this team, that looks like it can be solved but it can't? I mean maybe it's something we just change the climate of how we teach so that it's not – like I hate when teachers are always like oh I'm right and I got you. I don't like that arrogance in the classroom, y'know, but I would like them to know, I don't always know the answers and some mathematicians never know all the answers and that's life, that's why you're here.

[18] Megan in SG-ITH81905, interval 21, turn 388.

523	Melanie	And it'll depend on how you close it, it's okay to figure out that it didn't work, but eventually if you figure out what did work, I mean then you still have success and they don't feel like something went wrong, I don't think.
Excerpt 10. From ThEMaT-NEW-082206, interval 39.		

Particularly noteworthy here is the seeming contradiction between Catlynn, who says that such a move would be damaging if it happened "all the time", and Jillian, who argues that such a move would be damaging *unless* it were made a normal and expected part of classroom culture. But despite this apparent contradiction, Catlynn and Jillian in fact seem to be in agreement that such behavior is *not* normal. Catlynn seems to be describing school as she knows it, while Jillian appears to be speaking of school as she wishes it *could* be, but not as it actually is: the frequent use of qualifying hedges ("if", "maybe") labels her speech as counterfactual.

The second theme addressed in Lucille's long turn in Excerpt 9 — that the teacher runs the risk of seeming ignorant or incompetent — does not seem to have been a recurring theme in the study group corpus; I have not found any other examples of teachers raising it as a warrant against the decision to allow an incorrect statement to go unchallenged. On the other hand, the third theme (the risk of alienating or losing the "weaker" students in the class) was frequently cited in this connection. One of the greatest perceived risks that teachers associated with "misleading" students is the likelihood that some students will come away from the class with incorrect beliefs. This was expressed clearly in a subsequent study group, in which participants read aloud a script for a proposed sequel to "The Tangent Circle". In that script, the teacher draws a circle with two intersecting tangent lines, with a ray drawn from the two lines' point of intersection to the center of the circle, and (what appears to be) a diameter passing through the two points of tangency and the center of the circle (Fig. 3.2).

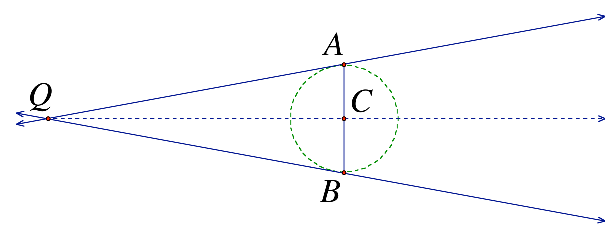


Figure 3.2. The teacher's diagram.

In the script students use this diagram to prove that the center of the circle lies on the bisector of the angle formed by the two intersecting tangents. After the proof is completed, one of the students (Lambda) notices that the diagram is flawed: the radii ought to be perpendicular to the tangent lines, and the two angles formed at *C* should not be right angles. The teacher acknowledges that the diagram contains an error, but points out that the proof that was written did not rely on the error. The script ends with the teacher asking the class, "So is the proof wrong, or is just the picture wrong? Or are they both wrong?"

In their discussion of this story, Glen argued against the teacher's use of an incorrect diagram:

1510	Glen	I just hate to put a drawing up that's misleading.
1511	Denise	Mm-hmm.
1512	Glen	Um, you know, again, you work so hard to get them to remember the correct ideas, and when you, as soon as I show 'em a wrong one there's going to be about 3 kids in the class that that's gonna be the only one they remember, out of the whole day. If I had a really, really excellent class, I could see using this type, where the vast majority of the kids could catch on to the error, discovering the error's a great idea, but, but if you've got an average class where you've still got kids struggling with the basic concepts, you're gonna confuse them all the more.

Excerpt 11. From ITH022206, interval 42.

Although he acknowledges that encouraging doubt and confusion about the truth can have a motivational effect on outstanding students or in a "really, really excellent class", as a teacher in a non-tracked classroom Glen feels an obligation to serve the needs of all his students, not only the ones he calls elsewhere "the bright ones". From one point of view, this position could be critiqued as condescending (in that it takes for granted that not all students have the intellectual capacity to handle challenging mathematics content) and essentialist (insofar as it regards intellectual capacity as something that the teacher must deal with, but has no power to change or stimulate). At the same time we recognize in Glen's comments another of the institutional constraints of schooling that the teacher must contend with, and that makes an elementary or secondary mathematics teacher a different kind of practitioner than a research mathematician, or even a mathematics teacher at the post-secondary level. The latter need not be overly concerned with teaching all students to succeed, and to a certain extent this becomes even more true as we look closer and closer to research mathematicians' practice. To take one example, it is clear from Parker's (2005) biography of R.L. Moore that, though Moore's courses were well-known for "converting" non-mathematics majors, an even larger number of students dropped his classes in failure and discouragement. Such an outcome is not desirable (or tolerable) for a secondary Geometry teacher. The need to teach all students is, like standardized testing, a fact of life that teachers must contend with, and one to which they may appeal when justifying instructional decisions that may seem at odds with the mathematical sensibility.

Summary

In the previous chapter I argued that one of the dispositions most central to the sensibility of mathematicians is a view of the mathematician as one who *wonders* what is true, and as motivated by encounters with situations in which the truth is unclear. The above excerpts from the study group and focus group data indicate some of the ambivalence teachers feel towards the idea that they can teach students to approach mathematics this way. On the one hand, many teachers recognize (at least in the abstract) the potential benefits of "healthy confusion" in stimulating student inquiry. This benefit, however, is bounded by a series of powerful constraints that teachers perceive on their freedom to encourage such "healthy confusion". First, teachers are extremely attuned to concerns of time: healthy confusion must not last too long or it becomes "just confusion", and on no account can it ever extend beyond the boundaries of a single lesson. Teachers who permit such confusion to last too long or become too muddled run serious risks: that they will accidentally reinforce misconceptions, which they will later have to "un-teach"; that their students will be unprepared for state-mandated standardized testing; that they will damage their relationship with students, eroding the trust that is an essential part of the teacher-student dynamic; that they will appear ignorant or ill-prepared to their students; and that they will in any event only succeed in stimulating the interest of a small number of the best students, neglecting their obligation to the rest. For all of these reasons there seems to be very little space in teachers' practice for cultivating this aspect of a mathematical sensibility in students: to do so would require, as Jillian suggested (Excerpt 10, turn 522), a radical change in the climate of the classroom and of the institutional conditions to which it is subject.

Modifying the hypothesis or conclusion of an implication

Beyond the characterization of mathematicians as individuals who wonder about the truth, one of the goals of the previous chapter was to identify particular dispositions (categories of perception and appreciation) that are typical of mathematicians. That is to say, within the essentially limitless set of things that mathematicians *could* wonder about, what kinds of questions do mathematicians perceive as worth doing, and with what types of appraisal are those questions (and work on them) appraised? In analyzing the mathematical narratives, I found that many of these dispositions are arranged in dialectical pairs. One such pair was *modify a hypothesis / modify a conclusion.* This pair of dispositions is deployed in various ways: as generalization, specialization, strengthening, investigating a converse and so forth. What all of these have in common is a focus on mathematics as the study of *conditional statements* — or, as Peirce (1882) put it, a recognition of mathematics as "the science that draws necessary conclusions". In the pages below, I examine the extent to which these practices (of strengthening and weakening hypotheses and conclusions of conditional statements) figure into the ThEMaT study group teachers' discourse about their classroom practice.

Considering the converse of a proposition or question

The teachers who participated in the ThEMaT study groups and the other GRIP focus groups spoke often of the important role conditional statements play in the intended curriculum of Geometry. In particular, teachers speak of exercises in which students are asked to rewrite conditional statements in an explicit "if-then" form, to state the converse and contrapositive of those conditional statements, and to know that a statement and its converse are logically independent (i.e. the truth value of one does not ensure the truth value of the other). This surfaced, for example, in a study group meeting (ITH102605) in which participants read and performed a story about three pairs of students working on a proof of the theorem "In a circle, two congruent chords are equidistant from the center." Notice that the "conditional" nature of this theorem is somewhat masked by its grammatical form: the "if" and "then" are not explicitly distinguished, and indeed it could be argued that the theorem is ambiguous[19]. Teachers were in general agreement that it would be a useful exercise to ask students to parse the statement into if-then form, and to ask students to prove both the theorem and its converse:

442	Tina	Okay, what we were just saying, if you put that in if-then statement, you can do an if-then statement of the converse. So you can have groups of proving either way. So do we want to prove one specific way? Or I mean, [Moderator: Wonderful question] are we trying to prove equidistance from the center or are we trying to prove if they are equidistant, then they are congruent?
443	Moderator	Do you think anything works? Do you want both directions? Do you want specific one?
444	Tina	Would it matter to any of you?
445	Moderator	What did you, what would you want your students to prove?
446	Cynthia	I wouldn't want them to think that it always works to prove it both ways. Ah, I mean, for other situations, not just, you know, not just this one.
447	Moderator	Not always
448	Cynthia	[inaudible]
449	Tina	But with our group, we've already done if-then statement and the converse as well [Cynthia: Right]. We know that sometime both are true [Cynthia: right], but sometime they're not.
450	Cynthia	Right.
451	Moderator	So would you let, would you want them?
452	Tina	That was my question to the group. I mean, would you put it in if-then statement or would you, you, do see what I am saying, if you, would you put it in if-then statement before you did this with this group? This is my question to

[19] In purely grammatical terms, the sentence has the same structure as "By a stream, two young children are napping." That is, the sentence asserts that that certain "things" (chords, children) with a specified property (being congruent, being young) also have another property (are equidistant from the center, are napping) — but any causal or implicative link must be inferred.

		the rest of you. [Denise says something inaudible] Would you just throw that up and have them prove it?
453	Moderator	Would you talk about the converse as well? Or just one direction?
454	Penelope	I think I would do something more like two-column proof and have them set up for themselves. So, "Okay, what's your given?" and then explain to me. I would have them probably go a little bit further, explain to me, "would this work in all cases in situations? What are some of the cases which this would work or which would not work?"
455	Tina	But my question is, would you set it up that you know that the chords are equal and you try to prove equidistance or would you prove that if they're equidistant from the center and then the chords are equal? Or would you let them in their own little groups figure it out and then find out whether they...?
456	Penelope	I guess I let them do their own and then go back to the theorem and ask them, now did you prove what the theorem asks?
457	Denise	Oh, I see.
458	Edwin	It's not very clear though (inaudible).
459	Denise	All I have is to set up, first, first I have to ask to write it in if-then because this is weird one. I know students will have trouble writing this. It will take more than two minutes to write if-then statement. But then once they did that I would have them prove it one way and then prove the reverse.
460	Moderator	So let them write if-then...
461	Denise	Yes.
462	Moderator	You let them write the if-then and then ask to do converse.
463	Denise	That'll be good practice.

Excerpt 12. From ITH102605, interval 22.

Tina's question to the other teachers (442, and repeated in 452 and 455) concerns the locus of accountability for teasing apart the hypothesis and conclusion of a theorem that is stated in an unparsed form: is this the teacher's responsibility, or should it be devolved to students? One possibility is that the teacher would translate the statement of the theorem into "if-then" form, identifying explicitly which property is assumed and which is to be proved. The alternative is that the teacher would "just throw that up" (452), i.e. pose the statement in its unparsed form, and let the students work on it, knowing that different students would be likely to interpret it variously. Although Tina does not voice a clear preference for one alternative over the other, her use of the word "just" in connection with the phrase "throw that up", connoting an almost careless approach to teaching, suggests that she locates the responsibility for parsing the statement with the teacher. At the very least, the fact that she reacts so strongly to the statement as presented in the animation (in its unparsed form) marks this as a breach in the norms of instruction.

With respect to the special role played by biconditional statements in Geometry, Cynthia's comment ("I wouldn't want them to think that it always works to prove it both ways", 446) shows that she recognizes the distinction between a theorem and its converse, and moreover that she expects her students to learn this distinction. Denise echoes this (459) and proposes that both directions of the theorem should be proved. On the other hand, while teachers recognize this as a mathematically worthy topic, they seem to have mixed feelings about devoting full class time to *both* a theorem *and* its converse. Later in the same study group meeting, teachers were asked what they would do after a group of students had presented a proof of the theorem. Despite broad agreement that it was worthwhile to explicitly state the converse, and verify its truth, teachers were in general agreement that proving the converse was not the most worthwhile way to spend the remaining class time:

725	Tina	I may have to show them that the converse is also true.
726	Moderator	So you would go, you would.
727	Tina	Yeah.
728	Moderator	Prove the one direction and then the converse.
729	Tina	I would show them this was an "if and only if". Yeah. You can go, it's biconditional, you could go either way first.
730	Moderator	And what do you do... so they prove, some group proved on the board one direction, and now what? You'd ask them to prove the other direction?
731	Tina	Could it be done the other way? If you knew, if you're given, prove this [the converse.
732	Moderator	And then, what are you going to...
733	Megan	I probably wouldn't let them work on that.
734	Tina	No, we'd [probably just do this as a] class
735	Megan	Yeah, I'd just say, you know
736	Tina	Could we do.

737	Megan	Could we prove this and then, somebody say.
738	Moderator	Could we prove this.
739	Megan	Who knows what the givens would be, and I'd write that up and then, okay, what are we trying to prove? And I'd write that. And then I'd ask questions and write it up, because you already spent a lot of time on this [Tina: mhm]. Sorry, time is a big deal for me on one theorem. I want them to be able to start their work before they leave, I really, that's sort of a big deal to me. [Tina: mhm] Everyday. I don't want to talk right to the end of the period, I want to be able to start their work. Then do a few problems before they walk out of the room. I try and do this everyday.
740	Moderator	So when you said that's, when you tell them to prove the converse or when you go, when you want to prove the converse yourself on the board, do you, I wasn't sure whether you were saying – would you let them state what the converse is? Would you just write the converse on the board?
741	Megan	No, I'll have them say what the converse is, and I'd have them say what the given is have them say what the prove, I might even say "somebody come up to the board and draw a picture for this." So the people are doing the proof, but you are sort of moving it along instead of letting them work at their little table. 'Cause you already did that, you know. Variety is the spice of [inaudible.] [Giggles.]

749	Researcher	So why is that? Is that just a matter of --
750	Megan	It's a time management thing. Cause, like she said, I do think she is right, I would want them to do the converse 'cause I keep talking about what a big deal that is, how powerful it is that they have a theorem that can go both ways, that that's a really powerful thing. That there is not too many things in life that are like that, so--
Excerpt 13. From ITH102605, interval 37.		

Megan's observation (to which Tina concurs) that not as much class time would be devoted to the converse is interesting: it is as if doing so would amount to spending twice as much time on a single topic ("time is a big deal for me on one theorem", 739; "you already did that", 741). What we see is that, despite the fact that Megan gives clear indication that she understands *mathematically* that a theorem and its converse are logically distinct (750), nevertheless from an *instructional* point of view she views them as the same thing. She also points to time restrictions as an argument against giving a full treatment to the converse; if forced to choose between, on the one hand, proving both directions of the biconditional statement in full (a mathematically valuable activity), and on the other hand giving students an opportunity to begin their homework before leaving class for the day (a pedagogically valuable activity), she chooses the latter.

Although both Tina and Megan make reference to the logical independence of a theorem and its converse, at the same time it is not clear from the data how seriously we should take Megan's characterization of a biconditional statement as "a really powerful thing" that is rare and important. It certainly is true from a mathematical perspective; is it also true from a curricular point of view? In a subsequent study group meeting Edwin remarks that in fact most of the theorems of Geometry are biconditional, and cites this as a possible reason why students expect theorems to work both ways:

917	Edwin	.... One of the things like I was saying, my book always does the one, and then always does the other like two days later, and I find they always find the first one is the important one, you know? And then because the book will say then here's the converse, and they'll do just that, they'll assume, wow, the converse was true, we already, they'll, I think, have already assumed that two days before, because they're so used to so many of them are if-and-only-if, that if you switched the order then the other one would be the one that would be important to them.....

940	Researcher	What I, what is interesting in what you all are saying, is that apparently this work of, you know, what is the if part and what is the then part, working on that with the students seems to be a good use of time, right? Or, or isn't it?
941	Denise	Of a biconditional statement?
942	Researcher	Yeah.
943	Denise	Not to me. Working on what's the "if" and what's the "then"…to me it doesn't matter because we all know that it goes both ways. So I say pick one, it doesn't matter. That'll be the first part and then you pick the other part, it doesn't matter, 'cause they're both equally important. That way it'll keep from making them think that one is more important than the other. You know what I'm saying?
Excerpt 14. From ITH041906, interval 36.		

Edwin's comment suggests that many students will mistakenly expect every theorem to be a biconditional, and to conflate a theorem with its converse. If he is correct — if most of the theorems for which students are accountable *are* biconditional

— then one would not expect students to routinely pose the question "Is the converse true?", as such a question would be purely rhetorical, the answer a foregone conclusion. And yet both the discipline and the curriculum make such a question important to ask. For this reason we might expect that teachers hold *themselves* always accountable for making sure the question is raised. Denise (943) disclaims this accountability, apparently on the grounds that ultimately it does not matter which direction is called the "theorem" and which is called the "converse", because the two are equally true and equally important. On the other hand, there is evidence elsewhere that teachers do hold themselves accountable for making sure that the question of the converse is raised. In one study group meeting, teachers collaboratively wrote a script for a story in which a class, having been taught the base angles theorem for isosceles triangles, undertakes a proof of the converse. Teachers joked about the possibility that the question of the converse might be raised by a student:

1272	Moderator	.... And so how would that go? So who would talk first? Anybod–Would, would the teacher start this off?
1273	Karen	You have the students say, "Mr. – whatever he is – is, is the converse of this theorem also true?"
1274	Moderator	Are you kidding or do you really mean that? [laughter]
1275	Esther	She's kidding. Unless it's Alpha. It could be Alpha saying that. [laughs]
1276	Moderator	Alpha could say that?
1277	Esther	Alpha's the only one that would say that.

1287	Moderator	(Okay, so let me --) so how does the teacher start it off? The teacher would ask a question?
1288	Megan	I think the teacher would have – I agree with you [looking at Esther], I don't think any kids (would say,) "Could we prove the converse now?" Even the nerdy kids I have. I think they would, you'd have to say, "Hey, do you think the converse of this is true?"

Excerpt 15. From ESP011006, interval 55.

Esther's joking comment that only Alpha — a character from the animation, one who Esther characterizes as the sort of student who sometimes makes unexpected contributions[20] — would raise the question of the converse, paired with Megan's response that "even the nerdy kids" would be unlikely to do so, suggests that the unlikelihood of such an event is understood by the study group participants as a reflection on the students they teach. That is, teachers would have to (note the modal verbs in 1288) raise the question of the converse because the students would not do so. Absent from this entire conversation is any consideration of the possibility that students might become *habituated* to raising this question, or that teachers might take upon themselves the responsibility for cultivating such a habit of mind.

In the script that the teachers wrote, the students hit an obstacle in their proof and are unable to resolve the obstacle. The moderator asked whether any students might take the failure of the proof as an indicator that the claim was not, in fact, true. From the point of view of mathematical practice, this would be a not unreasonable inference; one of the main themes of Lakatos (1976), for example, is that an analysis of a failed proof can support the construction of a counterexample. In scripting the continuation of the episode, however, teachers doubted this would happen; tellingly, Karen explained that this was unlikely to occur because "(I don't) think that we give them very many places where the converse is actually false" (interval 61, turn 1617).

[20] One interesting characteristic of the study group meetings is the extent to which the characters take on emergent qualities in the discourse of the teachers. The name "Alpha" is a placeholder, used in different stories for students with distinct qualities, but the teachers gradually build up a composite notion of who Alpha "is".

Taken together, the excerpts above create a mixed picture of how teachers relate to the mathematical practice of considering the converse of a mathematical proposition. On the one hand, teachers hold their students accountable for learning (at least abstractly) the difference between a theorem and its converse. A perusal of textbooks bears out the claim that the curriculum ascribes a significant role to this distinction: most geometry textbooks devote a significant portion of an early chapter to propositional logic, with an emphasis on if-then statements, converses, and contrapositives. Students are expected to know that "if" and "only if" mean different things, and that "if and only if" is a somewhat special phenomenon.

On the other hand, based on their statements in the study group, it appears that teachers do not expect their students to become habituated towards asking "Is the converse true?" when they encounter a new theorem. This is partly because so many of the theorems in the Geometry curriculum *are* biconditional. But nor do teachers customarily make it a practice to point out those cases in which the converse of a theorem is false. When both a theorem and its converse are true, teachers in our study groups report that they typically spend less time on the converse than on the original theorem, treating it as an embellishment of the first result rather than a truly independent second result, and that consequently students regard the converse as a less significant piece of knowledge than the original theorem.

Conditional thinking

Above we saw that teachers hold themselves accountable for including in their directions to students that the theorem be parsed into if-then form; they do not seem to expect that students would do so on their own. More generally, teachers seem to be in agreement that writing mathematical propositions in if-then form is something students will only do when expressly directed to do so by their teachers. In an early study group teachers viewed "Chords and Distances", an animation in which a teacher distributes copies of a worksheet with two diagrams of circles and some segments drawn in (Fig. 3.3), and asks them to "conjecture some theorems" about the diagrams.

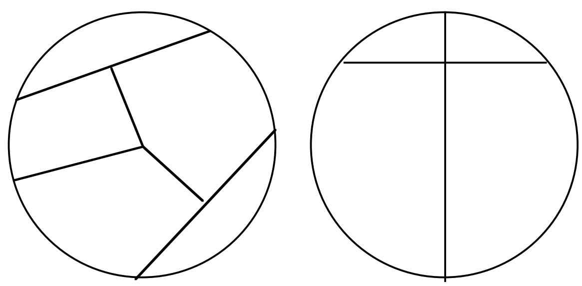


Figure 3.3. The diagrams in the "Chords and Distances" animation.

When asked what students would be likely to do in response to that prompt, many study group participants replied that some students might be likely to make some assumptions about the diagram, but that the better students in the class would know that they would not be justified in doing so, because without any "givens" a student is "dead in the water" (Tina, ITH092805, interval 7, turn 31). Moreover the teachers said that the students would regard both their (unjustified) assumptions, and their (proposed) conclusions, as factual descriptions of the diagram, rather than as conditionally related *possibilities*. If teachers wanted their students to make strategic hypotheses and consider what might be true when those hypotheses hold, the instructions would have to be explicit on that point:

56	Tina	I think have them write it as an if-then statement. I mean assume something, you know, assume something was true what could you prove from that? And, if you even, if you just gave them this and said okay, you have to write an if-then statement, assuming something is true what could you prove? What would you assume is true in this, what would you prove? [Denise: I agree] Then they would//
57	Denise	//If I [had given this to] my students and said write some conditional statements, then they would --
58	Tina	Right.
59	Denise	I agree.
60	Moderator	So you're saying that if you just wrote the "If something then something", they would be able to...
61	Denise	They would come up with a whole bunch of stuff. Might not make sense, but they would come up with a whole bunch of stuff.
62	Megan	You know what? Don't you think that//
63	Res. 1	So the key issue would be to tell them that you want an if-then statement.
64	Tina	Right. But I think, you would have you tell them they could assume something is true. Because with my kids I drill it, our books play tricks on them, and like make something look like a 90 degree and it's not, and they want them to think that it is but it's really not, so it drills in their head that even if you were given this, you can't assume that these two segments are equal or anything's perpendicular until somebody tells you it is. So if you just gave a sheet of paper to them, they would go, we don't know anything because the only thing we know is that's a radii. And those two are chords, and, you know, they would be pretty much dead in the water until you told them they could assume something was true.
65	Res. 1	If you told them, if you gave them some information, what would be the probability that they would come up with an if-then statement?
66	Tina	What do you mean, like, give them that they were perpendicular?
67	Res. 1	Well, perpendicular and the chords are equal.
68	Tina	Then I think they would come up with conclusions to it, because that would be their if part, if this is true then... I don't think they'd write it as an if-then statement, but they would write their conclusion, what they would come up with.
69	Cooney	Just write the conclusion.
70	Tina	Right. (3 sec pause) I don't think they'd put it as an if-then.
71	Res. 2	But in the situation that is given there, because no information is there, you don't think they would conjecture that anything is true about this, that figure.
72	Tina	Not unless I tell them they could. Mine are pretty well disciplined.

Excerpt 16. From ITH092805, interval 8.

In interval 8, Tina seems to be making several related points (to which the other teachers concur, elsewhere in the transcript). First, students are taught throughout the year not to make assumptions about the objects represented in diagrams without some explicit sanction to do so ("givens"); students do not choose their own givens, and without givens they are “dead in the water”. If the nature of a task is such that students *are* expected to make assumptions and see what conclusions follow from those assumptions, they need to be given explicit instructions to do so, and those instructions must contain a reference to “if-then” form or “conditional statements” in order to signal to students what they are to do. Even when students *do* have givens to work with, and are asked to make conjectures, the conjectures are most likely to be expressed as factual statements (“just write their conclusion”, 68), rather than as a conditional relationship. All of this together adds up to an argument that teachers do not expect their students to think conditionally or counterfactually about geometric objects represented in diagrams; they take as shared the notion that students are supposed to describe or prove what *is* true about certain objects, rather than what *could* be true *if* certain conditions were met.

These findings are consistent with earlier research on students’ interactions with diagrams (Herbst 2004) and on the role of the diagrammatic register (Weiss & Herbst 2007) in mediating students’ understanding of geometric objects. Those studies have argued that the conventions of the diagrammatic register are such that a diagram is taken to be a more-or-less accurate representation of a particular geometric object. According to these conventions, not everything true of the diagrammatic representation can be ascribed to the object represented; in particular, measurements made on the diagram are to be regarded with suspicion. Still, to the extent that the diagram is understood to represent an *actual object*, any claim that can be made about the object’s metric properties are either true or false. For example, any given quadrilateral either is a rectangle, or it is not; and although one may not know which is the case, students nevertheless know that it is surely one or the other. If there are enough givens to entitle the conclusion that it is a rectangle, then this may be asserted, and otherwise one must refrain from making any such assertions. This way of speaking of diagrammatic objects militates against making claims that are contingent on unverifiable possibilities, such as "*If* this were a rectangle, then…"

All of the above, however, is predicated on the assumption that students encounter a particular diagram as a representation of a particular object, rather than a generic representative of an abstractly-defined class of objects. Weiss & Herbst (2007) show that the latter is not unheard of in the Geometry classroom, but that its role tends to be restricted to the instructional situation of "installing theorems" (Herbst, Nachlieli & Chazan, submitted; Herbst & Miyakawa, 2008; Miyakawa & Herbst 2007a, 2007b). In this context, a diagram is typically accompanied by a verbal formulation of the abstract properties that the diagram supposedly represents; this verbal formulation marks the diagram as a generic object only, rather than a particular object. In the episode under discussion in the previous excerpt, no such markers of genericity were present in the task as posed by the teacher to the student. For this reason it is not surprising that study group participants were in broad agreement that students would be unlikely to make any claims about the diagram in a form that labels its properties as contingent possibilities, rather than actualities — unlikely, that is, unless they were to be explicitly directed to do so by the teacher.

Generalization and Specialization

In Chapter 2, I indicated how modifications to the conclusion and hypothesis of a (settled) conditional statement can produce generalizations and specializations. I illustrated this by showing how the proposition "*If ABCD is a parallelogram, then its angle bisectors form a rectangle*" can give rise to a range of new questions by iteratively generalizing and specializing. That example was chosen, in part, because a number of the ThEMaT study group meetings and GRIP focus groups were focused on some variation of the "angle bisector problem:" What can be said about the angle bisectors of a quadrilateral? The animations "The Square," "The Kite," and "The Parallelogram" (discussed above) all deal with one or more aspects of this problem. Each of those animations exemplifies how looking at a *special case* of a problem can help to explore a complex problem space. In light of this, it is interesting to see how study group participants responded to that aspect of those animations.

In the first study group meeting in which participants watched "The Square", participant Glen noted that the teacher had posed the problem in a general form ("What can one say about the angle bisectors of a quadrilateral?") and reacted negatively to Alpha's investigation of the special case of a square:

173	Glen	I would have stopped at some point fairly early on and said you know, a square is not, you know we want to talk about a quadrilateral. And this is a very specialized one, can we get back to the more general. I really stress to my kids, if I tell them to draw a triangle, I don't want them to draw an isosceles, because it may lead them to facts that look to be true but really aren't necessarily true in our problem. So I would have tried–I would have left his square on the board and tried to have gotten somebody else involved so we could have gone from just a plain old ugly quadrilateral and then eventually like you said worked down to the more specifics, the parallelogram, the rectangle, the square—
175	Greg	See I wouldn't even have let him come up to the board, cause I knew he was square. He said it right away. I wouldn't have had him come up, I would have said, well let's save that idea for later, let's deal with—
176	Mara	And you saw that they were open for 5 minutes, walking around the room, you already know—
178	Glen	So then you know, I can see both sides of our discussion here, you're going to stop this before it even happens, I may have let it happen and then kind of you move, show them the relationship--
Excerpt 17. From ThEMaT081905, interval 13.		

It is not completely clear on what grounds Greg and Glen object to Alpha's investigation of a special case. In part it appears to be based on the authority of the teacher: students are supposed to answer the question they are asked, not change it to a different one. On the other hand it also appears to be a matter of cultivating good intellectual habits: if students are supposed to be considering a general object, but use representations that have unrecognized special qualities, they may be misled in mistaking phenomena that are only true in certain cases as being true in general.

Glen's and Greg's critiques of Alpha seem to be based at least implicitly on the idea that Alpha's investigation of the square was not a deliberate and purposeful way of making progress in an open and uncharted problem space[21], but rather an accidental misinterpretation of the problem or a misguided attempt to make things easier for himself. Elsewhere in the corpus, in connection with a story in which students study the concurrency of medians in a triangle, we find evidence that teachers can regard a move such as Alpha's as both problematic and strategic:

134	Greg	I was surprised that more students didn't realize that by choosing an equilateral triangle or an isosceles triangle that that was a problem. I thought they would have picked up that right away.
135	Cynthia	Special case.
136	Greg	Right.
137	Cynthia	Right.
138	Researcher	Yeah actually one student says why is that a problem?
139	Cynthia	Yeah, maybe he was the only one though.
140	Researcher	So you think that students should know that that is problematic?

[21] It is perhaps worth noting here that, notwithstanding Glen's claim (173) that he would urge Alpha to "get back to the general", at no point in any of our study group meetings that discussed the angle bisector problem did any participant mention *any* properties of the general case as being worth pursuing in class. This was so even after we distributed to all participants a short document containing a more-or-less complete mathematical treatment of both the general case and a number of special cases.

141	Greg	I was just surprised that it either took them so long to figure that as a problem out or even one set of why is that a problem? I think even after you explained why it might have been a problem.
142	Researcher	Do the others share that sentiment, like the students should know that it's problematic to choose an equilateral triangle?
143	Denise	I don't know, because my students...every time I tell them that's a special case, that's not gonna work all the time, they still try to find the easy way out and pick these special cases that make the problem easier. So I don't know. And even why I say no that's a problem, that's a special case, I don't care if I say it every day for 15 days, someone's gonna say why is that a problem? Because they want an easy way out, they want–I mean–so I don't know. That didn't surprise me at all.
144	Researcher	In here, like, just to play devil's advocate, Megan had also said you will draw your own triangle. So couldn't they just say well this is my triangle? Or you would, but you would still expect them to choose a general one right?
145	Greg	Well I would think that they would first try an equilateral or an isosceles, because that's the easiest one to do. And then after that works for that one, let's see, can I do an obtuse or something else that--
146	Moderator	So they'd realize they're taking a special case, then they need to take a general case.
147	Greg	And I think a lot of them do that right away, take a special case first.
148	Moderator	Take special case first, but knowing that it's just a special case...then they--
149	Greg	I did. I felt the equilateral triangle would probably be the easiest one to see. And then once you see that then maybe let's try another one to see why did it work for that one.
Excerpt 18. From ITH011806, interval 10.		

This excerpt contains a complex set of evaluative stances regarding students who look at a special case when posed with a general question. Greg characterizes such a move as problematic, presumably for the same reasons cited above. Denise regards the move as an act of laziness ("they want an easy way out", 143). To this Greg responds that *knowingly* taking a special case to simplify a problem may be a very strategic move (149), *provided* that subsequently one returns to the general case. That the move be purposeful (rather than accidental and unrecognized) is an important one.

Denise's contention that specialization is something lazy students do finds an interesting counterpoint in a brief discussion from another study group meeting. In that session, participants considered "Conjectures about Quadrilaterals," a story in which a teacher had students investigate the quadrilateral formed by the midpoints of a rectangle, with the goal of discovering and proving that it is always a rhombus. Kappa, however, considers a *general* quadrilateral, and discovers that the "midpoint quadrilateral" (i.e., its dual) is always a parallelogram. Participants were quick to label Kappa as an advanced student:

783	Researcher	And, and you were saying, so, you were saying that going after Kappa's idea is like teaching to the top of the class.
784	Carl	Mm-hmm.
785	Researcher	Because Kappa didn't care to look at the more trivial cases, was that what you said? Like //
786	Carl	// Cause Kappa's moving to the general case.
787	Researcher	Mm-hmm. And, so, I //
788	Carl	// And that I think is-- all by itself illustrates teaching to a different level. Is saying we can make conjectures, we can, we can perhaps find proof in specific cases and I think that that's something that everyone in my class should be able to do. Take a square and do it. I mean, what do you get? You get another square. And I feel like, everyone's like, "Oh, okay, that's fine". You know, start with a rectangle, what do you get? And I feel like that's something I would expect them all to get in short order.
789	Researcher	But if the teacher follows Kappa's idea the teacher is sort of ignoring that bigger segment of the population, you're saying.
790	Carl	Mm-hmm.
791	Researcher	Are you saying that?
792	Carl	Yeah.
793	Melissa	I think so.
794	Carl	And pulling everyone up.
795	Researcher	Mm-hmm.
796	Carl	Perhaps.
797	Melissa	But maybe they're not ready to go into another, you know, [Carl: I mean, perhaps.] you might lose them more, and they're just like, "Oh god, there goes Kappa again."
798	Carl	Mm-hmm.
799	Melissa	Uh-huh. "I didn't understand the first part and now you're onto the (.) end."
Excerpt 19. From ESP101105, interval 22.		

These preceding excerpts, taken together, give evidence that teachers perceive a correspondence between the dialectical pair of mathematical practices "generalize / specialize" and the classification of students as "strong / weak". That is: Alpha, who specialized when he was not supposed to, is a weak or lazy student (Excerpt 17); Kappa, who did not follow the teacher's instructions and instead considered a more general case, is a strong student.

When this finding is set alongside the previously-noted obligation to teach all students (not just "the top of the class") we find an emergent argument for teaching to favor investigation of special cases (as more tractable for the majority of students) over a propensity towards generalization. And in fact when considering how the angle bisector problem might be deployed in classrooms, teachers indicated a preference for organizing the task as one of looking at special cases. We find, for example, the same idea expressed by two different teachers in sessions held more than a year apart:

231	Peter	I think if, if you had the software, you could have ten different quadrilaterals that you could have them, on each one of those, doing, do the angle bisectors and look and see what they got.... if you start with this, you get this, they weren't always the same, you didn't get the same thing at every singe shape. So then, if all the kids were working in pairs or whatever, then you could bring 'em all back together and collaborate. Every time we started with this, what did our angle bisectors produce? Every time we started with one of these, what did our angle bisectors produce?
Excerpt 20. From TWP020805.		

682	Tina	Well and see I've done something similar to this, and what I did is I actually drew up different quadrilaterals, like I drew like four rectangles on a page. I drew four squares on a page. I drew four – you know. I made, I made up my own sheets and then I randomly ran off however many I needed for the class and shuffled 'em, and then I handed out and said, "Okay, here you go. Everybody's got a sheet. Everybody's got a different shape. Draw the angle bisectors and come up with a conjecture you can make." And then what I do is say, "Find somebody who has the same shapes as you and see if you came up with the same conjec–" and you gotta get 'em moving that way a little bit, a little brain break and everything. And then I have them sit down and talk together and then we kind of post up one of the sheets and say, "Okay, who had this sheet? What conjectures did you come up with?" And then that way you kinda address 'em–and when you have fifty-five minutes it's a great way to kinda get through all of them quicker....
Excerpt 21. From TMW111506, interval 24.		

Although Peter and Tina differ significantly in the kinds of material resources they envision students using — Peter favoring dynamic geometry software, Tina describing the use of photocopied worksheets — the tasks they describe are similar insofar as they both call on students to discover the general case (so to speak) through an inventory of a number of special cases.

In summary, what we find is that the dialectical opposition between generalization and specialization finds its expression in the classroom in the form of a tension that teachers must negotiate. On the one hand, the participants in the study groups and focus groups say that students, when asked to draw general objects, tend to (unconsciously) draw special cases, and as a consequence will erroneously ascribe properties of the special case to the more general one. For this reason teachers need to resist the students' instinct to specialize, and remind the class to stay in the general. Teachers also recognize that looking at more general examples can help to differentiate between distinct objects that happen to coincide for some special case (e.g., the diagonals and angle bisectors of a quadrilateral; the medians and altitudes of a triangle; etc.) On the other hand they also perceive the general case as more difficult than the special cases, and not really suitable for the majority of students; this perception pushes them in the direction of providing scaffolding for students' work, by asking them to look at specific special cases rather than consider problems in their full generality. No one, not even the teacher, is held responsible for considering questions in as much generality as possible, for letting the question "Can this result be generalized?" take the inquiry as far as it can go.

Utility & Abstraction

As discussed in Chapter 2, another dialectical pair of dispositions (categories of perception and appreciation) that appear to be part of the mathematical sensibility is *Utility* and *Abstraction*. That is, at certain times and in certain contexts, mathematicians value questions that have applicability in a (not purely mathematical) domain (e.g. physics, biochemistry, economics), or take their inspiration from those domains; while at other times and in other contexts mathematicians may openly disparage "real-world relevance" as a criterion on which to judge the merits of a mathematical result and instead value mathematics precisely for its removal from such practical applications. As with the other dialectical pairs I discuss, these two oppositional categories are not in practice mutually exclusive; from the theoretical considerations discussed in Chapter 2, we would expect that most mathematicians might hold some allegiance to both of these positions, with one or the other coming to the foreground at certain moments.

In this section we investigate whether either or both these categories has salience for teachers of high school geometry: that is, in commenting on (theirs and others') practice, do teachers refer to the utilitarian or abstract qualities of things, and adduce those qualities as grounds on which to form an appraisal of a possible action. The question is an important one, especially when one considers the history of the high school geometry curriculum in the United States. As has been shown (González & Herbst 2006), for much of the past century the presence of geometry in the high school curriculum has been warranted on at least two competing grounds. On the one hand, from ancient times geometry has had its origins in the solving of practical problems (as attested by the etymology of the word "geometry" as "earth measurement"). With regards to the school geometry curriculum, as early as 1909 the *Committee of Fifteen on the Geometry Syllabus* began calling for an emphasis on application of geometry to “real world situations such as designing architectural elements, surveying, and sailing” (González & Herbst, p. 11).

On the other hand, as González & Herbst (2006) also report, from the end of the 19th century onward geometry was also touted as valuable for its capacity to cultivate students’ ability to think logically. In the 1893 *Report from the Mathematics Conference of the Committee of Ten*, “geometry, unlike algebra, was seen as an introduction to students of the ‘art of rigorous demonstration’.” (González & Herbst, p. 12, quoting Newcomb et al, 1893). From this point of view the most important aspect of the geometry course was not its *utility* in real-world problem-solving but rather its structure as an axiomatic-deductive body of knowledge (p. 8). This position can be taken as an expression of the *abstractness* value, and it too has found repeated expression throughout the last century, perhaps most notably in the “new math” reforms of the School Mathematics Study Group (Curtis, Daus & Walker, 1961).

González & Herbst (2006) identify four modal arguments that characterized the discourse around the nature and purpose of the geometry course during the 20th century: a *formal* argument (geometry teaches deductive reasoning), a *utilitarian* argument (geometry is essential knowledge in the workforce), a *mathematical* argument (geometry offers students the opportunity to experience the activity of mathematicians), and an *intuitive* argument (geometry provides a language for describing the world). Each of these four arguments finds its expression in recent policy documents such as *Principles and Standards for School Mathematics* (NCTM 2000). For the present purpose, the significance of these four arguments rests in the fact that two of them are related to the disposition towards *utility* and the other two are related to *abstraction.* As such they provide evidence that, at least as far as the *intended* curriculum is concerned, both of those dispositions have a presence in the high school geometry course.

The study group and focus group data bears these observations out, although the picture is a complex one. In the transcripts we often find teachers mimicking a stereotypical student asking "When are we ever going to use this in real life?" in a mocking tone. At the final ThEMaT study group of the first year (ESP062806), Megan was quick to concede that much of the substance of the course she teaches will have no direct application in students' lives: "And I say to kids all the time, 'I'm not – No one's gonna ask you to prove triangles are congruent in your life. I'm just going to admit that right now. No one's ever going to ask you to do that.'" Rather, for Megan the value of the course is located in its ability to teach students to think and learn independently: "I have faith in learning. That's really what I think I'm teaching. Faith in learning, that you can learn anything and you don't really need me, I'm just sort of showing you some tools but once I do that, you can just pick something you want to learn and you can learn it."

In a similar vein Megan frequently spoke of "proof" as something that students could apply to real-world "problems". In another session[22] she illustrates this with an anecdote from her classroom:

> .... But we've been talking about how you can use proofs in real life. Like I said to one girl, tell me something your mom wouldn't let you do, and she says "Wear a bikini top to school." So I was like, okay, let's go with that, so what would be your logical arguments against that? And she said,

[22] SG-ITH092805.

> “Everyone at school’s wearing a bikini top.” Which is, she was just using this as a far-out example, no one would wear a bikini top, but I said, this is perfect, that is not a logical argument, you’re giving something that is not a logical argument. Can somebody else give one? And this girl says, “I’m so hot at school I’m dehydrated and I can’t think straight”, something like that.... So [inaudible] more logical arguments, yeah, we’re trying to have fun with the whole, that this stuff is useful. Building a logical argument, I said, What are the givens you know about your mother? And she was talking about her mother, and I was saying, well you want to build your argument around your given things. So yeah we’re still like into that, we’re just starting proofs. (ESP092805, interval 26, turn 449)

Megan is not alone in disavowing the value of “proving triangles congruent” as useless, while simultaneously valuing the role of proof in teaching students to build persuasive arguments. In one typical exchange Karen and Carl agreed that geometry was useful not so much for the content of the mathematics, but rather for its capacity to teach students the need to provide support for claims:

63	Karen	I think that what they’re seeing – I mean I see proofs as like stories, where you’re bringing in the elements just as you need them and, you know, that they make a, they have a plot. And so, and I think it’s very hard for them to see any use in doing that with angles. I mean we really start to get the “when am I ever going to use this, you know why would I ever ever want to prove that two triangles are congruent?” And, um, and I’m trying to focus on that proof is a really important part of life. And that, that as you get to the idea of the basic structure of proof, and how to set it up so that you’ve got your givens and you make some reasonable arguments, that that’s a helpful thing. And, so I think it’s a way of connecting it into what they’re doing so that, so that they-so many people end up saying, I hear from people that are grown, well you know, I went through geometry, I don’t remember any of it, I didn’t learn anything, it was totally useless. And it was sort of like trying to get them to see it as part of a how we think.
64	Carl	I remember when–do you remember when we put together that proof project like a million years ago? Lucille and I, she had this proof project and the goal was, I think, was to sort of bring in these ideas of, everyday life proofs into the whole idea of doing like formal proofs. And I guess what–I do use it, but I rely on it less because I feel like I’ve hit–I’ve missed more than I’ve hit when I tried to do this project. And I think that my purpose is a pretty clear one, and that is to just show students how they do inductive and deductive thinking all day long, and they don’t pay attention to when they’re doing it. Um, but I do think that having those examples of those different kinds of thinking, those

		different styles of thinking in particular are really useful to them because I think that they recognize that I'm not just teaching them how to do formal proof, but I'm trying to teach them how to think in certain ways. Um, like Karen was saying, I think it is part of life and I do think kids often think of geometry sort of as less important perhaps, or less useful or whatever. But what ends up happening is kids end up saying to each other, like you know that's really one that you remember, that's the one you actually you know walk away and you actually–that's the math you're going to be doing. You're not going to do algebra when you're walking down the street, you're not going to be doing, you know, calculus. But you're going to be remembering relationships and remembering different kinds of thinking. But I can't say that I've come up with a successful lesson plan that has really hit home, kids say yeah I get that. I think that it's a little too abstract.
65	Researcher	So is the everyday logic helping them understand the geometric logic--
66	Carl	The need for it.
67	Researcher	--or is the geometric logic helping them understand the everyday logic?
68	Karen	I think it's a totally kind of new idea for them. I mean the English teachers yell at them all the time, put in evidence. And I think they're, and they watch all the law shows and you know, and they have a sense of what happens in the jury but I don't think they have a sense for what evidence is, or how to support a statement. And that's a crucial idea, that they get this way of understanding that you just can't make a statement without supporting it and what support actually looks like rather than you know just, "Well I think it's true".

Excerpt 22. From ESP110105, interval 8.

Karen and Carl's lengthy turns in the above excerpt are loaded with terms and phrases expressing a dichotomy between, on the one hand, the abstract *content* of geometric proofs, and on the other hand the utilitarian value of *the art of proving*. Karen distinguishes between proving "that two triangles are congruent" (which students see no use for) and proof as a "really important part of life". Carl speaks of "formal proofs", "algebra" and "calculus" as "the math you're not going to do when you're walking down the street", in contrast with "everyday life proofs" which students use "all day long".

But overlapping and coextensive with this dichotomy is a parallel one, in which deductive and inductive reasoning are opposed to each other and aligned, respectively, with formal (=useless) and everyday proofs. This is found explicitly in Carl's turn 64, and implicitly in Karen's turn 68, in which writing a mathematical proof is compared to writing an essay for an English class and to an argument before a jury in a court of law — both forms of argument that rely at last in part on amassing evidence and on rhetoric, rather than solely on rigorous deduction from axioms. Thus the participants' commentary on the utilitarian value of learning to prove contains an unrecognized contradiction: the value of learning deductive argument is that argumentation is valuable is many real-world contexts, but this is illustrated with examples of argumentation that are not purely deductive, calling into question the claim that deductive argument is really worth learning.

The above excerpts locate the value of the geometry course in its capacity to teach "logical reasoning", which is in turn warranted for its value in the "real world". On the one hand, teachers reject the "stuff" of geometry — its actual mathematics content — as not useful in the real world. On the other hand the need to justify the course on real-world grounds cannot be entirely neglected, if for no other reason than that students will continue to challenge the teacher with "Why do I need to know this?" Consequently the formal argument is turned around and made into a utilitarian one. We see here a striking accommodation between the opposing poles of utility and abstraction.

Teachers are not entirely insensitive to the value that the substance of geometry (its propositional content) has for real-world problem-solving contexts. In a discussion from a second-year study group teachers discuss the role that mathematics plays in blue-collar professions:

847	Researcher	Now you realize that some, some of those kids that, that – The people that come to your house and fix the, you know, that work on the pipes and do all the manual things are usually, have been the people that had difficulty with mathematics in school. But they have the capacity to reason geometrically at least in ways that, you know, are enviable, you know?

848	Melanie	Mm-hmm.
849	Researcher	And, you know --
850	Cynthia	And they don't see the connection between the two.
851	Researcher	We don't see the connection between the two! [group laughter]
852	Megan	No, they don't --
853	Cynthia	No, they don't. They don't see that, that like my cousin is a landlord he's got like, you know, thirty-some units which is a, a, quite a bit. And he doesn't see the math behind when he buys carpeting, paint and all that kind of stuff, of calculating how much he needs as being part of math and --
854	Researcher	Right.
855	Cynthia	-- you know I tell him, "John, you're much better than you think you are (because you're using) every day a real-life math, and he doesn't see that (connection to,) to what I teach or anything like that, and so it's just kinda interesting how --
856	Megan	(You're using math every day.) (He's in denial.)
Excerpt 23. From TMT101006, interval 26.		

Similar observations were made in other sessions by Greg and Karen (TMW032107, interval 34, turns 660-662), and by Robin (TWP050306, turn 404). In all of these instances teachers point to the practical value of geometry for workers in skilled trades (carpenters, plumbers, electricians, auto mechanics). At the same time they lament the fact that the usefulness of geometry goes generally unrecognized in our society, to such a degree that even those who use geometry routinely see little connection between what they learned in school and the mathematics they use in practice. This may speak to a sentiment among the study and focus group participants that the *utility* disposition is generally undervalued by a curriculum that places greater emphasis on *abstraction.* But at the same time, and somewhat paradoxically, the teachers themselves give virtually no evidence throughout the entire corpus of data that they themselves do anything in their practice to redress this perceived imbalance. On the contrary, whenever teachers discuss engaging their students in projects that would engage their students in real-world applications of geometry, invariably the activity is characterized as a fun change-of-pace from the normal classroom routine, one that is done to boost morale but for which there is little student accountability. For example, at the beginning of one study group meeting (ITH121505, interval 0), while waiting for other teachers to arrive, Tina discussed an annual activity in which her students design and build a putt-putt golf course. One could certainly regard the project of designing, building, and testing a putt-putt golf course as a learning experience, in at least two senses: a student could deepen his or her understanding of certain geometric content, and also gain a better appreciation for the utility of geometry in real life. But throughout her recount, Tina describes the activity as fun ("we had a blast", turn 195) rather than "educational" or "useful". Cynthia, who listens attentively to Tina's description of the activity and who asks her questions about the logistics of the event, concurs that "it'd be fun to do" but likewise says nothing about the project's educational value. Moreover, in her description of the annual event Tina states that it is customarily scheduled just before a vacation ("the day before Thanksgiving", turn 193, or "at Christmas", turn 203), further creating the sense that such events are an enjoyable diversion away from the normal activity of the classroom, rather than a central element of what is to be taught. As teachers describe such activities, they signal with their language choices at their goal is not for students to learn any specific mathematics, or even to gain an appreciation of the utility of geometry in real-world contexts, but simply for students to have "a blast".

In summary, we find that the twin dispositions of utility and abstraction have salience for the teacher — that is, teachers speak both of the usefulness of geometry, and of its disconnect from reality — but (as with the other dialectical pairs described above) that salience manifests in the teachers' practical rationality as a source of tension. Teachers feel obliged to point out the utilitarian nature of geometry, with its connection to real-world problem-solving, but neither they nor their students (nor former students who become users of geometry in the context of skilled trades) take that connection very seriously. On the other hand teachers seem somewhat embarrassed and apologetic about the abstract nature of the content, rejecting the content of proof as useless ("Nobody's ever going to ask you to prove two triangles congruent") while simultaneously justifying the teaching of proof by an appeal to real-world situations in which argument-building is important. This latter appeal deconstructs itself, in that the real-world contexts proffered are quite different in nature from the kind of arguments (logical deduction from assumption to conclusions) that typify proofs in geometry.

Surprise and Confirmation

The third dialectical pair discussed in Chapter 2 was *surprise / confirmation*. There it was shown that mathematicians place a high value both on results that run contrary to their expectations, and on results that confirm long-standing but unproven suspicions. This pair of dispositions is, obviously, closely related to the general theme of mathematicians as wonderers who work in areas of doubt and uncertainty; and, insofar as we have already seen that the high school geometry classroom provides only limited exposure for students to such areas (see earlier in the chapter), we might expect to find little recognition of these dispositions in the corpus of study and focus groups.

And, in fact, a search of the data found very few instances in which teachers said anything that seems directly related to these two dispositions. In part this may be explained by the central and particular role that visual intuition plays in high school geometry: The theorems taught in geometry are never really surprising, because they are statements of facts that are visually obvious. On the other hand the fact that the truth is never in serious doubt means that confirmatory results are not particularly important, either.

In one lengthy but very revealing conversation, teachers expressly acknowledged that exploration, and in particular the work of finding counterexamples to a conjecture, tends to disappear from their classrooms at roughly the same time that proof begins to take a prominent role. The following discussion excerpts are taken from two adjacent intervals in study group meeting ESP091305, in which teachers viewed "The Kite", an animation that shows students engaged in the question of whether the angle bisectors of a kite meet at a point (and, if so, where that point is located). In that context Carl noted that conjecturing and proof are allocated disjoint times within the school year:

459	Carl	..... And sometimes I feel like we talk about conjecturing and counter-examples at one point in the semester and then we're like okay, we're not gonna talk about this anymore, now we're talking about proof. That there's some sort of stopping place in there, I don't know if you guys do that too, but I, I think there's some conjecturing that goes on at some point but then it's like, okay, let's just really prove this [stuff]. I don't know if there's enough ambiguity left. You know, I don't know if, I think maybe I'm not being clear. Sometimes I just feel like everything we do is provable after a certain point in the semester, and we don't spend enough time looking at problems where things aren't provable, I don't think.
460	Karen	Well we don't spend very much time making conjectures because we have so many proofs to do.
461	Carl	Right, right. Maybe that's what //
462	Karen	// There's less time exploring, and then at the end we're stopping the proofs altogether //
463	Esther	// Yeah, we more or less tell the kids to show that this works. We don't do as much of asking them what they think works and whether or not they can verify it or prove it or explain it. We more tell them, this works, explain why, more. We don't, we don't do enough I don't think of.
464	Researcher	So why is that, that we don't do that?
465	Esther	Um...
466	Carl	Is the question that [the moderator] asked at the top of the hour foreign in your classroom at, like, the point it would be appropriate? Like, we can't start talking about kites until like December or January, right? Would you ever

		ask a question like that in December or January? Like, I'm, it looks so foreign to me. Like it's a great question. I wish I had time for it and everything. But, like, I don't feel like I ever would ask a question like that. Explore, figure out, and then maybe we'll prove something.
467	Lynne	Well, think of how much knowledge that they have to have already, when we're talking about //
468	Carl	// This is what I'm saying.
469	Lynne	// Side-side-side, so we have to have all the triangles congruence facts //
470	Karen	// We sure can't do it like next week or anything //
471	Carl	// No.
472		[3 sec pause]
473	Karen	Well, but, you know //
474	Esther	// It does take them to have a pretty strong basis before they do. //
475	Karen	// But I, I was like, I liked that question. And I was thinking //
476	Carl	// Me too. //
Excerpt 24. From ESP091305, interval 51.		

Within this first excerpt teachers affirm that conjecturing (as an activity) gives way to proof when the latter enters the curriculum late in the first semester. Consequently there is no natural time for a teacher to pose a question like that considered by the students in the animation "The Kite." In considering why this is so, teachers offer a variety of explanations for this state of affairs. Karen locates this as a result of the crowded curriculum: once students become accountable for learning proof (late in the first semester), "we have so many proofs to do" (460) that there is no time left for conjecturing. Lynne and Esther approach the problem from the other side: the reason why a problem like this could not be taken on in the early part of the year (when investigating conjectures is still a common activity) is that students do not yet have the prerequisite knowledge for such a task (467, 469, 474).

In something of a rebuttal to this, the following interval begins with a long turn in which Karen considers whether the problem could indeed be explored early in the year, and situates it within the unit on angle bisectors:

479	Karen	No, but we could it. We could do explorations. In the same, you know, like, put it in as part of teaching angle bisectors. I think it would make more sense. I think we have, well, what I feel is, I just have way too much content, too many things to cover and not enough time to do it. And then the other part is, I know students who seem to be trained to be stubbornly stupid. You ask them a question and they sit there [pause] and wait. And if they can out-wait me, then I'll tell them the answer. So I'm continuing on the training, you know another year of being stubbornly stupid, and eventually the teacher will tell you the answer and so all you have to do is wait and someone will tell you the answer. And it just seems like they're trained for it, that they're getting better and better training. To wait.

483	Carl	I feel like I'm driven by, the clock is ticking.
484	Lynne	But I think you kind of change your mindset too. Your mindset switches to something different and your motivations are forming coherent arguments and things like that, so you're not thinking, "I need to have my students make conjectures anymore." [Carl: And explore...] But, I think it's a great time when you're working with the shapes to talk about, okay, see if you can figure out their properties, what properties do you think are properties of the square?

491	James	What I, I think what Carl's saying is not so much the conjecturing because you can kind of introduce some of the proofs by getting, by trying to elicit some conjectures about it, but the counter-example. Where all of a sudden now, you know, to be able to just say, okay, I can think of a counter-example and, you know, now I don't have to do the proof. It would be nice to, if you're going along, if you're doing some proof [Carl: Yeah, it is //] to be able to have something //
492	Carl	// Yeah, it isn't that we don't conjecture, James's right on, it's not like we're not conjecturing in December, it's that we're not conjecturing falsely in December. Maybe that's what I'm saying.
493	James	Yeah.
494	Moderator	All of the conjectures were right.
495	Carl	All of our conjectures are right. Now we've just gotta write it down //
496	James	// But even the //
497	Researcher	// That's quite a coincidence, isn't it?
498		[Laughter]
499	Carl	Yeah, it sure is!
Excerpt 25. From ESP091305, interval 52.		

As in the previous excerpt, teachers offer a variety of explanations for the disconnect between conjecturing and proving. Karen (479) attributes it both to the overcrowding of the curriculum ("I just have way too much content"), echoing her comment from the previous interval, but also to her students' poor intellectual habits:

they are “stubbornly stupid”, unwilling to expend effort on a challenging problem, and able to outwait the teacher who must (because of the dictates of time and coverage) eventually capitulate and “tell them the answer”. Carl agrees that time compels the teacher to eventually resolve all questions (483). Lynne affirms that conjecturing and proving belong to different times of the year, and that once the former has been completed, it is no longer an important part of a teacher’s planning.

Carl’s amused recognition (to which James and Karen concur) that in December “all of our conjectures are right,” together with the observation that students feel little need to prove theorems because they are obvious, and Lynne’s description of students’ conjecturing the characteristics of special quadrilaterals (all of which are unsurprising and supported by visual intuition), combine to create a portrait of geometry as a subject in which nothing is ever surprising, and thus one in which neither confirmatory results nor unexpected counterexamples have much value — the former because they are trivial and ubiquitous, the latter because they are nonexistent. Thus it appears that neither “surprise” nor “confirmation” form part of the practical rationality of geometry teachers.

Now it is not at all the case that everything true in geometry is visually obvious; there *are* surprising results in Euclidean geometry. Napoleon’s Theorem, for example, states that if one begins with any triangle, constructs three equilateral triangles on its sides, and joins the centers of those triangles, then the resulting triangle will always be equilateral — a fact that many people find unexpected when they first encounter it. Pappus’s theorem about the segments joining six points chosen on two arbitrary lines (see chapter 2) is equally surprising. For a more recent example, Morley’s Trisection Theorem — which asserts that the *angle trisectors* of an arbitrary triangle meet at the vertices of an equilateral triangle (see Fig 3.4) — is striking not only for its content but also for the fact that it was not discovered until 1899, more than 2000 years after Euclid (Oakley & Baker, 1978). More examples could be provided *ad lib*. From another perspective, there are a number of famous *negative results* of Euclidean geometry that could be opportunities for students to encounter surprise: for example, the impossibility of squaring a circle or trisecting an angle with only compass and straightedge.

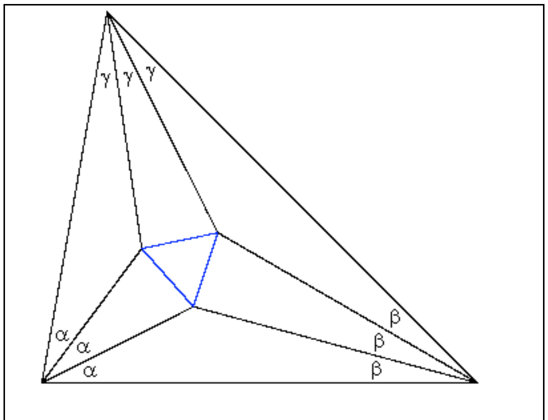


Fig. 3.4. Morley's Trisection Theorem.

Any of these phenomena could in principle be made an opportunity for students to encounter the values of surprise or confirmation. And yet, while some of them may be found in older textbooks, for the most part they are absent from the conventional geometry curriculum. They are certainly absent from the study group data — at no point in the corpus do teachers raise these or any other similar results as topics of conversation.

I raise these examples not because I wish to argue that any or all of them ought to be included in the prescribed geometry curriculum, but simply to illustrate my claim that there *are* surprising facts in Euclidean geometry; the fact that these and others like them are generally absent from the curriculum is thus evidence that the twin dispositions of surprise and confirmation play little role in the subject as it is conventionally taught. The study group data presented and discussed above suggests some reasons why these categories are mostly invisible to teachers. One possibility is to point to those institutional forces (textbook publishers, state and federal standards, school boards and department heads) that shape the curriculum. Teachers do not generally have ultimate authority to decide what to teach: they are bound by their professional obligations to follow curricula that are authored by others, and the time constraints under which they labor tend to crowd out any material that is not mandatory. (In this connection, see Karen's references to the overcrowded curriculum, as well as James's references to the STMP exam, earlier in this chapter.) Another possibility is to count this as a result of limits in teacher knowledge: obviously teachers will not teach content that they do not know themselves, and as none of the phenomena raised above are standard elements of a "geometry for teachers" course there is no reason to expect teachers to know this material. A third possibility is to observe that each of the mathematical phenomena cited above has a proof whose difficulty far exceeds that of the typical geometry theorem; in some cases the proof requires advanced methods from areas of mathematics that are entirely outside geometry. One can hardly expect such a proof to find a place in a high school geometry course. To some extent this might be mitigated in the future if dynamic geometry software succeeds in finding a role in the classroom: such software makes it possible for students to "confirm" surprising results, even ones as complicated as those above, without actually undertaking a proof. But even if this were to happen, the fact that the DGS provides instant feedback means that there will still be no long-standing but unproven beliefs to be either confirmed or disconfirmed.

Theory-building & Problem-solving

The final dialectical pair of categories of mathematical appreciation described in Chapter 2 was *theory-building / problem-solving.* Recall that *theory-building* places a high premium on the logical interconnections among the elements of a theory: which may be taken as postulates, which are logically equivalent to one another, and so forth. From this perspective, a re-organization of an existing body of knowledge (e.g. providing a new definition for an existing class of objects and showing that the previously-accepted definition is a provable consequence of the new one) would be regarded as a significant accomplishment. In contrast, Gowers (2000) describes combinatorics and related fields as areas of mathematics that are organized not by the interconnections among theorems, but rather by the presence of heuristics that can be applied to a broad range of context: in such fields, *problem-solving* takes the place of theory-building as an organizing principle.

I begin my analysis with the conjecture that the high school geometry class is by default organized around the work and value of problem-solving, rather than theory-building. Students learn to use algebra (and later, trigonometry) to find the measures of unknown segments and angles, and to calculate areas and volumes. Proof, as it exists in the high school course, is largely an opportunity not to develop the elements of a theory but rather to put in practice various proof-methods that students are accountable for learning (side-side-side, angle-side-angle, etc.; segment and angle addition arguments; and so forth), and so demonstrate that they have learned "how to prove" (Herbst, 2002). In this section I investigate the extent to which the study and focus group archive presents evidence that teachers place any value on the complementary value of theory-building.

The topic first appears in a discussion from an early study group meeting in which Megan raises the question of whether certain principles of triangle congruence that are designated as "Postulates" in her textbook might more appropriately be labeled as "Theorems", or indeed not included at all on the grounds that they add nothing to the theory. She recounts an episode in which she taught her students "Hypotenuse-Leg" (a method of proving two right triangles congruent by establishing that they have congruent hypotenuses and a pair of corresponding congruent legs).

119	Megan	.... But we had this picture up for Hypotenuse-Leg today, the two triangles. And I said let's talk about, 'cause it's a postulate in my book. So I said, well let's talk about this as a postulate. And some kids wanted to argue that it shouldn't be a postulate. It should be a theorem. And their idea was that postulates, we just believe those, and this they thought we could build a logical argument and they said, if we know two sides of a right triangle, we really know the third, because of the Pythagorean theorem. And they were rubbing into me, I have this book in my room, 300 ways to prove the Pythagorean theorem, they said see, we can prove that 300 ways! So then they said, so this shouldn't be a postulate. Once we have two sides and a right angle, we could argue that we really know the third side so the two triangles are equal by side-side-side or side-angle-side. Which is a pretty valid argument. One of the kids said no, you can't check every set of numbers. And another kid then went up to the board and put letters on them, he had a B here and a C here, and said instead of calling this A like you would for a right angle, it's the square root of C-squared minus B-squared. Which he's right. And so is this one. This is a great, you know, it was a great little thing that came out. And then how do you tell kids how you pick the postulates? You know, who picks that? Should it be a postulate or a theorem? It's a theorem in your book. Then why did they make it postulate?
120	Cynthia	Because your book company hasn't proven it yet and ours has.
121		[Laughter]
122	Edwin	Yeah, right.
123	Megan	Okay, thank you very much.
124	Penelope	They won't buy us new books.
125	Megan	Because I said, you're right, you know, and this is an interesting thing, because some books have side-side-side as a postulate, have just a few of them as postulates and all of the rest theorems, some just decided those are all postulates. [Cynthia: Right.] It's a really --
126	Tina	See that's what I was saying, when I taught out of the two different geometry books, that was a headache.
127	Cynthia	Yeah.
128	Megan	Oh, that's true, 'cause you had all that (inaudible).

129	Tina	(Because it was,) you know the same thing, it was postulate here, it was a theorem there and I would stand in front of the kids and be so confused I wouldn't know what I was talking about.
130	Megan	No, part of me is happy that they're thinking about the difference between a postulate and a theorem. [Tina: Right.] You know, that's a pretty valuable thing. But then, you know, how do I say, well why did the book pick that? You know I actually said I agree with you, this is pretty dumb. [laughs]
131	Tina	Well, I need to tell them that it's just what the authors picked. That's the way they chose to do it. It is a theorem, it is a theorem in our book.
132	Denise	I tell you my students, postulates, theorems, who cares? Use them to prove a proof. [laughs] They know you can use them all as reasons in a proof. And they're like ok.
133	Megan	Well did you, did your book have, this book also, in this section, it has three little minor theorems: Leg-Leg, Hypotenuse-Angle and okay... we were making fun of the names, "Ha!" We were singing like. I said they were not useful at all. I said I'm only telling you these because --
134	Tina	It's right angle.
135	Megan	You're right, it's "ra," leg-leg and "ha."
136	Tina	It just takes all three of them, it just takes the other one, the right angle out.
137	Megan	But why do they have those? They're so repetitive. I told the kids, this is a waste of time, it's the same as side-angle-side. So we wrote them all up, because they will show up in the answers. That's the problem. If you don't do them, kids doing their homework looks in the answers, it says "L-L" they're like what is that?!
138	Tina	See our book has taken those out.
139	Megan	That's good. See I never saw it any other book. The book I previously taught in didn't have it.
Excerpt 26. From ITH111605, interval 10.		

There is much in this excerpt that is relevant for the present discussion. First, we note that Megan (and her students) are sensitive to the fact that the theory, as presented in the textbook, is less economical than it could be: the identification of "hypotenuse-leg" as a postulate rather than a theorem, and the superfluous inclusion of "leg-leg" (which is really nothing more than "side-angle-side" in the case of a right triangle) are particularly questionable. The fact that students in Megan's class not only produced a more-or-less complete proof of "hypotenuse-leg", recognized that that proof was based on the Pythagorean Theorem, and argued on that basis that "hypotenuse-leg" ought to be classified as a theorem, rather than a postulate, seems like a clear instance of theory-building as classroom activity. And Megan indicates that she values it as such: not only did she allow what appears to be a significant amount of class time to be devoted to this discussion, she explicitly labels that activity as “a pretty valuable thing” (130).

But Megan’s appreciation of the value of this discussion is tempered (“*part of me* is happy”, 130, emphasis added). As much as she values that her students are engaged in a discussion over whether a given proposition ought to be classified as theorem or postulate, she is unsure how to respond to students: she appears to be genuinely mystified as to the grounds on which some propositions are deemed to be postulates while others are theorems. Significantly, though, Megan consistently refers to this as a decision to be made by a group of unspecified and somewhat mysterious others, referred to variously as “they” (120, 137) and “the book” (125, 130, 133). The possibility that the class might be entitled (or even expected) to take these decisions on themselves is both novel and problematic for her.

Megan’s comments are echoed by the other teachers in the conversation. For Tina, the important thing is for the teacher to conform to the will of the textbook publisher: different texts use different sets of postulates, those choices are essentially arbitrary (“it’s just what the authors picked”) rather than based on any grounds or values, and the teacher must simply adapt to the text. Finally, we note Denise’s position rejects the entire topic as unworthy of consideration. For Denise, it simply makes no difference whether a particular property is labeled theorem or postulate. Both are equally important insofar as students must learn to use them in a proof; but the designation of one of them as postulate and the other as theorem is of no significance. Denise thus embodies the problem-solving disposition in its purest form, with utter disregard for the theory-building disposition. In contrast, Megan seems to clearly recognize the importance of the theory-building perspective, and seems noticeably proud that her students have begun to exhibit the same disposition, but is unsure of how to validate her students' contributions in the face of the officially sanctioned curriculum. Tina represents a middle ground: she recognizes that the organization of a theory matters (i.e. it is something visible that she can talk about) but she has surrendered any agency on the matter.

Later in the same study group meeting Megan and Denise added to their earlier comments:

157	Megan	... But when we go over them and you talk about this is a postulate, so we can't prove it, and that's what bugs people about this one, is that I'm presenting it as a postulate, it says in the book it is, but they're saying wait, I can build a pretty good argument for that one, you know?
158	Denise	A lot of the postulates though you *can* prove, they're just something that you don't *have* to prove, it's just taken as true. A lot of postulates I found you can prove. That's why in some books they are theorems, because you can prove them and yet, you know what I'm saying?
159	Res. 1	So you mean, postulate means we don't have to prove it?
160	Megan	It's a commandment of math, that's what I say. [laughs]
161	Res. 2	So --
162	Denise	-- a theorem or something, but a postulate you use to prove theorems. Did I say that all right?
163		[Laughter]
164	Denise	Okay, yeah postulate is used to prove theorems. But I mean --
165	Megan	Most of them you can't. Like the ones, "two lines intersect in at most one point". Now those are like building block postulates and you just sort of look at it and say, we believe that, you know? But like I said, this one, they could build a pretty good argument for it. So, that's why [inaudible].
166	Tina	I'd be impressed if my kids came up and argued that fact, that it's a theorem not a postulate.
167	Denise	My kids wouldn't even try. Even if they thought of it they wouldn't let me know.
168	Megan	That's true, cause a few kids in every period, and one of them will bring something up, you can see them rolling their eyes, like here we go.
169	Res. 2	So you would just be happy that it happens but you wouldn't feel that you have to make that happen.
170	Denise	You mean that them making the distinction between postulate?
171	Res. 2	Yeah.
172	Denise	Oh, yeah I mean when we studied postulates, when I first introduced that word postulate to them, I want you to know what that is, and I want you to know what

		theorem is and I want you to know the difference. But you use them all to prove theorems. You use the theorems to prove other theorems, you use postulates to prove theorems. I mean you need to know that these can be reasons and you can use these as reasons in a proof, along with definitions and corollaries.
Excerpt 27. From ITH111605, interval 12.		

The difference between Megan's position and Denise's position becomes clearer in this excerpt. For Megan, a postulate is something that *cannot* be proven because it is too fundamental (the "building blocks", 165); she makes use of words denoting the possibility or impossibility of proving the statements ("can't", "can", "could"; 157,165). (Note also, however, the opposition present in turns 157 and 165 between "prove" and "build an argument". A postulate is something that cannot be *proven*; what students have done with respect to "hypotenuse-leg" is *build a convincing argument*, which seems to mean not quite the same thing.) In contrast, Denise uses words denoting the necessity, rather than the possibility, or proving a statement ("have to prove", 158): for her, a postulate is something that one *does not have to prove* because it has been labeled as such by an arbitrary choice of some authority. Note that Denise's characterization of postulates lacks any indicators of agency on the teacher's part.

If Denise's position can be critiqued as somewhat unmathematical, in part one can point to a century of geometry curricula in her defense. Although mathematicians working at the foundations of geometry have a long tradition of exploring the independence and interdependence of various axiomatic formulations of the subject (see in this connection the discussion of R.L. Moore's work in the previous chapter), Geometry curricula have largely avoided the matter altogether. Virtually every Geometry curriculum written for school use in the last century has included redundant axioms for ease of presentation and clarity. For example, one common practice has been to include “ruler and protractor axioms” into the curriculum. These axioms build the metric properties of the plane in to the theory from the beginning, and render most of the other axioms unnecessary. On the other hand, most textbooks do not discard the traditional axiomatic system, perhaps because doing so would have produced a course that was not recognizably the same as the one that had been taught for the previous half-century. The result is a theory that is overloaded with redundant postulates that are rarely used and even more rarely considered critically. Our study group teachers often seemed not even sure what were the postulates of their textbook:

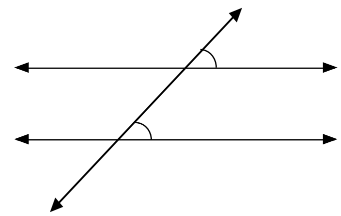
827	Mitch	I think the book I’m using now has a compass postulate. I think it calls it a compass postulate, that when the, when the pointer and the pencil are separated and you mark it--
828	Megan	Compass postulate.
829	Esther	I’ve never heard of that!
830	Mitch	--and you move it over, you’ve given the same distance.
831	Megan	Well you have a ruler postulate, our book has that. And I actually made fun of that.
832	Mitch	Yeah, it’s (connected--)
833	Megan	(Because) it’s very lengthy and long and I say to kids, what does this actually say?
834	Mitch	It says you can construct a ruler.
835	Megan	And you can usually get one kid to say, you can make a ruler and measure things? I’m like, Yes! That’s what it says. Do we really need this as a postulate? But that’s, compass postulate would be just like that.
836	Mitch	Well I think–yeah, or you like put the compass on the ruler...I don’t remember how it’s stated in my--
837	Megan	You can use both together!
838	Mitch	--it might even be the ruler postulate, but if we think of it as putting the compass on the ruler, and then we have equal lengths.
839	Megan	That’s hilarious.
840	James	Because they didn’t--
841	Moderator	In our book they had the ruler postulate and the protractor postulate.
842	Megan	Oh yeah, that’s protractor–that’s right, I forgot about that.
843	Moderator	But we didn’t have a compass postulate.
844	Mitch	Well, my–in fact my memory might not be quite right but what we were doing is putting it on the ruler. So I guess that’s using a new tool--
845	Megan	That’s terrible.
846	Mitch	--a new tool to make distance.

Excerpt 28. From ESP020706, interval 18.

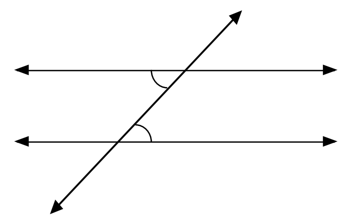
From a mathematical point of view it is hard to understand how one could possibly teach any proof-based mathematics course without knowing exactly what axioms underlie the theory. This is all the more surprising in the case of Geometry, which for centuries has drawn attention to questions of the independence and completeness of its axioms. That experienced teachers of Geometry express clear uncertainty about the axioms of the course they teach is indicative of the low value they place on holding students (and themselves) accountable for the theory-building disposition.

Despite this there are natural places within the Geometry curriculum where one might expect that opportunities for cultivating the theory-building disposition might be situated. For example, consider the four propositions stated and illustrated in Figure 3.5, below. These propositions state various properties of particular pairs of angles formed by intersecting a pair of parallel lines with a transversal. All four of these properties are common elements of typical high school Geometry curricula — in fact, in most textbooks these four properties occur within a few pages of one another. It seems safe to say that each of the properties asserted in Table 1 are all taken as objects of the didactical contract that binds Geometry teachers and students together: students are expected to know each of those four properties, and teachers hold their students accountable for that knowledge.

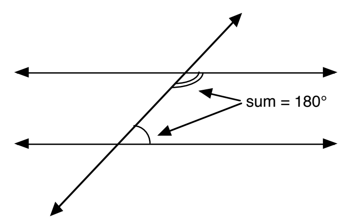
1. If a pair of parallel lines is intersected by a transversal, then the corresponding angles are congruent.



2. If a pair of parallel lines is intersected by a transversal, then the alternate interior angles are congruent.



3. If a pair of parallel lines is intersected by a transversal, then the consecutive interior angles are supplementary.



4. If a pair of parallel lines is intersected by a transversal, then the alternate exterior angles are congruent.

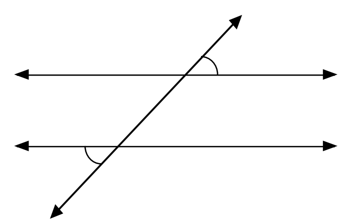


Fig. 3.5. Some equivalent statements about parallel lines.

And yet what is important from the theory-building point of view is not that the four properties in Fig. 3.5 are *true*, but rather that they are all *equivalent* to one another. In the first place, they are "true" only relative to some set of postulates, not in any absolute sense: in non-Euclidean geometries they are all false, and in fact they are all also equivalent to the Euclidean parallel postulate. The conventional high school geometry curriculum obscures this by taking one of these properties *as* a postulate, *in addition* to the Euclidean parallel postulate. But beyond noting that the designation of one of these four properties as a "postulate" is *unnecessary*, it is also important to note that the selection of *one* of them to be so designated is *arbitrary.* That is to say, if *any* of the four properties were taken as a postulate, then the remaining properties could *all* be proven as easy consequences. Moreover there are subtle interrelationships between these postulates and other consequences of (and conditions that imply) Euclid's parallel postulate, such as those presented in Fig. 3.6. The network of implications among those properties is complex; some (but not all) of them are equivalent, some (but not all) of them can be proved from the other axioms of geometry (and thus hold in geometries other than Euclid's).

1. Parallel lines are coplanar lines that never intersect.
2. Parallel lines are always the same distance apart.
3. If a line segment intersects a pair of lines forming two interior angles on the same side that sum to less than 180°, then the pair of lines intersect on that same side.
4. Given any line, and a point not on the line, there exists *at least one* line that passes through the given point and is parallel to the given line.
5. Given any line, and a point not on the line, there exists *not more than one* line that passes through the given point and is parallel to the given line.
6. If a pair of parallel lines is intersected by a transversal, then the four angle-pair properties indicated in Table 2 hold.
7. If a pair of lines is intersected by a third line such that any one of the four angle-pair properties indicated in Table 2 hold, then the pair of lines is parallel.
8. The sum of the three angles in any triangle is 180°.
9. Parallel lines have the same slope.

Figure 3.6. A collection of interrelated properties.

But these interrelationships are not typically part of what students are held accountable for studying. Indeed none of the contemporary high school geometry textbooks I have reviewed (Larson, Boswell & Stiff, 2001; Jurgensen, Brown, & Jurgensen 1990; Schultz, Hollowell, Ellis & Kennedy, 2001) suggest that the organization of the theory could be other than the form in which it is presented. Even

Lang & Murrow (1980) and Moise & Downs (1991), which use far more sophisticated and parsimonious organizations of the theory, present the theory of parallel lines as a *fait accompli* — the work of organizing the theory is carried entirely by the authors; students need not concern themselves with it. Serra (1997) is a notable exception: In a note to the reader Serra (p. 727) states explicitly that the designation of one of the properties in Figure 3.5 as a postulate is purely arbitrary, and he encourages students to explore alternative restructurings of the theory. Also in Serra's textbook we find explicit attention to the network of interrelationships among the properties of Figure 3.6 (pp. 745, 759). However, as Serra defers all discussion of postulates, theorems, and proof until the last 60 pages of an 800-page book, it is unclear to what extent students are actually held accountable for this content in classrooms that make use of that text. In any event, it remains true that in none of the other works are students expected to investigate these matters; in all other cases the textbook authors do that work themselves, giving no hint that the theory could be structured differently. Questions about the organization of the theory are simply not part and parcel of what the geometry curriculum is about.

One ThEMaT animation was designed specifically to probe how teachers respond to a story in which theory-building plays a prominent role with respect to the theory of parallel lines. In "Postulates and Parallel Lines", a teacher introduces a new postulate ("If two parallel lines are cut by a transversal then corresponding angles are congruent") and proves two theorems ("If two lines are parallel then alternate interior angles are congruent" and "If two parallel lines are cut by a transversal then same-side interior angles are supplementary"). Following the proofs, the teacher briefly remarks that the three propositions could have been done in a different order, with one of the theorems taken as a postulate and the postulate proven as a theorem. In addition to the story as just described, three alternate endings were also represented as separate animations. In the first alternate ending, the teacher undertakes to do precisely what she had previously only described as a possibility: she erases the board, announces that she is starting over, and shows how the same three propositions could have been organized into a theory differently by taking the first theorem as a postulate, and proving the original postulate as a theorem. The animation depicts the lesson running smoothly. In the second alternate ending, the teacher follows the same course of action as in the first alternate ending, but things do not go so well: students seem confused and angry by the "starting over" move, ripping out the pages from their notebooks and crumpling them on their desk. In the final alternate ending, the teacher states briefly that the theory *could* be organized in a different order, and invites students to write up the details for extra credit.

Two study group meetings were devoted, almost in their entirety, to a discussion of this story and its alternatives. Both meetings followed the same agenda: a prompt asked participants to discuss how they teach their students the difference between a theorem and its converse. After some discussion of this, the main version of "Postulates and Parallel Lines" was screened and discussed. Participants were asked to comment particularly on the ending, in which the teacher states that the theory could have been developed in a different order. Later in the session, each of the three alternate endings was viewed and discussed in turn.

These two meetings were the only ThEMaT study group sessions that were structured around an animation that was specifically designed to elicit teachers' responses to a representation of the Theory Building disposition, and the animation proved successful at doing so: Of the 119 intervals in the two meetings, 41 (just over a third) included some significant discussion of the status of propositions (as "postulate" or "theorem"), and of the possibility of restructuring an existing theory into a different configuration of postulates, theorems, and definitions. Because of the large volume of the data it is not possible to report here all of the relevant commentary that teachers offered; I will present only selected excerpts.

Participants in both study group meetings were quick to identify the animated teacher's intention ("I guess her point was that any one of those could've been a postulate"; Raina, TMT030607, interval 15, turn 318). They also were in broad agreement that the teacher was less than successful in achieving that goal. They differed, however, in their assessment of whether the goal was a *worthwhile* one, and it is this disagreement to which I now turn.

Almost immediately after completing the viewing of the first version of the story, Jillian opined that "It was a terrible idea. Absolutely terrible idea." (TMT030607, interval 15, turn 322). When asked to expand on her strong criticism, Jillian explained that the problem had to do with a mismatch of (what she perceived to be) the teacher's goal with the *timing* of the story: a lesson on the properties of the angles formed by parallel lines would fall "so early in the year", and "at this point in the year it's absolutely terrible" because students would be completely unable to understand the teacher's point (324, 326). She conceded that the "lesson" (by which she seems to mean the learning goal) could be a valuable one, but concluded with "I would never do it with my kids" (328).

Immediately after Jillian offered this strongly-worded opinion, Lucille offered her own experience as a counter-argument: "I have done it… I don't see anything negative to it… It didn't bother me that she said any one of these could've been, um, postulates" (330, 335, 343). To justify her appraisal, Lucille cited something she said she had learned from participating in previous study group meetings: "that every textbook is a little bit different" (339), with a different presentation of definitions, postulates, and theorems, and that some students may encounter these different presentations — for example, in online resources, or in the textbooks of friends who attend other schools (339, 365). On hearing this explanation, Jillian immediately reversed her previous opposition to the story, fully endorsing the teacher's goal but continuing to criticize the execution of the goal ("It was so bland", 360, 366).

Other teachers offered different reasons why a lesson like that depicted in the animation would be unnecessary, undesirable, or unlikely to succeed. One common opinion voiced by many of the teachers in both study group meetings was that, fundamentally, the difference between a postulate and a theorem is of no consequence. Articulating this view, Greg said "I really don't, personally I don't care if they think it's a postulate or a theorem" (TMW032107, 419), echoing Denise's comments above from a different session (see Excerpt 27). The same sentiment was expressed later in the same meeting by Tina and Cadie (863-865), and in the other study group meeting by Lucille: "When they're proving I don't really care whether they say postulate or theorem. That they're all like kind of equal in status" (TMT030607, 562). None of the teachers participating in either study group meeting voiced an opposing viewpoint, despite the fact that in both meetings the moderator attempted to garner some sympathy for the position that the status of a proposition as postulate or theorem is an important thing for students to learn. Instead, teachers repeatedly endorsed the view that what matters is for students to learn how to *use* the various properties of parallel lines in order to do proofs and other exercises. That is, they value the properties of parallel lines for the role they play in a diverse range of problem contexts — a perspective consonant with the *problem-solving* disposition, as articulated by Gowers (2000).

There was also a broad consensus among most of the teachers that factors beyond their control make it extremely unlikely that they would devote class time to a proof that two conditional statements, one dubbed a *postulate* and the other a *theorem*, are in fact equivalent. Teachers cited time constraints (TMW032107, 648, 978), an overcrowded curriculum (TMT030607, 861, 887-893), and students' intellectual immaturity (TMT030607, 328, 863, 874). On this latter point, one strand of arguments in particular seems noteworthy. Melanie, in TMT030607, interval 16, raised the concern that "in these upcoming years all students are going to be doing this type of proof and, I mean, my special ed kids – I'm really concerned with how much they're even going to even understand" (398). From an outsider's perspective, it may seem quite extreme for a teacher to indict a proposed course of action on the grounds that it would be too difficult for special education students[23]. But Melanie's concern was echoed much later in the same meeting by both Lucille ("Well, when we talk about some of the concerns of students with special needs … y'know it's, it's an interesting thing because what they're

[23] The teachers seem to take for granted that Special Education students are less intellectually capable than others, a position that oversimplifies the wide range of conditions that Special Education students may have. As the goal here is to report teachers' perceptions of what is and is not viable in the classroom, I defer to their own use of the term.

expecting us to do is raise the level or the rigor of what's going on and yet they're asking us to do it with everybody", 675) and Jillian ("You're gonna tell the special ed kids that [the postulate could be a theorem, if you took something else as a postulate] though right? You're not, I don't think. I just think that's not going to make any sense", 780). These three teachers call our attention to the way that institutional factors — such as, for example, a completely de-tracked curriculum in which special education students are mainstreamed into the same Geometry course as advanced students (398) — can be experienced by teachers as constraints on what kinds of teaching moves are viable in their classroom.

When teachers were asked to respond to the particular pedagogical strategy of erasing the board and "starting over" (as shown in the first two alternate endings), there was a broad consensus that students would be confused (TMT030607, interval 24) and even angered by such a move (TMT030607, interval 33). Instead they proposed various other ways the teacher could have conducted the lesson (TMT030607 interval 24, TMW032107 interval 60). The teacher might use one half of the board for the initial development of the theory, and use the other side for the "alternate" development; the teacher might distinguish the two possibilities using different colors, or surround it with a "cloud" marking it as a brainstorming exercise or thought experiment only. They suggested using phrases like "alternate universe" and "through the looking glass" to label the second development of the theory as a distinct possibility. Throughout all of these proposals, teachers consistently use modal verbs denoting possibility and subjunctive constructions ("we could have done it this way", "if we had done it differently"). All of these strategies convey the notion that the theory *could have been* developed differently, while at the same time preserving the initial development in the official record of the class's work (the board, students' notebooks). In contrast, the "erase and start over" move depicted in the animations was characterized as signaling to students an actual restart: what is erased is not only the literal board, but also the students' collective memory of the lesson to that point (González, 2009). Participants reacted negatively to this, stressing that what the teacher needs to communicate is that she is "just taking a different perspective" (TMT030607, interval 47, 951), rather than starting over.

Underlying all of this discussion, and surfacing explicitly throughout it, was the question: What makes some statements a "postulate" and others a "theorem"? In the animation, the teacher says that a postulate is "so obvious that we just take it for granted," and, later, that "postulates are just things we assume to be true." In the study group meeting, Raina objected to the characterization of postulates as "something so obvious that we just take it for granted"; she asked, "What makes the first statement more obvious than any of the other ones that they would consequently prove?" (TMT030607 interval 15, 318). Later in the same session, Jillian and Cynthia reacted to the animated teacher's phrase "just things we assume to be true" on the grounds that "It seems so arbitrary" and "It sounds like they're not important" (interval 40). But participants' own speech revealed a wide variety of ways of speaking about postulates. They used various metaphors, such as the foundation of a house (TMT030607, interval 24, 565) and the (arbitrary but inflexible) rules of a game (TMW032107, interval 24, 473). As noted above, study group participants were in widespread agreement that students should not be held accountable for knowing which statements are postulates and which are theorems; despite this, teachers still held some fidelity to the notion that not everything should be taken as a postulate, that "we kind of like to keep the postulates at a minimum and do the bulk of our work deductively" (Jillian, 842). Teachers acknowledged that the curriculum they teach takes as postulates many more statements than is really necessary ("Most of the postulates in our books… [are] just theorems that are beyond the scope of this course", 859), a state of affairs they describe as being "lazy" (TMT030607, interval 33, 748) in order to make things "easier" (TMT030607, interval 23, 515).

In their discussion, participants reveal a basic tension they experience with respect to theory-building. On the one hand, they appreciate the value of producing many theorems from a small number of postulates, and want their students to appreciate this as well (TMT030607, interval 40, 842; TMW032107, interval 24, 565). On the other hand, they do not hold themselves responsible for using only a minimal set of independent assumptions ("That's what college is for"; TMT030607 interval 25, 571). Instead, they defer to the authority of the textbook and its authors, who determine what may, and what may not, be taken as true without proof.

It may be that individual teachers' knowledge of mathematics may influence how they perceive the role of postulates and theorems in mathematics. In some cases teachers made arguments that suggest a serious misunderstanding of this role. Karen, for example, distinguished between "axioms" (which she understands to refer to statements that cannot be proven) and "postulates" — which she understands to refer to statements that are "one level up" in the hierarchy, statements that *can* be proven, but for which the proofs are difficult and hence omitted for ease of presentation (TMW032107, intervals 24 and 33). In a more extreme case, Denise endorsed the notion that a postulate is an obvious truth, and used this to argue that *none* of the properties of parallel lines should be called postulates, on the grounds that (unlike other postulates, such as the fact that two points define a single line) none of the properties of parallel lines are obvious (TMW032107, interval 24, 456-458) — a view that calls to mind the long history of attempts to find a proof for Euclid's parallel postulate, motivated in large part by the sense that it alone among the postulates was not truly "self-evident" and hence ought to be provable (Greenberg 1980).

In contrast, both study group meetings contained at least one moment in which a participant made a connection to larger questions of theory in geometry and other branches of mathematics. Raina reported that she normally includes, as part of her teaching the parallel postulate, a discussion of the many (failed) efforts over the century to prove it, and how those efforts typically include an assumption that is later recognized as equivalent to the parallel postulate itself — for example, "you can take the… triangle angle sum being 180, is logically equivalent to the parallel postulate" (TMT030607, interval 16, turn 372). Like the teacher depicted in the main branch of the animation, Raina does not pursue a proof of this equivalence when teaching; she merely informs her students that such a proof is possible. None of the other teachers present in that study group meeting respond directly to Raina's report. In the other study group meeting, Karen cited Gödel's Incompleteness Theorem to justify why it is acceptable for Geometry curricula to include many redundant postulates that are amenable to proof: "We did this whole race to prove all of mathematics, and we wanted to limit our postulates… but that sort of went out the door with Gödel anyway" (TMW032107, interval 34, turn 636). Again the other teachers present seem to have little to say in response to this connection; one of them, Madison, says only "I don't remember that".

These two episodes provide some evidence to suggest that advanced mathematical knowledge, and in particular knowledge of the history and philosophy of mathematics, may play a role in supporting teachers' ability to speak in terms of the theory-building disposition. But even if this is so, it is also noteworthy that neither Karen nor Raina, or indeed any of the teachers who voice some appreciation of the values of mathematical economy, speak of *holding students accountable* for studying the equivalence of different formulations of an axiom, or for investigating the independence of proposed axioms. Although they give indications that they recognize these as mathematically valuable activities, they offer no evidence that they carve out a role for these activities in instruction.

Thus we see that, even when teachers give evidence that they place some value on the theory-building disposition when they consider it from the perspective of *mathematicians*, they consistently alienate themselves from that disposition when considering it from the perspective of *teachers of mathematics.* This illustrates well the sense in which practical rationality — the collection of categories of perception and appreciation with which practitioners view and appraise the world — is a characteristic not of the individual practitioner, but rather the practice as a whole. The value that practitioners attach to theory-building is that it may be important for mathematicians (and they may include themselves in that group), but it is *not* important for high school teachers and their students. Time constraints, the intellectual limitations of their students, and the obligation to conform to the authority of the curriculum all are cited as reasons why teachers reject the idea that they might make theory-building a central classroom activity as portrayed in "Postulates and Parallel Lines".

Discussion

In this chapter I have examined a large corpus of conversations among experienced high school geometry teachers gathered in monthly study groups to find evidence of how they relate to some of the various dispositions that were proposed, in the previous chapter, as characteristic of mathematicians' professional vision. I now summarize my findings and discuss their significance in the light of the larger questions of this dissertation.

In Chapter 2, I showed that mathematicians describe themselves as problem-posers who actively seek areas of doubt and uncertainty in which to explore. Mathematicians use various generative moves — modifying hypothesis and conclusion, generalizing, specializing — in order to generate new questions and hypotheses. They bring various frames to bear in appraising the value of questions and results: they may be appreciated for their applicability to contexts outside of mathematics, or for their abstract qualities; for confirming or disconfirming a result long believed true; for revealing relationships among the elements of a theory, or for providing a set of heuristics that can be applied in a wide range of problem-solving contexts.

In this chapter I have turned to a corpus of records from a collection of study group and focus group meetings of experienced geometry teachers, and looked for evidence of the way these practitioners relate to the dispositions of mathematicians. In these study groups, practitioners considered various teaching scenarios in the form of animations and other representations of teaching — representations designed to elicit their own practical rationality. The results are revealing. Although many teachers in the study group meetings were quick to acknowledge the potentially generative role that doubt and uncertainty can play in motivating student inquiry, they were consistently vocal in their criticism of teaching stories in which a teacher pursues such doubt and uncertainty. In voicing this criticism, teachers pointed to the conditions of their work: these include time constraints, the need to cover the curriculum and prepare students for mandatory standardized testing, student absenteeism and attentiveness, the heterogeneous qualities of their classroom, and students' intellectual immaturity. They also spoke of the nature of their professional obligations to their students: ultimately, teachers perceive their role as obliging them to, eventually, resolve all questions and clarify what the truth is. A teacher who fails to do so is guilty of violating the tacit didactical contract, and risks jeopardizing both their relationship with their students and their status as authorities. Because students know that teachers will eventually step in and clear up all mysteries, the potential of doubt and uncertainty to motivate student inquiry is undermined and neutralized.

Many of the excerpts I have cited and discussed in this chapter suggest that, unlike mathematicians, geometry teachers do not view mathematics largely as the study of relationships among contingent possibilities (what *could be* true, what *would be* true under certain conditions); rather, they regard it as the study of the factual properties of diagrams, which represent objects that are acknowledged to be idealizations but nevertheless are regarded as "real" in some sense (a kind of tacit and unexamined Platonism). For this reason, many of the generative moves that underlie mathematical inquiry are found only rarely in classroom practice: students are not expected to seek a conclusion that might come from strengthening or weakening a particular hypothesis, or to seek a hypothesis that might be sufficient for a particular conclusion. Because diagrams by their nature collapse a conditional statement into a simultaneous presentation of concomitant facts, even the disposition towards considering whether the converse of a proposition is true — a disposition that has an avowed place in the structure of the geometry curriculum — gets little play in actual practice. And because "proof" is allotted a designated time within the curriculum, and is viewed by teachers as *a thing to learn* rather than as a tool to know with, conjecturing and proving are experienced in classrooms as distinct and disjoint activities.

With respect to the pair of dispositions *Utility* and *Abstraction*, the data testifies to a fundamental paradox in how teachers speak of the value of the course they teach. Teachers speak of the abstract quality of what they teach in almost apologetic tones (this sentiment finds its expression in the frequent refrain, "In real life, nobody's ever going to ask you to prove two triangles are congruent"); implicitly, they seem to agree that the course must be justified in terms of its real-world utility, which they locate in the value of *proof itself* (rather than on the content of what is proved), citing countless examples of contexts in which argumentation is a useful life skill. This is consistent with the observation just made above, that proof is perceived as a thing to learn, rather than a tool to know with. The fact that the examples offered to justify the "real-world" utility of proof are all illustrations of argumentation that are not purely deductive undermines this argument, but this goes unrecognized by teachers. Notwithstanding this, teachers acknowledge that the content of geometry *does* have "real world" utility, particularly for architects, engineers, and workers in the skilled trades, but this utility plays only a marginal role in their classrooms.

As was noted above, teachers consistently cite factors beyond their control as reasons why the potentially generative functions of doubt and uncertainty are seen to pose greater risks than benefits. A consequence of this perspective is that little to no space exists within the Geometry classroom for the mathematical dispositions of *Surprise* and *Confirmation*: teachers feel obliged to ensure that all questions are resolved by the end of every class period, an obligation that is incompatible with the practice of allowing open problems and conjectures to exist over extended time periods. In arguing that there are no surprises in high school Geometry, I do not mean to suggest that every result in Geometry is immediately obvious (although many are); rather the claim is that a particular pedagogical practice — allowing the emergence of conjectures (some true, some false) that remain unresolved for days, weeks, or months, eventually to be either confirmed with a proof or disconfirmed with a counterexample — might conceivably be endorsed by teachers who appreciate the value of surprising or confirmatory results, but is instead universally rejected as inappropriate for the high school setting.

In the final section of this chapter I discussed a number of study group conversations in which teachers spoke about issues related to the disposition pair *Theory-Building / Problem-Solving*. I showed that teachers were, by and large, uninterested in the organizational structure of Geometry. Many teachers expressed uncertainty over which statements are designated as postulates, which as theorems, and which as definitions in the curricula they teach. They were aware that different textbooks may structure the theory differently, but they regarded that choice as an essentially arbitrary decision of the authors, and indicated that they did not regard an investigation of the different possibilities as appropriate material for classroom work. More than one teacher told the study group that they expressly tell their students not to worry whether a statement is a postulate or a theorem; all that matters, they teach, is that students know how to *use* the statement for the purpose of problem-solving (here "doing a proof" is included as a special kind of problem). Some teachers were aware that many of the statements designated as "postulates" in the curriculum are, in fact, provable consequences of other postulates, but they attach little significance to this, and as a matter of course do not engage their students in discussions of the independence of postulates. As in the other cases, teachers cite time constraints, an increasingly overcrowded curriculum, and their students' lack of intellectual sophistication as reasons why classroom theory-building is inappropriate.

Across all of the excerpts cited and discussed in this chapter, there are two noteworthy themes that warrant further elaboration. The first is that in many of the conversations in the study group meetings there was at least one teacher who was able to temporarily step outside, so to speak, the perspective of the practice, and view matters from a point of view more akin to that of the mathematical sensibility. But even these teachers, when asked to comment on how they might deal with an imaginary classroom scenario, voiced the same opinions as their colleagues. This is significant, in that it confirms the notion that practical rationality is not a characteristic of the *individual practitioner* but rather a collective phenomenon, a characteristic of the *practice*. An individual teacher may, at times, look through different sets of lenses, but when they put on their "teacher goggles" (so to speak) they perceive and value differently than they do when they wear "mathematician goggles".

The second significant theme that runs across all of the examples in this chapter is the remarkable *lack of agency* that teachers exhibit when they discuss their practice. On questions of mathematical value, the teachers in this study consistently subordinate their own desires to the authority of the textbook, to the (perceived) deficiencies of their students, to the clock and the calendar and the many other institutional factors over which they have no control. The study group participants' responses tend to come in the form "We cannot do X because Y", or "We have to do X because Y", where Y is some institutional reality (and, in particular, not one of the dispositions that comprise the mathematical sensibility). One might imagine teachers arguing in favor of a particular course of action *despite* the institutional factors that militate against it, on the grounds that it is mathematically appropriate to do so — but I have been unable to find examples of this in the corpus of study group meetings.

This last observation serves to motivate the next two chapters. In them, I turn to a collection of documents produced over a three-year period by a teacher who, it will be shown, consistently created assessment items in which students were held accountable for elements of the mathematical sensibility. The analysis of these assessment items, considered individually and over time, provides another, complementary vantage point for the questions that were posed in Chapter 1: Can the geometry course be used to cultivate, in students, the dispositions that are characteristic of the mathematical sensibility? What factors oppose such a use? And, finally, how does that intended use adapt over time to those opposing forces?

Chapter 4

Looking for the mathematical sensibility in a corpus of examination questions

Introduction

In the preceding chapters of this dissertation I analyzed a collection of mathematical narratives to identify a number of key dispositions that appear to be characteristic of the "mathematical sensibility" — the practical rationality that typifies the professional vision of a mathematician (Goodwin, 1994). Following that analysis, I turned to a corpus of data from a series of study group meetings among experienced geometry teachers to see what role, if any, those dispositions play in their own practical rationality: that is, to see whether those same dispositions have a customary place in the teaching of high school geometry. In this chapter, I turn to analyze a corpus of geometry examination questions that was constructed by a teacher[24] in a small, private high school in an affluent, suburban Midwestern community in three school years between 2001 and 2005. My initial motivation for undertaking this analysis was to provide an "existence proof" — to document that geometry instruction *can* be taught in such a way that students are held accountable for learning the elements of the mathematical sensibility.

---

[24] It is probably a good idea to reveal here that the anonymous teacher referred to in this chapter is actually the author of the dissertation. However, throughout this work I use the device of referring to "the teacher" in the third person, in part to help preserve the traditional distinction between subject and object. For more on the challenges and affordances of self-study, see below, p. 193-199.

But beyond merely establishing that such a role for the mathematical sensibility is *possible*, I intend to examine *how* such a role can be carved out, and in particular to identify the constraints that threaten the viability of such a role in the ecology of the classroom.

Before elaborating fully on the theoretical perspectives and methods for this study, I begin with two motivating examples. An informal examination of these examples will bring into focus what I mean by one of my main claims in this chapter: that an assessment item can exemplify, or instantiate, one or more of the mathematical dispositions. The examples will also serve as an opportunity to introduce one of my key theoretical resources in this and the following chapter, namely the construct of the implied teacher. For the first example, consider the item in Figure 4.1:

In class I claimed that any quadrilateral has exactly twice as much area as its *dual* (remember, the dual is the figure formed by joining midpoints of adjacent edges). Prove this in the special case where the original figure is an isosceles trapezoid.
Figure 4.1. An assessment item.

Notice that, as written, this item includes more than just the statement of the task (which could easily have been reduced to "Prove that an isosceles trapezoid has twice the area of its dual"). It also includes an introductory sentence that serves as a motivation for the task: specifically, this task exists in order to confirm, at least partially, a claim that had been previously made but (it appears) was never proven in class. Thus the item includes alongside the task a kind of *commentary on the task*. This commentary communicates several things, simultaneously, to the student:

- that previously-made claims ought to be supported by a proof
- that the student need not remember the definition of the word “dual” (because it is provided in parentheses)
- that doing a proof in a special case can be worthwhile, even though it falls short of proving the general case

The commentary thus calls the student’s attention to two important mathematical dispositions: *Confirmation* and *Specialization*. In addition to this, the item’s text includes (via the use of the phrase “in class I claimed”) a memory of a prior classroom experience that is taken as shared; an explicit reference to the item’s author, i.e. the teacher; and a characterization of that as a certain kind of individual — one who makes claims, and comes back to them later to seek confirmation. Less explicitly, the very existence of the commentary indicates something about the teacher’s expectations of the student: namely, that a student in the midst of taking a test may be expected to care about more than just “what am I supposed to do now”.

The theory of systemic functional linguistics (SFL) (Halliday 1994; Martin & Rose, 2003) provides a language for discussing the way this (short) text operates. SFL identifies three distinct “metafunctions” of language: the *ideational* (what is being talked about), *interpersonal* (the interaction of “speaker(s)” and “audience”), and *textual* (the mode and organization of a text). From the point of view of the ideational metafunction, this is a text about the areas of certain quadrilaterals. From the point of view of the textual metafunction, this item makes a connection to a prior “text” (where here I use the word somewhat loosely to include prior classroom interactions that may not have been recorded in written form). But from the perspective of the interpersonal metafunction, this assessment item is a text about who the teacher is, and who the teacher expects the student to be. Seen from this perspective, it is as if the teacher is trying to say to the student that this problem needs to be done, in part, because of who the teacher and student are, and because of the relationship between them as enacted through ongoing interaction: in short, it says that “you” (the student) and “I” (the teacher) have unfinished business to take care of.

This self-insertion of the teacher into the assessment item, and its narrative quality, personalizes and temporalizes the question in a way that seems odd from the customary perspective that depicts mathematics as timeless and depersonalized (Brousseau, 1997, chapter 1). But perhaps this should not come as a surprise. Students do not, after all, do mathematics problems solely for the purpose of learning mathematics; they also do so because of a didactical contract that joins teacher and student in a network of explicit and implicit mutual obligations. Whenever a student does (or refuses to do) an assigned task, that action indicates compliance with (or rejection of) the roles that are institutionally mandated for the student and teacher.

This brief discussion serves to introduce two of the main themes of this chapter: that an assessment item can (through commentary and word choice) signal the importance of one or more of the mathematical dispositions; and that an item (through the textual and interpersonal metafunctions) can imply a specific teacher-role and student-role. I return to both of these points later in this chapter.

For my second motivating example, consider the problem in Figure 4.2. In this problem, students are presented with a diagram of what appears to be two parallel lines crossed by a transversal. Four of the angles formed by those lines are labeled with algebraic expressions in two variables, and the student is asked to determine whether or not the lines are actually parallel. Note first how differently this problem uses language in contrast to the previous example. Unlike the example in Fig. 4.1, this second problem contains no commentary, no personalization of the teacher, no reference to past experience. The interpersonal metafunction here appears to construct an entirely different relationship between teacher and student, one in which the teacher's individuality is elided altogether. Nor does the text contain any words that seem to clearly mark any of the mathematical dispositions of the previous chapter.

In the diagram below, are the lines parallel? Why or why not?

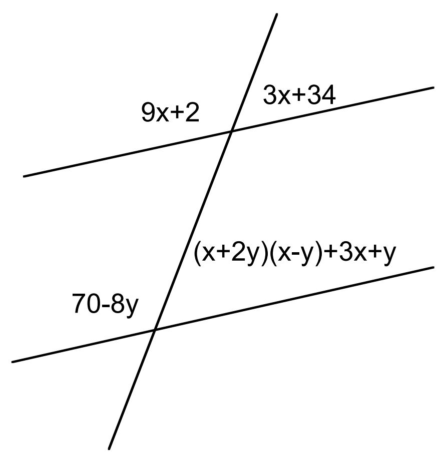


Figure 4.2. The *Possibly Parallel Lines* problem.

But it would be incorrect to conclude that this item says nothing about the mathematical dispositions, or about the roles of the teacher and student. As I will now show, that information can also be found through an analysis of the possible solution strategies that a student might pursue. This problem admits at least two solution strategies, either of which in principle would have been available to a typical student in a

Geometry class. One strategy, which I call the *direct approach*, is conceptually very straightforward but entangles the reader in some rather thorny algebraic manipulations. In this approach, the student notices that the two angles at the top of the diagram constitute a linear pair, and hence that

$$(9x+2)+(3x+34)=180$$

which, after some modest algebra, yields $x=12$; and hence the two angles in question are 110° and 70°, respectively. So far there the problem is fairly commonplace and not very interesting. Continuing in this fashion, one next turns to the lower pair of angles. Following the same reasoning (and making the substitution $x=12$) one obtains

$$(70-8y)+(12+2y)(12-y)+3\cdot 12+y=180$$

But whereas the reasoning that produces this second equation is the same as that behind the previous one, the algebra needed to carry this equation forward to a solution is much more difficult. If students are diligent and careful, eventually they would arrive at the quadratic equation

$$2y^2-5y-70=0$$

which has the two (irrational) solutions

$$y=\frac{5\pm 3\sqrt{65}}{4}.$$

After all of this, one determines the measures of the lower pair of angles to be $\left(60+6\sqrt{65}\right)°$ and $\left(120-6\sqrt{65}\right)°$. Finally one reaches the conclusion that the corresponding angles in the diagram are not congruent, and hence the two lines are not parallel.

So much for the direct approach, which is conceptually very straightforward but requires a great deal of algebraic facility (and entails a not insignificant risk of error). The second strategy, the *indirect approach*, is far simpler from a procedural standpoint, but conceptually subtler. In this approach, one begins as in the direct approach by using the top pair of angles to find $x = 12$ and then determining the upper angles to be 110° and 70°. At this point, however, rather than continue onwards to determine the value of y and the measures of the lower pair of angles, and hence determining whether the lines are parallel, one instead makes a *strategic hypothesis*: suppose that the lines *are* parallel. If so, then the lower left-hand angle must also be 110°; and from the equation

$$70 - 8y = 110$$

we quickly discover that $y = -5$. Next, we substitute the (known) values for x and y into the expression for the angle on the lower right, obtaining $(12-10)(12+5)+3\cdot 12-5$, which is easily evaluated to be 65°. We reach the conclusion that if the half-lines on the left side of the transversal are parallel, then the half-lines on the right side are not; or, equivalently, that the only way to make the left-hand pair of corresponding angles be congruent is to put a kink in the bottom "line".

In summary, the indirect approach almost entirely circumvents the need for the difficult algebraic manipulations required for the direct approach. Instead, the student makes a strategic hypothesis, deduces consequences from that hypothesis, and ultimately rejects the hypothesis on the grounds that it leads to a contradiction. In other words, the indirect approach calls for the student to engage in a kind of reasoning that corresponds to the method of *indirect proof*; it also captures well one of the heuristics a mathematician might use in attempting to discover whether something is true or not (Polya, 1957).

In terms of the mathematical dispositions of the previous chapter, note that the moment at which the strategic hypothesis is made could be characterized as a deployment of the move "consider the converse": Rather than seek the correct values of x and y to determine whether corresponding angles are congruent, instead one assumes a desired conclusion (that the two left-hand angles are congruent) and seeks conditions under which that conclusion would hold. In addition, the problem places students in a mathematical context in which the truth of the matter is ambiguous: it is not clear until one has finished the problem whether the lines are parallel or not. Precisely because of this, the indirect approach described above is a risky one: if one were to follow the steps outlined above but *not* reach any contradiction, one would not be entitled to reach any conclusion whatsoever, and would need to start over with some other strategy.

The problem thus places the student in a situation of ignorance somewhat analogous to that of the research mathematician, who genuinely may not know ahead of time whether the theorem he is attempting to prove is true or false, or whether the particular problem-solving strategy he pursues will bear fruit or lead to a dead-end. The "direct approach" solution navigates through the problem-space in a more conservative fashion: the student executing this approach makes no unfounded assumptions and thus takes no "risks". The cost of this conservatism, however, is a high level of procedural difficulty and a pronounced risk of error. Seen from thus perspective, the "Possibly Parallel Lines" problem can be understood as holding students accountable for being able to deploy the mathematical dispositions of making strategic assumptions, considering the converse problem, and seeking the conditions under which a conclusion holds as heuristics for navigating through uncertainty.

The second motivating example was intended to illustrate the fact that geometry problems at a relatively accessible level can nevertheless become opportunities for students to call upon rather sophisticated forms of reasoning. Of course, the fact that the opportunity *exists* does not mean that students will *exercise* it; there is, after all, always the direct approach. But the direct approach is costly, and students who pursue it face a kind of tax on their work.

Looking back across the two motivating examples, we see that an assessment item can represent an instance of a mathematical disposition in more than one way. On the one hand the item can *explicitly* mark a disposition as salient, either through word choices or through commentary; alternatively an item can *implicitly* mark a disposition as valuable, in the case where there is a "favored" solution in which such disposition is deployed. This distinction aligns with an important distinction from Chapter 2. Recall that practical rationality consists of both categories of perception and categories of appreciation, and recall also that I have included the mathematicians' "generative moves" among the categories of perception, in that they constitute labels for what a mathematician sees as actions that may be performed at a particular moment. The categories of appreciation, on the other hand, are labels that describe how a particular piece of mathematical work (including both problems and solutions) may be valued. In the two examples, we see that categories of appreciation are instantiated through word choice and commentary; these tell the student not only what he or she is to do, but why it is important or worthwhile that it be done. On the other hand, categories of perception (the generative moves) may be located in the analysis of the possible solutions of a problem.

Theoretical perspectives

The intentional fallacy and the "implied teacher"

In the discussion of the previous examples I made some informal comments about the teacher and his apparent intentions as revealed through text, and before proceeding it is important to address directly the role that such questions play in my analysis below. It is always risky to make claims about a teacher's intentions on the basis of what is observed to have happened, and in any event this study is not meant to be an analysis of an individual teacher, but rather of *teaching* as evidenced by a collection of artifacts. And yet the artifacts themselves are provocative, and it is somewhat natural to speak of them in terms of intention. For this reason, I begin my discussion of theoretical frameworks with a perhaps unusual turn to a discussion of some developments in 20th century literary criticism.

Prior to the mid-20th century, a standard element of literary criticism was an inquiry after the intention of the author, who was held to have primacy in determining the meaning of a text. Wimsatt & Beardsley (1946/1999) dub this the "intentional fallacy", and offer as an example of the fallacy Goethe's three questions for critics: "What did the author set out to do? Was his plan reasonable and sensible, and how far did he succeed in carrying it out?" (p. 483). The intentional fallacy was named, and rejected, by the founders of the so-called "New Criticism," who argued that the only meaning that matters, or that is even accessible to the critic, is that which can be found *in the text*, rather than in the mind of the author: "One must ask how a critic expects to get an answer to the question about intention. How is he to find out what the poet tried to do? If the poet succeeded in doing it, then the poem itself shows what he was trying to do" (Wimsatt & Beardsley 1946/1999, p. 481). From this point of view, while information about the author's life may be illuminating, this properly belongs to the realm of literary biography, rather than literary criticism.

But despite the rejection of the intentional fallacy, literary critics have still found it useful to introduce various surrogate notions to take the place of the author. In his 1961 work *The Rhetoric of Fiction*, the literary critic Wayne Booth introduced the notion of the *implied author* of a text. The implied author is a theoretical construct, distinct from both the actual author of a text and the narrator. When a reader encounters a text, that text is likely to suggest to the reader certain characteristics or qualities of its author, and thus an impression of the author is formed in the mind of the reader. This impression may correspond only approximately, or not at all, to the *actual* author who put pen to paper and constructed the text; and certainly it may be that the author implied by a text to one reader may be somewhat different from the author implied by the same text to a second reader. Still, despite the somewhat conjectural nature of the implied author, as a theoretical construct literary critics have found it a natural and useful way to discuss matters of intent without the need to make claims about an actual biographical individual— when, in any event, such claims may be difficult or impossible to substantiate.

Booth describes the implied author as "the picture the reader gets of [the author's] presence". Booth observes that

> It is a curious fact that we have no terms either for this created "second self" or for our relationship with him. None of our terms for various aspects of the narrator is quite accurate. "Persona," "mask," and "narrator" are sometimes used, but they more commonly refer to the speaker in the work who is after all only one of the elements created by the implied author and who may be separated from him by large ironies. "Narrator" is usually taken to mean the "I" of a work, but the "I" is seldom if ever identical with the implied image of the artist....
>
> It is only by distinguishing between the author and his implied image that we can avoid pointless and unverifiable talk about such qualities as "sincerity" or "seriousness" in the author.... we have only the work as evidence for the only kind of sincerity that concerns us: Is the implied author in harmony with himself — that is, are his other choices in harmony with his explicit narrative character? If a narrator who by every trustworthy sign is presented to us as a reliable spokesman for the author professes to believe in values which are never realized in the structure as a whole, we can then talk of an insincere work. A great work establishes the "sincerity" of its implied author, regardless of how grossly the man who created that author may belie in his *other* forms of conduct the values embodied in his work. (pp. 73-75)

By putting forth the construct of the implied author, Booth brackets as irrelevant any consideration of the biography of the *actual* author, while simultaneously enabling discussion of "intent" without the need for presumptuous and unverifiable claims about an actual person's state of mind. The implied author is of course himself a fiction, one co-created by the critic who recognizes traces in a text of what is presumed to be the text's creator. It goes without saying that different readers of the same text may encounter different implied authors there; however, insofar as the implied author has no qualities other than those that can be inferred on the basis of textual evidence, it is to be expected that any two implied authors of a common text will at least have much in common.

Booth goes on to identify the implied author's opposite number, namely the *mock reader*. Any text, Booth argues, presumes a certain kind of reader; it is this reader whom the implied author has in mind when creating the text. Booth quotes Walter Gibson in saying that “a book we reject as bad is often simply a book in whose ‘mock reader we discover a person we refuse to become, a mask we refuse to put on, a role we will not play” (p. 138):

> We may exhort ourselves to read tolerantly, we may quote Coleridge on the willing suspension of disbelief until we think ourselves totally suspended in a relativistic universe, and still we will find many books which postulate readers we refuse to become, books that depend on “beliefs” or “attitudes” — the term we choose is unimportant here — which we cannot adopt even hypothetically as our own.

A perspective similar to Booth’s is found in the work of the semiotician Umberto Eco. In *The Role of the Reader* (1979), Eco renounces “the use of the term /author/ if not as a mere metaphor for «textual strategy»” (p. 11). That is to say, for Eco the only “author” that can be spoken of with any definiteness is the author who we can identify through the traces he leaves in the text, the methods he employs for creating meaning. Elsewhere Eco (1994, p. 50) distinguishes between the *intentio operis* and the *intentio auctoris* (respectively, the intent of the work and the intent of the author). The actual, biological individual who constructed the text ceases to be relevant the moment the text leaves his hands and enters those of the reader: the *intentio operis*, the textual strategy, is Eco’s analogue for Booth’s implied author. Corresponding to Booth’s mock *reader* is Eco’s own notion of the *Model Reader*. The Model Reader is a model of the possible reader, one presumed by the author in construction of a text. Notice, though, that in saying that the Model Reader is “presumed by the author”, it is important to keep in mind that by “author” is meant only a textual strategy, not an actual flesh-and-blood person. Thus both Eco’s “author as textual strategy” and “Model Reader” are, like Booth’s "implied author" and "mock reader", purely hypothetical constructs, descriptions of a presumptive author and reader that may differ substantially from the *actual* author and readers of a given work. These constructs provide a metaphor-rich language for discussing the strategies employed by a text in terms of intention, without requiring (or even tolerating) any claims about the author's real desires.

Booth describes the experience of rejecting a book as bad as the result of a mismatch between the book's mock reader and its actual reader; Eco takes this point in a more radical direction, arguing that a successful text will *transform* the actual reader (in Eco's parlance, the *empirical reader*) into the model reader. Eco illustrates this claim with an example from Sir Walter Scott's novel *Waverley*:

> The author of *Waverley* opens his story by clearly calling for a very specialized kind of reader, nourished on a whole chapter of intertextual encyclopedia.... But at the same time [the text] *creates* the competence of its Model Reader. After having read this passage, whoever approaches *Waverley* (even one century later and even — if the book has been translated into another language — from the point of view of a different intertextual competence) is asked to *assume* that certain epithets are meaning «chivalry» and that there is a whole tradition of chivalric romances displaying certain deprecatory stylistic and narrative properties.
>
> Thus it seems that a well-organized text on the one hand presupposes a model of competence coming, so to speak, from outside the text, but on the other hand works to build up, by merely textual means, such a competence. (pp 7-8).

In the analysis of a corpus of mathematics assessment items, or indeed any artifacts of teaching, I propose that it is valuable to speak of analogues of the implied author and mock reader — two constructs I will refer to as the *implied teacher* and the *implied student*[25]. The implied teacher is the observer's answer to the question, "What is the intention of this teaching?" The implied student is the answer to the question, "What is a student held accountable for being capable of doing here?" In asking and answering these questions we must scrupulously avoid making claims about the *actual* teacher and students (following Eco, I will henceforth refer to these as the *empirical teacher* and *empirical students*); to do so would be to fall victim to the intentional fallacy (Wimsatt & Beardsley, 1946/1999)[26].

In particular, Eco's perspective on texts — in which the function of a text is to transform the empirical reader into its Model Reader, by creating in the empirical reader the competencies needed to understand the author's message — takes on a particular salience in the case of written instructional materials. That is to say, the purpose of such materials is *explicitly* to build in each and every empirical student the capacities of the implied student. In the present context, it amounts to the claim that assessment items constitute an opportunity to learn.

Assessments in mathematics education

In the preceding pages, I have suggested that the constructs of the implied teacher and implied student may be useful for describing a collection of assessment items and what they may say about teaching. Such constructs can make possible a description of

---

[25] I use "implied student" rather than "mock student" or "model student" because the latter two have unintended connotations.

[26] In this connection, it is worth noting that teacher educators who make use of videocases have in some cases noted that preservice and inservice teachers may be reluctant to make judgments about a teaching episode on the grounds that it is impossible to evaluate such an episode without knowing the teacher's intentions — a version of the intentional fallacy.

the didactical contract, in terms of what the teacher does (and does not) hold his students (and himself) accountable for. But it must also be stressed that while the same perspective could fruitfully be brought to bear on a study of other teacher-authored documents —homework assignments, in-class worksheets, unit review guides — the texts in question for the present study are none of those: they are, specifically, assessment items, and as such constitute a particular *genre* of instructional text. Just as literary genres (e.g. Western, Detective, Science Fiction, Romance) are defined by certain normative expectations (conventions), so too are assessments characterized by certain conventional features. I turn now to a review of the literature on the functions of assessment in mathematics education.

There is an extensive and deep literature on assessment in mathematics education. Kulm (1990) notes that assessment is a tool with multiple instructional uses, including diagnosis, monitoring, and evaluation of student learning, and that "in recent years... the evaluation and comparison of achievement between groups of students, school districts, states, and nations have become the central focus of assessment" (pp. 1-2). In addition to these functions Kulm lists as primary purposes of assessment "feedback for the student" and "communication of standards and expectations" (p. 4). Kulm's focus is on the twin questions of *what* should be assessed, and *how* it can be assessed. A crucial component of Kulm's analysis is his contention that not all *forms* of assessment are equally well-suited for assessing different kinds of knowledge. While the traditional in-class written examination may suffice very well for measuring students' computational and procedural fluency, alternative forms of assessment — including oral presentations, group project work, open-ended take-home assignments, and student portfolios — may be better able to provide insight into students' higher-order thinking and problem-solving skills.

Niss (1993) frames the purposes of assessment as: (1) provision of information (to the student, to the teacher, to the system), (2) the establishing of bases for decisions or actions (filtering and selecting individuals for various functions, licensing, ordering, etc.), and (3) the shaping of social reality (disciplining students, teachers and institutions; subordination to power, authority, and ideology; etc.). In addition to the *purpose* of assessment, Niss describes an assessment as a metaphorical "vector" whose components include the *subject* (who is assessed), *object* (the content and ability being assessed), *items* (the kinds of output), *occasions, procedures and circumstances*, *judging and recording,* and the *reporting of outcomes.* With respect to the second of these, Niss notes that "In an increasing but limited number of cases, objects of assessment include... *heuristics and methods of proof...* [and] *problem solving*" but that "we rarely encounter... *exploration and hypothesis generation* as objects of assessment (p. 15)." Niss also distinguishes between *continuous assessment* and *discrete assessment*: the former are generally of the "formative" type (in that they provide feedback to the teacher at the same time that the teacher is engaged in instruction) while the latter are typically of the "summative" variety.

The authors of the various chapters in the volume edited by Lesh & Lamon (1992) echo Kulm's concern that the *form* of assessment be well-suited for the *content* that is to be assessed. In that volume, the editors (Lesh & Lamon, 1992, p. 6) stress that "the authors are not simply concerned about developing new *modes* of assessment. They are primarily concerned about changing the *substance* of what is being measured." Goldin

(1992), for example, discusses five versions of the same question (see Fig. 4.3 below), and shows how seemingly minor variations in phrasing can result in assessments of very different kinds of knowledge. To be specific, Goldin's "Problem 3" assesses the extent to which a student is familiar with the notion of "average" or "mean", as well as his or her understanding of the properties "commutative" and "associative" as applied to binary operations, and finally the student's ability to transfer that knowledge to an unfamiliar context in which the truth is in doubt. Problem 3a removes the dependence on prior acquaintance with the mean by including an explanation, but at the same time "it requires more than the transfer of the concepts of commutativity and associativity to a new domain. It requires, and assesses, the student's ability to *construct* the new domain from the given verbal description of a new operation." (p. 79, emphasis added).

Problem 3. Let the symbol @ stand for the *average* or *mean* of two numbers. For example, we shall write 6@8=7, because 7 is the mean of the pair 6 and 8. Is the operation @ commutative? Is it associative? Explain why or why not.
Problem 3a. Let the symbol @ stand for the *average* or *mean* of two numbers. This is found by adding them and dividing their sum by 2. For example, 6@8=7, because 6+8 is 14, and 14 divided by 2 is 7. Is the operation @ commutative? Is it associative? Explain why or why not.
Problem 3b. The operation of addition (+) is *commutative* because when two numbers are added, their sum is the same in either order. For example, 6+8=14 and 8+6=14. Addition is also *associative*, because when three numbers are added it does not matter which pair is added first. For example, (6+8)+2=14+2=16, while 6+(8+2)=6+10=16. Now let the symbol @ stand for the *average* or *mean* of two numbers. This is found by adding them and dividing their sum by 2. For example, 6@8=7, because 6+8 is 14, and 14 divided by 2 is 7. Is the operation @ commutative? Is it associative? Explain why or why not.
Problem 3c. Let the symbol @ stand for the *average* or *mean* of two numbers. For example, we shall write 6@8=7, because 7 is the mean of the pair 6 and 8. Give an example (using two numbers) which illustrate the commutative property of the operation @. Give an example (using three numbers) to show that @ is not associative.
Problem 3d. Let the symbol @ stand for the *average* or *mean* of two numbers. This is found by adding them and dividing their sum by 2. For example, 6@8=7, because 6+8 is 14, and 14 divided by 2 is 7. Draw a number line and, using two numbers as an example, show the meaning of @ with a diagram on the number line. Then explain what your picture suggests about whether @ is or is not commutative.
Fig. 4.3. Four variations on an assessment item. After Goldin (1992).

Similarly, Problem 3b removes the dependence on prior acquaintance with the commutative and associative properties; Problem 3c restores the need for prior knowledge, but both removes the uncertainty inherent in the original Problem 3 and replaces the call for an explanation with a more modest call for an *exemplification*; and Problem 3d calls for a particular kind of pictorial representation that may or may not have been part of a student's response to the other four variations.

In Chapter 5, I perform a similarly fine-grained analysis on a number of exam questions from the corpus of data under examination. In particular I will show how certain exam questions that were repeated on tests in each of the three years of the data sample underwent subtle changes in phrasing and emphasis that amounted to changes in what was at stake for students.

Much of the literature on assessment is concerned, quite properly, with what we might call the *retrospective* function of assessments. That is to say, the purpose of assessment is understood as being the measurement (or other characterization) of what has happened up to the present: what a student has learned, what he remembers of what he has learned, and what he can do with what he remembers of what he has learned. But as the example from Goldin suggests, assessment items can have a secondary, *prospective* function as well. The student who finds Problem 3b on an assessment is likely encountering the commutative and associative properties for the first time (at least, the problem does not hold them accountable for having seen or understood it previously); but it may not be for the *last* time. On the contrary, the introduction of these concepts on the assessment may set the stage for work that will be done in subsequent days or weeks. In this way examination questions can foreshadow material from the coming chapter;

they can introduce new vocabulary and definitions that will be used in subsequent classroom work. From this point of view, assessment items do not only measure what *has been* learned; they also constitute an *opportunity to learn* new material (Bell, Burkhardt, & Swan 1992a, 1992b; de Lange 1992; Chazan & Yerushalmy 1992).

Reconstructing practice from assessment items

As has been noted above, the literature generally regards assessment items as instruments used to discover information about students. From this point of view the value of an assessment item is only realized when students are actually assessed, and the informational value of the assessment resides within the students' responses to it. In this chapter, however, I take a quite different position with regard to the analysis of assessment items: I look to these items for the purpose of a *post facto* reconstruction of classroom instruction itself. That is, in the absence of other data sources (e.g. field records of instruction) that may contribute to our understanding of what and how a particular teacher taught, I contend that a corpus of examination questions can provide some insight into the classroom practice of the teacher by answering the question, "What did the teacher hold his or her students accountable for learning?" In other words, I use the assessment items not to make measurements on or describe the knowledge of *students*, but to reflect back on the teaching practice, as evidenced by the textual strategies — the implied teacher of the assessment.

This perspective rests on some assumptions about teaching that derive from an understanding of the didactical contract (Brousseau, 1997). Because education is both institutionalized and compulsory, students and teachers hold certain (tacit) expectations of one another. One of those expectations is that teachers prepare students for assessments by ensuring that there is some close correspondence between what is taught and what is assessed. When a teacher asks, on an exam, a question that students (individually or in the aggregate) are not capable of answering, the contract is breached: either the students have failed to live up to their responsibilities, or the teacher has failed to live up to his. In either case we can expect that an institution would respond, and that either the teacher or the student would be expected to make changes going forward.

For this reason a corpus of exam questions, taken from a single teacher, can be taken as evidence of what the teacher intended to teach, or (more precisely) as a proxy for what students were intended to learn. What students *actually* learned, of course, is an entirely different matter, and to study that one would need access not only to the questions asked, but also to the responses students generated. Moreover, even this would not tell us the whole story. Students do not take examinations in isolation; on the contrary, teachers often interact with their students during exams, responding to questions, clarifying instructions, giving feedback and in some cases offering hints to those who are struggling. In this manner a students' experience of working on an examination question may include scaffolding not present in the written text of an item. It may even be that challenging examination questions undergo a degradation analogous to that documented by studies of mathematical tasks in classrooms. Such studies have shown that curricula intended to engage students in high-level reasoning tend to be deployed in ways that reduce the conceptual demands to the level of merely procedural proficiency (Stein, Smith, Henningsen & Silver, 2000).

Despite this, there is value in looking closely at curricular tasks in their own right, disconnected from the work students do and the interaction between teacher and students in the deployment of the tasks. Earlier studies of curricular tasks have largely focused on textbook exercises rather than an exam questions, but some of the same considerations certainly apply here. Mesa (2004) writes that textbooks are an expression of the *intended curriculum*; the same is true of examinations, which embody what students are supposed to have learned from each curricular unit, and what the teacher believes he or she actually taught. Mesa describes the study of textbooks as a hypothetical exercise:

> What *would* students learn if their mathematics classes were to cover all the textbook sections in the order given? What *would* students learn if they had to solve all the exercises in the textbook? *Would* they learn the particular mathematical notions that are presented in the textbook? *Would* that learning work well in their future work in mathematics? (Mesa, p. 256; italics in original)

The analogous questions about examinations (e.g., What *would* students have had to learn in order to answer this question successfully? What new mathematics *would* a student learn if they were to answer this question?) invites us to investigate the tasks in much the same way that Simon (1975) explores the problem space of the Tower of Hanoi problem. In that analysis, Simon performs an *a priori* analysis of a problem and describes multiple strategies that might be brought to bear on its solution. Each strategy is in turn analyzed for the kinds of demands it makes on one's visual perception and short-term memory. In this way his analysis shows that different individuals may have equal success in solving the Tower of Hanoi problem, and yet learn very different things from the experience. My analysis above of the Possibly Parallel Lines problem (Fig. 4.2) is in this same spirit: it shows that a student might solve the problem by persistently applying algebra techniques to a somewhat complex problem, or alternatively might reduce the algebraic complexity by deploying some of the mathematical dispositions. The implied teacher and implied student can be regarded as the product of carrying out the hypothetical exercise described in the passage quoted above. As such they amount to a retrospective characterization of the *intended* curriculum— they describe what was supposed to have been taught and learned.

Beyond the value that lies in examining individual assessment items to see what they tell us about the intention behind them, there is additional evidence in the corpus that allows us to draw some inferences regarding the alignment between what was *intended to be taught* (as indicated by the questions) and what was *actually taught and learned* (as would be indicated by students' responses). This evidence exists because the corpus under examination contains not a single assessment or even a single *year's* worth of assessments, but a multi-year corpus of examination questions written by a single teacher, in which many questions are repeated in successive years. For this reason we may take the position that the teacher, in giving and scoring assessments, receives *feedback* from his class that could inform his behavior in subsequent years. For example, if the teacher were to discover through grading exams that there was a mismatch between what he thought had been taught and what students actually learned, it is likely that the implicated question would disappear from, or undergo modification in, subsequent years' examinations — or alternatively that the teacher will take greater pains in subsequent years to prepare students better to answer the question. On the other hand if an assessment item "succeeds" (i.e. if it produces roughly the kind of response the teacher was hoping for) then it is likely to reappear in subsequent years with little or no modification. For this reason in the next chapter I will look not only at individual items on single assessments, but at the *history* of each assessment item, for evidence of how viable it proved to be in the ecology of the classroom. A question that continues to appear on exams in multiple years may legitimately be taken as a reflection of the teaching that preceded it. On the other hand a question that drops out of the corpus after one year, or recurs with modifications, may be taken as evidence of some kind of mismatch between what the teacher hoped for and what was actually achieved.

To make this more concrete, suppose a teacher were to include Goldin's Problem 3 (Figure 4.3) on an exam in one year, and in the following year were to use the variation Problem 3a instead. The thrust of my argument here is that such a *change* in the problem would testify to more than would a consideration of the individual variations as separate items. It would be reasonable to infer, on the basis of such a change, that the teacher's assumptions in the first year (e.g. that students already were familiar with the notion of *average*) proved not to be true, so that in the second year the teacher found it prudent to provide additional scaffolding for those students who needed it. At the same time the fact that the teacher provided the scaffolding in 3a, but not the scaffolding in 3b, suggests that either the teacher consistently intended to hold students accountable for knowing the algebraic notions of "commutativity" and "associativity" and was therefore unwilling to provide definitions for these concepts, or alternatively that the results of the first year's exam showed that students were generally capable of demonstrating knowledge of these properties and therefore the teacher did not find it necessary to provide additional scaffolding.

Coherence and adaptation in the assessment items corpus

More generally, one can look at the evolution of a single assessment item across multiple years, and — by noticing what changes and what does not — draw some reasonable inferences about the kinds of feedback that a teacher experienced in the use of the task. I return to this theme in Chapter 5, in which I discuss adaptation in the corpus across multiple timescales.

All of the foregoing discussion relies in part on the presumption that it makes sense to regard a corpus of assessment items as a *text*. The legitimacy of this presumption could itself be challenged. From one point of view, such a corpus resembles a novel or biography far less than it does, say, a recipe book, or the classified advertisements in a local newspaper. That is, assessment items are typically short (on the order of at most a few sentences) and independent of one another. The questions in most exams can be done in any order, and the content covered on one exam may have little in common with that covered on another. One would not expect to find much of an implied author in a heterogeneous collection of independent, disconnected paragraphs; on what grounds is it legitimate to approach the corpus of assessment items in the way have I described?

To answer this question I turn again to the linguistic theory of *systemic functional linguistics,* or SFL (Halliday 1994; Martin & Rose 2003), and in particular to the theory of *cohesion* (Halliday & Hasan 1976; Hasan 1984). Recall that SFL identifies three distinct "metafunctions" of language: the *ideational* (what is being talked about), *interpersonal* (the interaction of "speaker(s)" and "audience"), and *textual* (the mode and organization of a text). The key notion of SFL is that any text can be analyzed for how it uses language in the service of these three metafunctions. *Cohesion*, meanwhile, refers to the use of lexical and grammatical relationships to produce a coherent text, that is, to hold a text together and give it meaning. Cohesion is, so to speak, the means by which the end of coherence is achieved. Thompson and Zhou (2001) note that most studies of cohesion focus exclusively on the use of language to signal logical connections among propositional content, and thus stress the ideational and textual functions of cohesion. However, they point out, cohesion can function along the interpersonal axis as well:

> Halliday (1994: 338) mentions that certain types of cohesive resources involve interpersonal rather than ideational meaning, but he does not explore the implications. In general, the important concept of evaluative coherence — the way in which, for example, writers work to convey a consistent personal evaluation of the topic they are dealing with ' has received little attention, and there has also been little investigation of the corresponding role of the evaluative lexis in creating cohesion (though see Hunston 1989, 1994). What we wish to do… is to explore one of the ways in which explicit evaluation in a text functions to create texture and structure. (p. 123)

In particular Thompson and Zhou study the use of disjuncts to create an evaluative stance. Evaluation is regarded as part of the interpersonal metafunction, in that it conveys the author's opinions about what is written. But, they argue, disjuncts also function as *cohesive signals*, tying a text together. The continuity of an author's evaluative stances is one of the elements of a text's coherence.

More generally, cohesion can serve the purposes of any or all of the three metafunctions: that is, cohesion can exist at the ideational level (e.g. repeated references to the same thematic content), the interpersonal level (e.g. repeating a phrase that characterizes the relationship between the author or reader, or the author and the text), or the textual level (e.g. comments such as "As we saw in the last chapter..."). Indeed the recognition of something as "a text" (rather than as a disconnected collection of words) rests on cohesion along all three metafunctions. When such cohesion is lacking, the "text" will lack coherence. A page of classified advertisements from a newspaper has some ideational coherence (in that many of the advertisements are for the same or similar products) but essentially no *textual* coherence (in that typically no advertisement refers to other advertisements on the same page, even indirectly) or *interpersonal* coherence (in that there is no consistent authorial "voice" that ties advertisements together). This is not to say that all three metafunctions are not present in the individual ads — certainly advertisements may position the buyer and seller in a mutual relationship — but only that there is no cohesion that can create coherence across the entire page. For this reason it would be unnatural to speak of the page as a whole as a single "text" with a (single) implied author — although it may make perfect sense to speak of *each individual advertisement* as a text with its own implied author.

Thus one way of justifying my treatment of this corpus of exams as a "text" is to identify cohesion within it: that is, to locate questions in which the author makes explicit self-reference, and questions that refer back to other questions on the same exam, or to previous exams, either by revisiting thematic content or by making explicit reference to past experience. One of the tasks to be undertaken in the subsequent analysis, then, is to document that such cohesion exists in the corpus under examination.

The problem of self-interpretation

The framework described above — in which "author" and "reader" are taken to mean "implied / model author" and "mock / model reader", with an explicit disavowal of any claim about the empirical author and reader — is complicated when the empirical reader of a text — the one forming the interpretations — happens to coincide with the empirical author of the text. In this case, we must contend with some difficult questions: Is the (actual) author of a text capable of judging the meaning of a text? Or is there a kind of conflict of interest that precludes an author from an unbiased reading of "his" own text?

On the other hand, one might approach the situation from the opposite end: Does the empirical author stand in a *privileged position* when it comes to interpreting the meaning of that text? The author, after all, can claim, in a way that no others can, to know what was meant by a particular phrase, or why a certain character was given a specific name, and so forth. Other readers must content themselves to make do with the "traces" of the author left behind in the text, but the author himself has the inside story, so to speak. Does the author's own interpretation of his text have any particular preeminence?

Eco deals directly with this question in "Between author and text" (1992). Writing from both the perspective of a philosopher of language and semiotics, and from the vantage point of an author of novels (most notably *The Name of the Rose* and *Foucault's Pendulum*), Eco addresses the question, "Can we still be concerned with the empirical author of a text?" Eco enumerates more than a dozen instances in which readers of his own novels have "discovered" references and allusions that, as author, he "knew" were not intended. He discounts some of these "findings" as unconvincing and false. Nevertheless in other cases he concedes that the text he produced supports the contention that its "Model Author" (i.e., an explicit textual strategy) intended something that, as empirical author, he knows was not designed**.** In the end he concludes that "there are... cases in which the empirical author has the right to react as a Model Reader" (p. 79), but reaffirms that even in this instance the empirical author possesses no privileged position from which to interpret: "I have introduced the empirical author in this game only in order to stress his irrelevance and to reassert the rights of the text" (p. 84).

Eco's discussion of the interpretative difficulties that accompany the situation of an empirical author who acts as a Model Reader gains particular significance for us in the context of the teacher-researcher tradition. The teacher-researcher is both an observer of and a participant in the activity under investigation, a position that suggests the possibility of gaining valuable insight into teacher's decision-making that would be otherwise unavailable to researchers, while simultaneously raising questions about subjectivity, bias, and conflict of interest. Of course, all research is subjective to a degree; but the complexity of disentangling oneself from the phenomena under investigation is much greater when one stands squarely in the middle of it. In a survey article on teacher research, Cochran-Smith & Lytle (1990) discuss the benefits that such research affords, as well as the risks that are inherent to it. In particular, Cochran-Smith & Lytle argue that teacher research should be regarded as something fundamentally distinct from traditional research on teaching:

> Regarding teacher research as a mere imitation of university research is not useful and ultimately condescending. It is more useful to consider teacher research as its own genre, not entirely different from other types of systematic inquiry into teaching, yet with some quite distinctive features. But it is also important to recognize the value of teacher research for both the school-based teaching community and the university-based research community. (p. 4)

It is interesting to note that Cochran-Smith & Lytle do not directly address the problems of subjectivity, bias, and conflict of interest. In a subsequent review article a decade later (Cochran-Smith & Lytle 1999), the same authors elaborate further on the distinctions between teacher research and other research on teaching. They note that some scholars have critiqued teacher research on grounds that Cochran-Smith & Lytle refer to as *methods critiques*. These methods critiques deal directly with the questions of bias and subjectivity. In this respect they cite Huberman (1996), who argues that

> understanding events when one is a participant in them is excruciatingly difficult if not impossible, thus challenging the possibility of the teacher functioning as a researcher in his or her own school or classroom setting.... Huberman argues that the fact that teachers may have intimate insider information about teaching does not negate the need for research methods that are 'minimally reliable' in order to safeguard against 'delusion and distortion'. (Cochran-Smith & Lytle, p. 20)

In contrast to Huberman's concern for documentation and reliable research methods, Cochran-Smith & Lytle also cite Berthoff (1987), who proposed a view of the teacher as what she called "RE-searcher", who "did not need more 'findings'... but more dialogue with other teachers that would generate theories grounded in practice". (Cochran-Smith & Lytle 1999, p. 15). From this perspective, the teacher has access to a unique "data set" – namely, his or her own first-hand experiences – which is no less valuable than the more objectivized "data" that research normally privileges.

These theoretical concerns find their echoes in the writing of individual teacher-researchers. Wilson (1995), reflecting on her own experiences as a teacher-researcher, concludes, "I think learning to do research made me a better teacher." And yet, echoing the "methods critiques", Wilson does acknowledge that "the limitation and problems of such research deserve critical attention," adding (in a footnote) in particular the need to ask the questions: “What are our blind spots, places where our defensive selves make it impossible for us to ‘see’ clearly?” (pp. 20-21).

Ball (2000) discusses teacher research through an analysis of four case studies: her own work, and the work of Magdalene Lampert, Ruth Heaton, and Martin Simon. Ball argues that such inquiry, which she refers to as “working on the inside” and as “first-person research”, has struggled to find its place within the larger spectrum of education research. As Ball says in the opening sentence of her essay, such research presents both unique benefits and serious pitfalls. Although Ball does not focus her attention explicitly on the problems of bias and subjectivity in such first-person research, her analysis suggests that one question that must be asked is, “Can a teacher-researcher accurately represent him- or herself after the fact?” It is difficult enough — perhaps not even fully possible — for *any* researcher to disengage his or her personal feelings and opinions when observing teaching; how much more so is it difficult for the teacher-researcher to remain objective when holding the mirror up to his or her own practice?

Perhaps in response to this tension, teacher-researches have used various mechanisms to create distance between their “research selves” and their “teacher selves”. Ball recalls how in her own dissertation (Ball 1988), she described a teaching episode from her classroom, in which she described herself throughout as “the teacher” without ever identifying the lesson as her own experience:

> The description of the lesson continues for almost four pages, with detailed account of the discussion, the teacher’s moves, and the children’s work. The section ends with a commentary on the mathematics lesson, but nowhere do I reveal that I am using my own practice as a source for understanding the nature of teaching, that I am “the teacher”. (p. 372)

Such *third-person masking* is a strategy for coping with the conflicts that are intrinsic to the dual role of the teacher-researcher. Ball is not alone in her use of third-person masking; to take another example, Steffe (2004) includes a detailed analysis of a teaching experiment, in which "the teacher" is referred to throughout in the third person; a reader might read the entire paper and never come to the realization that "the teacher" is, in fact, Steffe, the author.

A fascinating variation on this strategy is employed by Heaton (1994). As Ball describes Heaton's own dissertation research, "Heaton invented a methodological strategy that afforded her separation from herself by using different voices to represent different points on her trajectory: Ruth 1, the self that was doing the teaching at the time; Ruth 2, the self that was reflecting on that teaching at the time; and Ruth 3, the self that made sense 3 years later of the teaching she did **during** that year." This artifice, the trifurcation of the self into multiple personae, helps to signal the various roles that a single individual may play. It would be overreaching to claim that this tactic, or other uses of the third-person masking strategy, actually *solves* the problems of bias and subjectivity; rather it attempts to ameliorate those problems by *flagging* them and making them explicit.

My introduction of the construct "implied teacher" can be regarded as another instance of third-person masking. In my view, however, the implied teacher (and his opposite number, the implied student) are not "masks" for the empirical teacher and student; they are *entirely independent constructs* with a (hypothetical) life entirely their own. As I seek to describe the intentions and values of the implied teacher, it is important for me to stress that *no claims of my own intentions or values as empirical* *teacher have a role in this analysis.* With the distance of time (some of this teaching occurred seven years before the writing of this analysis), in the absence of a journal or lesson plans, and with no access to students' work, it would be impossible to make credible claims about what I actually intended or what students actually learned. The most that can be claimed is that the documents themselves imply a certain kind of teacher, who in turn anticipated a certain kind of student. It is to these documents that I now turn.

Data sources and methods

The remainder of this chapter, and the next, are devoted to an analysis of a corpus of examination questions that was constructed by a teacher of Honors Geometry in a small, private high school in an affluent, suburban Midwestern community in three school years between 2001 and 2005. The corpus of questions is, as will be shown, unusual in a number of respects. In part because of the small size and young age of the institution at the time (the school began operations with only 52 students in the year 2000), and in part because in each year there was only a single Honors Geometry class, the teacher was given nearly complete freedom to develop his curriculum as he saw fit, without regard to any external mandates — although there was a preselected textbook (Larson, Boswell, & Stiff, 2001). This freedom is significant, in light of the fact that the teachers in the ThEMaT study groups (see Chapter 3) frequently cited institutional constraints as reasons why it is not viable for them to hold students accountable for learning the mathematical dispositions: it suggests that in this class the possibility, at least, may have existed to cultivate in students the elements of a mathematical sensibility.

The corpus at many timescales

Instruction is a complex phenomenon that can be examined over many different timescales (Lemke 2000): the task, the lesson, the unit, the year, and so on. Each timescale brings different information to light and allows asking different questions. In the present study, I analyze the corpus of assessment tasks at four different scales:

(i) The scale of the *individual item*, as it might be encountered by a single student, one at a time. How does the individual item represent an opportunity to learn elements of the mathematical rationality?

(ii) The scale of an *entire assessment*. Students do not encounter assessment items singly, but in the form of quizzes, tests, or projects that typically bundle many questions together. Each such bundle is located at a particular point in the trajectory of the school year, and there are certain descriptive features that are associated to the assessment as a whole: How many questions did it have? How much time did students have to work on it? What resources (notes, calculators, etc.), if any, did students have access to? A description of the corpus at this level will allow us to see if there are any long-term changes in the basic accountability structure of assessments in the corpus.

(iii) The scale of the *single year*. Within each year, there may be certain items that refer explicitly back to problems on earlier assessments, or hint at problems that may appear on future ones. Such "item chains" create textual cohesion (Halliday & Hasan 1976) across the year. Cohesion of this sort helps justify the stance that a full year's set of examinations constitutes a single "text" with a coherent implied teacher and implied student.

(iv) The scale of the *full corpus*. Many questions from the first year of the corpus reoccur with little or no change in the second and third years, while others undergo modifications or drop out entirely. Thinking of the history of a single item as it evolves from year to year helps to foreground the way in which the teacher's initial intentions were shaped and redirected by experience. From this perspective, the appropriate unit of analysis is not the individual item or the item thread, but rather an equivalence class of variant versions of a single question across the three-year corpus; I will refer to this as an *item class*. Cohesion across item classes realize the coherence of the implied teacher: the repetition of items (with or without alterations) serves to enable the identification of two tests in different years as nevertheless "the same" in some sense. (This cohesion is not visible to the implied student, because a successful student would not return to take the exams again in subsequent years). Examining item classes also provides some (admittedly indirect) evidence for how well the enacted curriculum aligned with the teacher's intentions in a given year, in that changes within a class can be regarded as adaptations, and questions that are stable across multiple years can be regarded as "successful" from the point of view of the teacher.

These different grainsizes are represented schematically in Fig. 4.4, which represents the corpus at different scales. In Fig. 4.4, the dark horizontal regions represent the three years of the corpus; within each year, the white rounded rectangles represent assessments, and the small dark circles represent individual assessment items. The dark horizontal arrows linking each assessment to the next signifies the evolution of the corpus as it is experienced by the implied teacher (for whom each new assessment constitutes an opportunity to deploy lessons learned in the previous); the thin jagged lines linking items within a single year represent ideational cohesion from the perspective of the student in a single given year; and the vertical lines represent item classes.

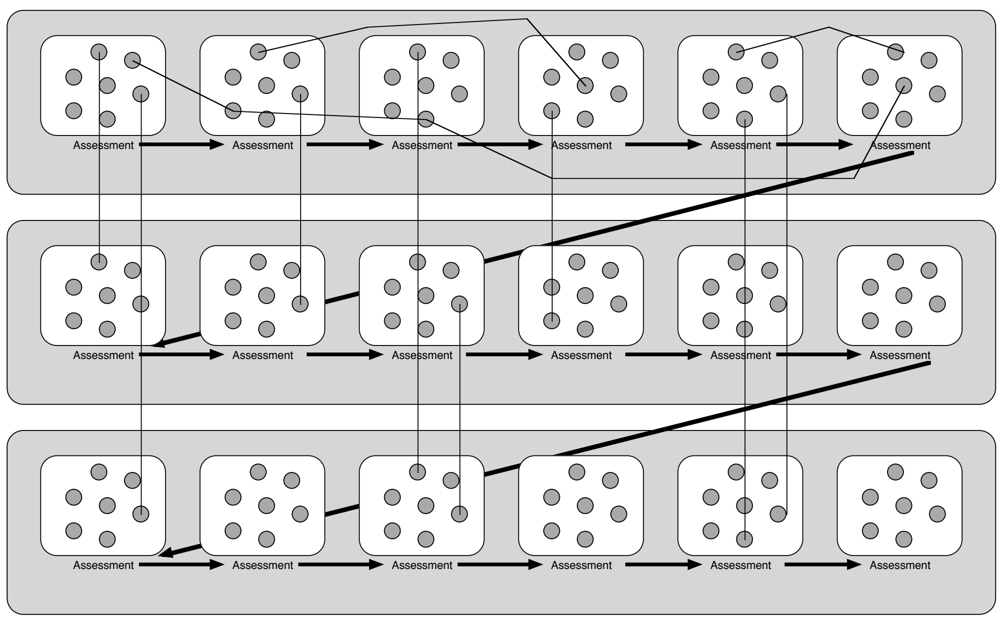


Figure 4.4. The corpus at many timescales.

One fairly coarse-grained way of representing the corpus is with a table, as in Fig. 4.5. This view of the data tells us nothing about the content of the individual questions, but does give some sense of the size and arrangement of the corpus, and certain large-scale patterns are already evident. Reading across the rows, we can note that the number of questions per exam consistently declined over the three-year history of the archive.

Additionally, the scope of the material covered drops from year to year. Moreover, the use of take-home problems as part of an examination — introduced in Year 2 solely for the midterm and final exam — becomes, in Year 3, part of the normal characteristics for an exam. Most strikingly, the use of regular chapter exams seems to have disintegrated entirely by the middle of Year 3. These phenomena will be discussed in more detail in Chapter 5.

	Number of items		
	Y1 (01-02)	**Y2 (03-04)**	**Y3 (04-05)(i)**
Chapter 1	22	17	14
Chapter 2	18	17	17
Chapter 3	17	14	12
Chapter 4	12	8	8 (iii)
Chapter 5	(ii)	11	12 (iii)
Midterm	31	10 (iii)	9 (iii)
Chapter 6	15	11	
Chapter 7	(ii)	(ii)	(iv)
Chapter 8	15	10	(ii)
Chapter 9	11	(ii)	(ii)
Chapter 10	18	(ii)	(ii)
Chapter 11	7	(iv)	(iv)
Chapter 12	(ii)	(iv)	(iv)
Final exam	27	18 (iii)	21 (iii)

Figure 4.5. An overview of the corpus.
Notes. (i) The exams for chapters 1, 2 and 3 in Year 3 each came in two versions(designated A and B) that differed only superficially (e.g. different sequences of problems, different numbers and equations, etc.). To avoid double-counting these items, only the Version A items were analyzed. (ii)No exam given; material from this chapter was included on a subsequent exam. (iii) Included a significant take-home component. (iv) Material not covered in this year.

Individual Items

A relational database containing the text and relevant images for each assessment item, along with metadata locating the item in the overall corpus, was created using Filemaker Pro database software. In transforming the corpus into a database, it was necessary to give some consideration to the question of what constitutes an *individual item*. As is customary on assessments, each examination question is prefaced with a number indicating its sequence in the test; thus we speak of Question #1, Question #2, etc. This enumeration creates a default division of the corpus into individual items. However, a difficulty emerges when we consider questions with multiple parts. Should such a question be entered as a single item, or should each of its constituent parts be entered as a separate item?

As an operational decision, I used the following criterion: If the separate parts of a question were *independent* of one another (i.e. if they could be done in any order and with any possible combination of correct and incorrect responses), then each part was entered into the database as a separate item. On the other hand, if any of the parts of a question required the result of an earlier part, then all parts of the question were coded as a single unit. This criterion is illustrated by the two multi-part problems in Fig. 4.6. In the multi-part problem of Fig. 4.6a, parts (a) and (b) are entirely independent of one another. In contrast, Fig. 4.6b shows a multi-part problem in which there is dependence: specifically, parts (b) and (c) cannot be answered without first solving part (a), and any errors made in a solution to part (a) would propagate through the solution to the rest of the problem. In one case, four questions from the Year 1 midterm exam (specifically #26-29), although numbered as separate items on the assessment itself, were nevertheless entered into the database as a single item because the latter questions could not be solved without the result of earlier ones. (Significantly, this same set of questions reappeared on subsequent years' midterm exams, reformatted as a single, multi-part question.)

A center pivot irrigation system uses a fixed water supply to water a circular region of a field. The radius of the watering system is 560 feet long. (a) If some workers walked around the circumference of the watered region, how far would they have to walk? Round to the nearest foot. (b) Find the area of the region watered. Round to the nearest square foot.	Triangle *ABC* has vertices at *A*(0,0), *B*(3,4), and *C*(5,0). (a) Find *AB, AC,* and *BC*. (b) What is the perimeter of triangle *ABC*? (c) What is the area of triangle *ABC*?
(a) A multi-part question coded as two independent items.	*(b) A multi-part question coded as a single item.*

Figure 4.6. Examples of multi-part questions in the assessment items corpus.

Based on this criterion above the corpus consists of 405 individual items, with 193 items in Year 1, 112 items in Year 2, and 96 items in Year 3. In the Findings section below, selected individual items will be presented with an accompanying analysis of how they might be solved, and how those solutions inform the question of what students were held accountable for. Such an analysis is in the spirit of Simon's (1975) analysis of the Tower of Hanoi problem, or Goldin's (1992) analysis of the way variations of a single question can hold students accountable for different knowledge (see the discussion of Table 1 earlier in this chapter). The size of the corpus, however, makes it impractical to analyze every item in this fashion; an analysis of the corpus requires the aggregation of items into classes, and the development of some systematic method of coding the classes.

Item Classes and Codes

Within the database, each item from Year 2 and Year 3 was identified as either new or a repetition (or adaptation) of a problem from a previous year. This led to the partitioning of the 405 items into 239 distinct *item classes*, each of which is an equivalence class of problems from different years. (To be clear, problems that repeated within a *single* year — for example, a question on a midterm exam that nearly replicates a question from a previous chapter exam — were not identified as part of the same item class.) The determination that two items belong to the same item class was made by identifying re-used words and phrases, the overt goal of the problem (i.e. what students were supposed to do or produce), as well as textual cues such as the placement of the item relative to the assessment as a whole. Item classes will function as the principal unit of analysis in Chapter 5, which documents adaptation in the corpus over time. In the present chapter, however, I examine the corpus at the level of the individual item, in order to quantify the extent to which the various mathematical dispositions can be found in the assessments. For this purpose, it is necessary to develop and deploy a set of codes for the assessment items.

Researchers have employed various classification schemes for coding tasks. For example, Caldwell & Goldin (1987) classified word problems according to two independent axes (abstract/concrete and hypothetical/factual), thus identifying each word problem as belonging to one of four categories. Stigler et al (1986), also working in the context of word problems, used the framework of Carpenter & Moser (1982, 1984) to classify textbook problems into 20 categories according to the natural methods by which they could be solved. More recently, Li (2000) compared the kinds of problems U.S. and

Chinese textbooks include in lessons on adding and subtracting signed integers. Li analyzes these tasks along three independent axes: mathematical features (does the solution require single or multiple computations?), contextual features (is the problem purely mathematical, or does it contain illustrative content?) and performance requirements (what type of response is called for, what is the cognitive requirement?). Mesa (2004), in a study of functions, draws on the construct of *conception* developed by Balacheff and Gaudin (2002) to code each task in her sample with a quadruplet (P, R, L, Σ), where P is a set of problems, R is a set of operators, L is a representation system, and Σ is a control structure. This results in an extremely subtle and sophisticated coding protocol that allows for a careful mapping of the concept space in a large corpus.

For this study, in order to describe the role these items play in creating an opportunity to learn the mathematical rationality, the items were coded with a set of 16 descriptors based on the mathematical dispositions identified in previous chapters. A list of these descriptors is found in Fig. 4.7. The first 12 descriptors correspond directly to dialectical pairs of dispositions that were described in Chapter 2. The four additional descriptors were employed to try to identify items that constitute an opportunity for students to experience an encounter with the unknown — to wonder about what is true, rather than to prove or disprove claims that are already identified as such. For example, items that were coded for *Existence* focus on the issue of whether a certain kind of object, or an object with a certain combination of properties, could exist. *Converse* was used to identify those items which foregrounded the question of whether the converse of a property was true. *FindHyp4Con* (an abbreviation for “find hypotheses for a conclusion") and *FindCon4Hyp* ("find conclusions for hypotheses") described items that foreground the conditional relationships among contingent possibilities.

Generalize	*Specialize*
Utility	*Abstraction*
Surprise	*Confirmation*
TheoryBldg	*ProblemSolving*
Simplicity	*Complexity*
Formalism	*Platonism*
Existence	
Converse	
FindHyp4Con	*Find Con4Hyp*

Fig. 4.7. Descriptors used to code assessment items.

In general, items were coded after specific descriptors if there was evidence in the text or diagram that the disposition was of some salience. Before describing the codes in more detail, it is helpful to anticipate how, exactly, such salience can manifest. Recalling the two motivating examples from the beginning of this chapter, I observe that there are at least two ways in which a disposition might be salient for a student working on a particular task: the disposition might be *explicitly marked* in the wording of a problem or in accompanying commentary, or it might be *implicit* in the work that students are expected to do to solve a problem. Consider the two problems shown in Fig. 4.8. The first of these (4.8a) calls explicitly for an investigation of whether a particular property, proved for one set of hypotheses, generalizes to a larger class of hypotheses: the word "generalize" is a marker of the salience of the disposition *Generalization* for this problem. This problem not only holds students accountable for being able to generalize; as a speech act (Searle, 1975), it also signals to students that generalization is something important, something one should care about. The second problem (4.8b) likewise calls for a generalization, but in a very different manner. No explicit marker signals the value of generality or calls students' attention to the fact that they are engaged in an act of generalization. And in fact it is possible that a student might actually attempt to solve the problem in 8b directly, by drawing a 22-sided polygon and attempting to draw and count all of its diagonals. However, as with the Possible Parallel Lines problem (Fig. 4.2), the unwieldiness of such a solution, and the likelihood of error, imposes a penalty on those students who avoid generalization, while the relative simplicity of a generalization-based solution strategy rewards those who have learned to deploy this key mathematical disposition.

In class we proved that the sum of the interior angles of any convex polygon is $180(n-2)$. Does this generalize to non-convex polygons? Give an argument for, or against, your answer.

8a

In this question, we will investigate how many diagonals a polygon has.

- Draw diagrams of polygons with 3, 4, 5, and 6 sides, and fill in the following table.

N (# of sides)	*d* (# of diags)
3	
4	
5	
6	
7	

- How many diagonals does a 22-sided polygon have?

8b

Figure 4.8. Two tasks that call for generalization.

In order to code the assessment items, I adapt notions taken from Doyle (1988) and Herbst (2006). Doyle (1988) defines an *academic task* in terms of four components: (1) a product to be generated by students, (2) operations students may use to produce the product, (3) resources supplied by the teacher, and (4) a weight that indicates the relative significance of the task in the accountability system of the class. To these four components, Herbst (2006) adds two important perspectives. First, Herbst attends to the temporal nature of student work — to the way in which a problem does not only specifies operations and resources but also, implicitly, the unfolding of those operations over time. Within this observation in mind, Herbst distinguishes between the "problem" (the specification of a goal and the provision of resources by the teacher) and the "task" (the universe of potential solution paths, modeled as the execution of a sequence of operations over time). Second, Herbst introduces the notion of *instructional situation*, a conceptual frame that participants "use to know who has to do what and when, so that whatever they do can be used to claim the fulfillment of their contractual obligations" (p. 316). In the Geometry class, for example, "Doing proofs" is one instructional situation, and "Installing theorems" is another (Herbst & Brach, 2006; Herbst & Miyakawa, 2008). A single problem may, in two different instructional situations, call for entirely different kinds of student work. Thus understanding the meaning potential of a task requires not only an examination of Doyle's four components, but also an anticipation of how the task might be deployed over time, and how that deployment is mediated by the instructional situation in which the task is located.

My discussion above suggests the significance of another (potential) component of a task, as exemplified by the commentary embedded within Figure 4.8(a). This commentary, explicitly marking one or more mathematical dispositions, provides a kind of framing distinct from the instructional situation Herbst considers: that is, it provides a conceptual frame for telling the student why a task is worth doing *relative to a set of values that are not necessarily instantiated by obligations in the didactical contract.* That is, a student attempting to solve the problem in Figure 4.8(a) will probably not be helped in that work by the knowledge that the problem can be framed as a "generalization"; nor does labeling the problem as such tell the student much about what he or she is accountable for doing (that information is found elsewhere in the statement of the task). But framing the problem as a "generalization" tells the student why, from a mathematical point of view, the problem is *worth doing*.

Thus every problem in the exam items corpus can be described in terms of the following five components:

- The goal to be produced (e.g., the result of a calculation to be performed, a proof to be written, a yes / no question to be answered)
- The resources provided to the student in and through the text. These include diagrams, hints, scaffolding (e.g. breaking a complex problem down into distinct parts to be answered in sequence), etc.
- Framing of the problem. This is an articulation of the importance of the problem, or why the problem is worth doing. This can be found in the form of commentary, as in the example of Figure 4.8a, but also through the use of words or phrases that are not strictly necessary for stating the goal, as in the example of Figure 4.1 at the beginning of this chapter.
- Stakes of the problem. How many "points" does a successful solution to the problem earn the student? Problems worth more points have higher stakes than a problem worth fewer points. The designation of some problems as "Extra Credit"

also marks them as lower stakes, in that not all students are expected or required to complete those.[27]

- The solution space of the problem. For each problem, I considered different sets of operations that a student might perform to reach a correct solution, as well as operations that might lead to solutions that are incorrect but plausible-seeming.

I examined the framing and the resources for each item for words or phrases that could mark one or more of the categories of appreciation; additionally I considered the solution space of the problem to see whether an application of one or more of the generative moves (categories of perception) would be particularly useful in solving it. (Other aspects of the item, such as other resources and the stakes, will be used in Chapter 5 in my study of item evolution over the three-year corpus.) Based on these considerations I applied descriptors both to tasks that mark a disposition explicitly, as well as those for which there is a preferred solution method that makes use of one or more of those dispositions.

Coding the 405 items in the corpus requires some clearly articulated criteria for when to employ each of the 16 descriptors. First, the descriptors are independent and non-mutually exclusive; any individual item can be coded after any combination of descriptors, or after none of them. A single item can also be coded with both members of

---

[27] Note that “Stakes” emcompasses but is somewhat more general than Doyle’s notion of “Weight”. A 10-point “mandatory” problem and a 10-point “extra credit” problem on the same test have the same weight — a student who does the former but not the latter earns the same grade as a student who does the latter but not the former — but nevertheless the stakes are different; students are “supposed to know” how to solve the mandatory problem, whereas they are not “supposed to know” how to solve the extra credit problem.

a single dialectical pair of dispositions. I describe next the criteria used for each of the descriptors.

As illustrated above, the descriptor *Generalize* was applied to any item that signals the mathematical value of generalizing results or questions. This signaling can be explicit in the wording of the problem (e.g. marked by the word "generalize" or phrases such as "in general"), or implicit in one of the likely methods of solution (e.g., a problem that can be solved using multiple methods, but for which the approaches that avoid generalization carry some penalty in the form of additional difficulty or likelihood of error). Note that neither the *proof* of a generalization, nor an explicit statement of the generalization in some general form, need be present (cf. Fig. 4.8b). The same criteria were used for the descriptor *Specialize.*

In the case of the second dialectical pair (*Utility / Abstraction*), any problem that placed mathematical knowledge in a purportedly "real-world" context (however contrived) was coded with *Utility*, because the use of such a context implies that the teacher sees value in the role that geometry can play in solving real-world problems, or perhaps that the teacher anticipates his students to see such value — in other words that Utility is an appropriate category with which to appraise the mathematical content of a question. On the other hand, the mere absence of such a real-world context does not, in itself, warrant the use of the keyword *Abstraction*: this latter was only used when the problem calls attention directly to something that is impossible (or seems impossible) to experience directly in the "real world" — for example, references to the idealized "zero thickness" of lines, or to processes or objects that go on "forever", or to geometry in more than three dimensions, or to non-Euclidean geometry, etc[28].

The dispositions of *Surprise* and *Confirmation*, as described in the previous chapters, are ways of framing results that either confirm or disconfirm previously-held beliefs. For this reason, these codes were only used when the text of a problem made explicit reference to conjectures and/or claims made at a moment that is prior to, and disjoint from, the moment at which the claim or conjecture is definitively settled. This could be done in a retrospective fashion ("In class I claimed that…. Now let's do a proof of it.") or in a prospective fashion ("Make a conjecture about…. In the next chapter, we will see if this conjecture is true.") Questions of this nature signal to students the value of revisiting previously-held but unconfirmed beliefs to determine unequivocally whether or not they are true.

The *Theory-Building* descriptor was used for items that call attention to the status of, and relationships among, the elements of a theory. For example, problems that call for students to use, or avoid, a particular theorem in order to prove another theorem signal that not only the postulates and theorems themselves, but also their independence, dependence, and equivalence, is something important to care about. The descriptor *ProblemSolving* was used for problems that call attention explicitly to a particular *method*[29]; for example, "What would you need to know in order to prove these triangles

---

[28] It is true that non-Euclidean and higher-dimensional geometries have enormous real-world utility, and in principle could be appraised as such; but for a student who does not know of those applications, it is likely that such contexts appear to be interesting (if at all) precisely because of their seeming removal from the real world.

[29] In Chapter 3, I claimed that problem-solving is a default (unmarked) disposition for the Geometry course, and in fact if every problem that called for the use of some

congruent by SAS?", or "Use the distance formula to show that the two segments are congruent." Likewise, the codes for *Simplicity* and *Complexity* were applied to items that explicitly labeled a piece of mathematics as simple (e.g. with an instruction like "Simplify your answer as much as possible") or as complex (as might be conveyed through a reference to "the complicated diagram below").

With respect to the pair of descriptors *Formalism* and *Platonism*, the former code was applied to any task that could be solved by ignoring all or part of the semantic content of the problem, and focusing instead on its syntactic structure. This might occur, for example, when a problem uses words or phrases for which there is no shared meaning — nonsense words, or terminology that students are not expected to understand — but which can be treated as "placeholders" and manipulated in a purely formal sense. It also can occur when the problem can be solved by relying solely on notational conventions and ignoring other sources of meaning (such as diagrams and/or numerical measurements). In contrast, the code for *Platonism* would be applied only in cases where some reference is made to a "real object" that exists independently of its representations. Such a reference must be explicit; otherwise any task that is accompanied by a diagram and a statement about the properties of the object represented would be coded for *Platonism*, and this would drown out the phenomenon we are attempting to document. As in the case of the other dispositions, the question here is whether the assessment items constitute an opportunity for the teacher to communicate to

---

problem-solving method had been identified as such, it is likely that the entire corpus would have been coded with this descriptor. For the purposes of the present chapter, I coded only those items that make *explicit* reference to a problem-solving method.

students a Platonist view of mathematics, or to hold students accountable for learning to view mathematics in such a way.

The criteria for applying the final four codes — *Existence*, *Converse,* *FindHyp4Con* and *FindCon4Hyp* — are relatively straightforward. *Existence* was used to mark those problems that ask a question of the form "Does ___ exist?" or "Can there be a ____?". *Converse* was used for any problem that foregrounded the difference between a conditional statement and its converse (such foregrounding does not necessarily require the use of the word "converse"). *FindCon4Hyp* and *FindHyp4Con* were used whenever a set of properties was specified, and students were asked to find either necessary or sufficient conditions for the given properties.

In order to determine reliability, a subset of 40 items (10% of the total corpus) was coded independently by a colleague using the same guidelines detailed above. For each of the dispositions, the number of disagreements was tallied, and an inter-rater reliability calculated as (number of agreement) / 40. Figure 4.9 summarizes the results.

Disposition	Inter-rater reliability
Generalize	85.0%
Specialize	77.5%
Utility	85.0%
Abstraction	100.0%
Surprise	100.0%
Confirmation	100.0%
TheoryBldg	92.5%
ProblemSolving	95.0%
Complexity	97.5%
Simplicity	95.0%
Formalism	90.0%
Platonism	100.0%
FindCon4Hyp	37.5%
FindHyp4Con	97.5%
Converse	100.0%
Existence	92.5%

Figure 4.9. Reliability of coding per disposition.

The extremely low reliability of *FindCon4Hyp* appears to have been the result of providing the colleague with an overly-broad description of that code; in 24 of those 25 disagreements regarding *FindCon4Hyp*, the colleague had marked the code and I did not. For that reason the results reported in the next section are, if anything, an underestimate of the presence of the dispositions in the data. The overall inter-rater reliability across all codes was found to be 90.5%.

Findings

Statistical profile of items

Because each item could be coded after more than one descriptor, or none at all, a first question to ask of the data is to what extent the codes were applicable at all. How many items received codes? How many received more than one code? Figure 4.10 summarizes the data by showing, for various values of *n*, the number of item classes that received *n* codes. It is perhaps worth noticing that slightly more than 61% of all items were not coded after any of the descriptors at all; of the 405 items in the data, only 155 were coded after one or more of the descriptors. On the other hand, a total of 265 codes were assigned throughout the corpus, a "density" of roughly 2 codes for every 3 item classes. If we restrict our attention only to those items that received codes, we find an average of 1.71 codes per item; and if we restrict our attention still further by discarding as outliers the four items with more than 5 codes each, we still find an average of 1.54 codes per item.

N	Items receiving n codes	% items receiving *n* codes (of *N*=405 total)
0	250	61.7%
≥1	155	38.3%
1	82	20.2%
2	61	15.1%
3	4	1.0%
4	3	0.7%
5	1	0.2%
7	2	0.5%
9	2	0.5%

Figure 4.10 Number of codes assigned to items.

Overall the impression given here is of a collection of assessment items that accords a substantial role to the dispositions coded for. In order to gain a sense of whether or not this role is *unusually* large, it is necessary to seek comparative data from an analogous corpus of assessment items. For this purpose, I applied the same coding protocols to the "Chapter Tests" printed at the end of each chapter in the textbook by Larson, Boswell & Stiff (2000). This book, which was the assigned text in all three years of the Honors Geometry course from which the assessments item corpus was drawn, represents a normative set of expectations for students in the course; the degree to which the assessment items corpus departs from that normative set of expectations is what needs to be measured. And indeed the difference is pronounced: of the 229 items contained in the 12 chapter tests in that text, only 38 were coded with one of the descriptors (about 16.5%, less than half the density in the assessment items corpus), and no items were coded with more than one descriptor. For additional corroboration, we can look back to the analysis done in Chapter 3, in which it was shown that geometry teachers gathered in study groups reported that they generally do not hold their students accountable for work that could exemplify many of the dispositions. Figure 4.10 suggests that, in this respect, the teaching implied by the assessment items differs substantially from what is normal in the teaching of Geometry, as evidenced by the textbook and as revealed in the records of those study groups.

Figure 4.11 breaks the data down by descriptor and shows the number of items identified with each. As can be seen, two of the codes most frequently assigned to items are *Formalism* and *Theory Building*, while the dialectical partners of these (*Problem-Solving* and *Platonism*) are among the least commonly assigned codes. For comparison, Figure 4.11 also shows the comparable data from the coding of the chapter tests in the textbook.

Descriptor	**# of items in corpus (N = 405)**	**# of items in textbook (N = 229)**
Generalize	19 (4.7%)	1 (0.3%)
Specialize	13 (3.2%)	0
Utility	19 (4.7%)	27 (6.7%)
Abstraction	5 (1.2%)	0
Surprise	5 (1.2%)	0
Confirmation	10 (2.5%)	0
Theory-Building	28 (6.9%)	1 (0.3%)
Problem-Solving	5 (1.2%)	0
Simplicity	2 (0.5%)	0
Complexity	4 (1.0%)	0
Formalism	33 (8.2%)	0
Platonism	0	0
Existence	33 (8.2%)	0
Converse	24 (5.9%)	4 (1.0%)
FindHyp4Con	13 (3.2%)	0
FindCon4Hyp	52 (12.8%)	5 (1.2%)

Figure 4.11. Number of items coded with each descriptor.

On the basis of these results a few observations can be made. First, it is clear that the assessment items in the corpus make use of different textual strategies than do the comparable items from the textbook chapter tests; that is, they presume a different set of competencies from the student, and imply a different set of intentions on the part of the teacher. In particular it is clear that the items in the corpus allocate a significant role for the mathematical dispositions. On the other hand it must be noted that not all of the dispositions are equally well-represented here: the relative rarity of items coded for Simplicity and Complexity is noteworthy. In the next several pages I turn to a more detailed discussion of some of the codes. For each of the codes discussed, I will provide examples of items from the corpus that were coded with the descriptor.

Generalize and Specialize

25 items were coded as exhibiting the dispositions of generalization, specialization, or both. Figure 4.12 shows two items from such classes. The item in Figure 4.12a calls for students to prove a property that is true of all triangles — a relationship among the orthocenter, circumcenter and centroid — in a special case for which the proof is particularly easy. In this sense it resembles the first motivating example from the beginning of this chapter, in which a general claim that had been previously made in class is to be proven in a special case.

Figure 4.12b contains two items that were coded independently in the database because they could have been solved in either order. The student, however, would have encountered these two items on an exam in the form of a single, two-part problem, and the default for such a student would have been to solve them in sequence. Thus the implied student would have found the value of each angle in an equiangular dodecahedron, and then generalized the result to find a formula expressed in terms of the parameter n.

[Year 3, Unit 5 exam, #1]	[Year 3, Ch. 3 exam, #9a]

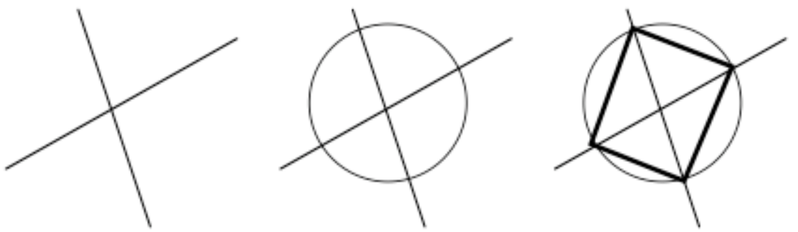
Draw a right triangle. Where, exactly, is the orthocenter? Where, exactly, is the circumcenter? Prove that, in the special case of a right triangle, the orthocenter, circumcenter, and centroid are collinear, with the distance from the orthocenter to the centroid exactly twice the distance from the centroid to the circumcenter.	A *dodecagon* is a 12-sided figure. In an equiangular dodecagon, what is the measure of *each* angle? [Year 3, Ch. Exam, #9b] Find a formula for the measure of *each angle* in an equiangular *n*-sided polygon.
(a) An item from Year 1, coded for *Specialize.*	(b) Two items from Year 3; the second was coded for *Generalize.*

Figure 4.12. Problems coded for (a) *Specialize* and (b) *Generalize.*

The first of these two tasks holds the student accountable for knowing how to find the sum of the angles in any polygon: such an implied student would find the indicated sum, and divide the result by 12. The fact that the second item reads "Find a formula..." implies that students are not expected his students to *recall* such a formula but to *derive* one. This second item was coded with the descriptor *Generalize* because in order to solve it, a student would repeat the calculations performed on the earlier item, operating on *n* in the same way that he had previously operated on 12. Notice that while the question in Figure 4.12a refers *explicitly* to specialization, the second item in Figure 4.12b only *implicitly* calls for a generalization.

One might ask whether the items in Figure 4.12 were "typical" of the items in the corpus. In this connection it is worth noting that both of the items in Figure 4.12 appeared for the first time in Year 3. Thus all three of the items might be considered "singletons", i.e. items that belong to classes with no other members. On the other hand, the item in Figure 4.13 (below) appeared on an exam in Year 1 and re-appeared, with no modifications, in Year 2. In this item, students are expected to recognize that a method for constructing a parallelogram actually results in a rectangle; the presence of the phrase "special parallelogram" signals that a special case is under consideration. Note that this problem does not call for students to *perform* a specialization, but merely to *recognize* one.

Consider the following method for constructing a parallelogram:

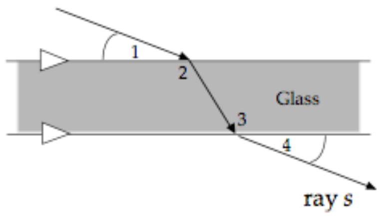


What kind of "special parallelogram" results? Prove your answer.

Figure 4.13. A problem coded for *Specialize.*

Utility and Abstraction

The corpus contained 24 items that were coded after one or both of the pair of dispositions *Utility* and *Abstraction*. As was briefly stated above, items were coded for *Utility* if they contained any kind of real-world context, no matter how contrived or artificial, on the grounds that the presence of such a context at all constituted an implicit endorsement of the value of utilitarian mathematics. (Recall that the word "utility" is here restricted to refer to the usefulness of mathematics in *non-mathematical* contexts.) Figure 4.14 shows three items from classes that were coded for *Utility*. Note the variation among them: the first item takes its context from physics, the second from architecture, and the third from a fanciful blend of cartography and super-hero criminology[30].

<table>
<tr>
<td>When light enters glass, the light bends. When it leaves glass, it bends again. If both sides of a pane of glass are parallel, light will leave the pane at the same angle at which it entered. (See diagram) Prove that the path of the exiting light is parallel to the path of the entering light.

</td>
<td>The outer wall of Fort Jefferson, which was originally constructed in the mid-1800s, is in the shape of a hexagon with an area of about 466,170 square feet. The length of one side is about 477 feet. The inner courtyard is a similar hexagon with an area of about 446,400 square feet.

(a) What is the ratio of the areas of the small hexagon to the large hexagon?

(b) What is the scale factor of the small hexagon to the large hexagon?

(c) How long is a side of the inner courtyard?</td>
<td>When I was a kid I used to watch "The Superfriends" – kind of a dumbed-down version of the Justice League. Here's a typical Superfriends scenario: While the heroes are meeting in the Hall of Justice, suddenly the Troubalert computer sounds an alarm: the Legion of Doom has just robbed three banks in Metropolis!

Batman assumes that the Legion of Doom's secret hideout is probably equidistant from the three crime scenes. Explain how, using only a bat-compass, bat-straightedge, and a map, they can use the techniques of geometry to pinpoint the Legion of Doom.</td>
</tr>
<tr>
<td>A</td>
<td>b</td>
<td>c</td>
</tr>
<tr>
<td colspan="3">Figure 4.14. Three problems coded for Utility.</td>
</tr>
</table>

In addition to the variation in contexts, each of these three problems holds students accountable for learning different kinds of knowledge. The item in Figure 4.14a holds students accountable for knowing how to use properties of one pair of parallel lines to prove that another pair of lines is also parallel. The item in Figure 4.14b holds students accountable for understanding the notion of scale factor, and for knowing that

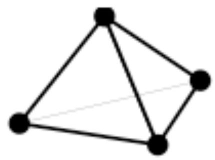
[30] It could very well be argued that the context of the item in Fig.14b is no less contrived than that of the item in Fig. 14c. Under what conceivable circumstances would one know the area of two hexagonal regions without knowing the lengths of the sides?

the ratio of the areas of two similar polygons is the square of their scale factor. The item in Figure 4.14c holds students accountable for knowing that the circumcenter of a triangle is equidistant from its three vertices, and that it can be constructed as the intersection of the bisectors of any two of the angles in the triangle. The three items share very little, apart from a tacit endorsement of the principle that mathematics can be useful in the "real world".

The fact that only 19 of the 405 items were coded for *Utility* makes it clear that the default, in this corpus, is for tasks to be presented without a real-world context. But the absence of such a context does not, in and of itself, mean that an item should be coded for *Abstraction*. The goal of the coding was not to identify problems that are abstract, but rather to identify problems that *call attention to abstraction as a value*. There were very few (only 5) item classes that seemed to do this. One of them is reproduced in Fig. 4.15. It will be immediately noticed that that item is *long*: it filled an entire page in its original context, and contains more than 300 words of text, nearly twice as many as the three items in Fig. 4.14 combined. Moreover the text contains several markers of the value of "abstraction": the phrase "in higher-level mathematics", the reference to "higher dimensions", and the explicit statement that four-dimensional simplices cannot be drawn all point to the notion that the notion of simplex is an abstract one, disconnected from the world of direct experience.

Extra Credit. (5 points) In higher-level mathematics, a triangle is sometimes called a *2-simplex* (plural *simplices*). As the name suggests, there are also 1-simplices, 3-simplices, and so on. Simplices can be glued together to form complicated surfaces (called *complexes*). The study of simplices is called *simplicial geometry*.

Here are the basic definitions:



A 3-simplex (3-dimensional tetrahedron.)

- If you have *one point*, you call it a 0-simplex.
- If you have *two points*, then (as long as they are not 'coincident', i.e. not the same point) you can form a line segment; this segment is called a 1-simplex.
- If you have *three points*, then (as long as they are not collinear) you can form a triangle; this is called a 2-simplex.
- If you have *four points*, then (as long as they are not coplanar) you can form a tetrahedron; this is called a 3-simplex.
- In higher dimensions, you can form 4-simplices, 5-simplices, etc.

Notice that every simplex contains within it "smaller" simplices. For example, the *sides* of a 2-simplex are 1-simplices, and the *vertices* of a 2-simplex are 0-simplices.

Similarly a 3-simplex has *faces* that are 2-simplices, *edges* that are 1-simplices, and *vertices* that are 0-simplices.

(a) Fill in the chart below. *Note*: You should be able to fill in the first three rows just by drawing some pictures and counting. The last row involves 4-dimensional geometry, so you can't draw a picture; you to use inductive reasoning to make a prediction based on a pattern in the first three rows.

	How many 0-simplices does it contain?	How many 1-simplices does it contain?	How many 2-simplices does it contain?	How many 3-simplices does it contain?	How many 4-simplices does it contain?
1-simplex					
2-simplex					
3-simplex					
4-simplex					

Fig. 4.15. A problem coded for *Abstraction*.

One might look at the problem in Fig. 4.15 as an impressive introduction of sophisticated mathematical content into the secondary Geometry course — it goes well beyond the normal curriculum and beyond the 3rd dimension. On the other hand, an analysis of the task from a student's perspective leads to a somewhat more reserved assessment. In order to complete the first three rows of the table, the implied student need only:

- be able to understand the “definitions” (really, examples) of simplices
- be able to draw a picture of a segment, a triangle, and a tetrahedron
- be able to count to 6

With these minimal skills alone, a student could fill in the first three rows, as shown below in Fig. 4.16. In order to complete the table, a student must then (as the problem says) use inductive reasoning to make a prediction based on the pattern in the first three rows. There are a number of things a student might do at this point. The pattern in the first column seems evident, and a student might reasonably (and correctly) be expected to fill in a “5” for the first cell in the last row of the table (an example of the “Generalize” disposition). A slightly more subtle pattern runs down the diagonals of the pattern: a student who sees them might be expected to guess that a “5” goes in the second-to-last cell in the bottom row, followed by a “1”.

	How many 0-simplices does it contain?	How many 1-simplices Does it contain?	How many 2-simplices Does it contain?	How many 3-simplices Does it contain?	How many 4-simplices Does it contain?
1-simplex	2	1			
2-simplex	3	3	1		
3-simplex	4	6	4	1	
4-simplex					

Fig. 4.16. The first three rows of the table from Figure 4.14.

But after this, it is unclear what else a student should do. If the student is familiar with Pascal’s Triangle, he or she might be expected to recognize a variation of it here; absent such a familiarity, there really is very little a student could be expected to do to fill in the remaining two cells correctly. It is even more difficult to imagine how a student might answer part (b) of the item, which calls for a general conjecture about the number of *k*-simplices contained in an *n*-simplex. To answer this a student would need to know something about binomial coefficients, or (at the very least) have a familiarity with Pascal's Triangle. And yet nowhere else on this, or any other exam in the corpus, is there any indication that the teacher has taught that content and expects students to know it. It will probably not have escaped the reader's notice that the whole task is explicitly designated an *extra credit problem*, and thus (almost by definition) cannot be taken as an indicator of what the teacher held students accountable for. Thus while the first few parts of the item are more or less trivial (requiring little more on the part of the student than visual perception, counting, and pattern recognition), the latter parts of the item appear to be aimed at the exceptional students who may have learned some mathematics in an outside-of-class experience, for example on a Mathcounts® or other "Mathlete"-type activity[31]. Such students might or might not recall a previously-encountered formula or table, but in the event that they do, they are not expected here to do anything with it other than recognize and reproduce it. In summary, despite the fact that this problem appears on its surface to introduce students to high-level content, on closer examination it is seen to be a marginal example that makes only superficial demands on students' knowledge. It embodies the value of *Abstraction* — but little else.

Surprise and Confirmation

Only 5 item classes in the corpus were coded after *Confirmation*, one of which was also coded for *Surprise*. It will be remembered that Confirmation, in Chapter 2, was used to refer to a positive appraisal of a mathematical result on the grounds that it

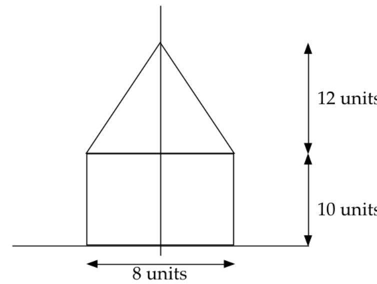
---

[31] Mathcounts® is a nation-wide competitive mathematics enrichment program for middle-school students (http://mathcounts.org).

confirmed something that had long been suspected true, but not yet proven; Surprise was used to value results that did the opposite, by disconfirming something that was expected to be true. Both of these dispositions require, as a prerequisite, that open questions be allowed to exist for a significant period of time: if every question is resolved either affirmatively or negatively almost immediately when it is raised, there can be neither surprises nor confirmation. In Chapter 3 it was shown that Geometry teachers in the ThEMaT Study Group were almost uniformly opposed to leaving questions unresolved for extended periods of time, suggesting that Surprise and Confirmation have little place in their customary practice; the scarcity of examples in the corpus of assessment items suggests that the teacher here is not dissimilar in that respect.

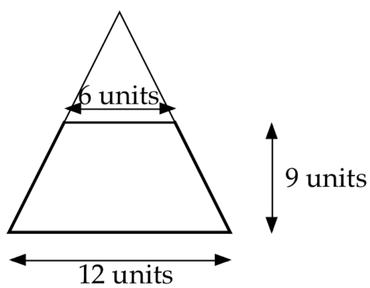
One of the 5 item classes coded for *Confirmation* was already discussed above: it is the question about the area of the dual of an isosceles trapezoid (see Fig. 4.1 at the beginning of this chapter). Note that that item specifically references the fact that the property had been claimed as true but not yet proven, and suggests that to produce a proof (at least of the special case) would amount to paying off a kind of debt. Another item coded for *Confirmation*, also a “singleton” from Year 1, is in Fig. 4.17. In its length and its tone, it is somewhat reminiscent of the simplices problem (Fig. 4.15).

15. One way to find the centroid of an irregular region is to break it into smaller regions whose centroids are known and then take a *weighted average* of the centroids. For example, in the figure at right, the centroid of the rectangle is at (0, 5) (halfway up the rectangle), and the centroid of the triangle is (0, 9) (one-third of the way up the triangle). To find the centroid of the combination, one might think you just take the average of 5 and 9 — but this is *wrong*, because it neglects the fact that the rectangle is bigger than the triangle. Instead, we multiply each coordinate by the area of the figure (the "weight") and then divide by the total area. So, in this case, the



correct formula is $\dfrac{(\text{area of triangle})(\text{centroid of triangle})+(\text{area of rect.})(\text{centroid of rect.})}{\text{total area}}$. In this case, $\dfrac{(48)(9)+(80)(5)}{128}=6.5$, so the centroid is at (0, 6.5).

Using this method, you can find the centroid of an isosceles trapezoid. Refer to the figure below at left:



(a) What is the height of the small triangle? (3 pts)

(b) What is the y-coordinate of the centroid of the small triangle? (3 pts)

(c) What is the y-coordinate of the centroid of the large triangle? (3 pts)

(d) What are the areas of the small and large triangles? (4 pts)

(e) What is the area of the trapezoid? (2 pts.)

(f) Use the formula

$$\frac{(\text{area of small}\,\Delta)(\text{centroid of small}\,\Delta)+(\text{area of trap.})(\text{centroid of trap.})}{\text{area of large}\,\Delta}=\text{centroid of large}\,\Delta$$

to find the centroid of the trapezoid. (5 pts)

(g) Check that your answer matches with the formula I gave in class, $y=\dfrac{h}{3}\left(\dfrac{a+2b}{a+b}\right)$, where a and b are the long and short bases (respectively), and h is the trapezoid height. (5 pts)

Figure 4.17. An item coded for *Confirmation.*

I will return to this example later, in my discussion of *Problem Solving*. For the present, the item in Fig. 4.17 is significant principally for part (g), which asks students to confirm (in a special case) a formula that was presented as true, but not proved, in class — much like the problem in Fig. 4.1. Rather than asking for a proof that something is generally true, this item merely asks for confirmation that a particular value, calculated via one method, is *consistent* with a general formula. Despite this difference, both of these problems imply teaching in which it is important that claims not be taken solely on faith — and yet also in which confirmation of such claims may be delayed until a later date.

Very few items in the corpus (only 5) were coded for *Surprise*, and in fact those five items really consist of three item classes: one item that appeared in Year 1 and again (without modifications) in Year 2, one item that appeared in Year 3 only, and one item that was the take-home problem from the final exam of both Years 1 and 2. This take-home problem is a true outlier in the corpus, both in terms of the number of dispositions embedded within it (it was coded with 9 of the 16 available descriptors) and in terms of the way in which it violates many of the conventions of assessment. A fuller account of this exercise is given in the next chapter.

Theory Building and Problem Solving

As Fig. 4.11 shows, 33 items were coded with the descriptor *Theory Building*, making it one of the most commonly-applied codes. In general an item received this descriptor if it made reference to the relationship between two or more theorems, postulates, or definitions. Three examples of such items can be found in Fig. 4.18. The first example (Fig. 4.18a) asks for students to show that any one of the properties of parallel lines may be used to prove all of the other properties — in other words to show that the designation of one of those properties as a "postulate" and the others as "theorems" is an arbitrary matter of convention. Note that this is very close to the content of the ThEMaT animated classroom scenario "Postulates and Parallel Lines", summarized in the previous chapter. It will be recalled from Chapter 3 that the members of the ThEMaT study group were nearly unanimous in their agreement that they do not typically hold students accountable for proving the equivalence of the properties of parallel lines; nor do they make questions of the organization of those properties into a theory part of the classroom discourse. In the context of that discussion it is particularly interesting that the author of the exams corpus holds students accountable for *precisely* the same content that the study group teachers rejected.

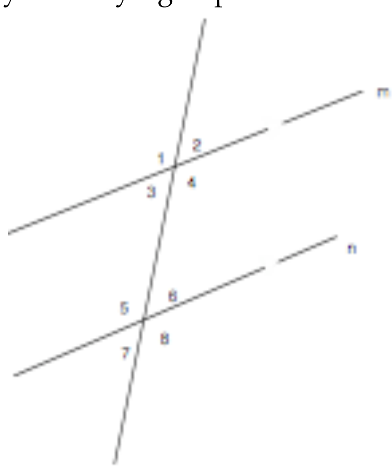
Using any one of Postulate 15 and Thm. 3.4-3.6, prove any other one. (e.g., you could use Corresponding Angles to prove Alternate Interior, or vice versa.) You may wish to refer to the diagram below. Make sure you state what you're trying to prove!  	The proof of the Polygon Interior Angles Theorem is based on the … (a) Alternate Interior Angles Theorem (b) Triangle Sum Theorem (c) Midpoint Theorem (d) Definition of a regular polygon	Consider the following conjecture (this is a variation on the Angle Bisector Theorem): "Let AOB by any angle, with points *A* and *B* equidistant from the vertex. If *P* is any point on the angle bisector of AOB, then *P* is equidistant from *A* and *B*." (a) Draw a sketch illustrating this conjecture. (b) How is this different from the "usual" angle bisector theorem? Explain. (c) Is the conjecture true? Convince me with either a proof or a counterexample.
A	b	c

Figure 4.18. Three problems coded with the descriptor *Theory Building.*

The second example of *TheoryBldg* (Fig. 4.18b) similarly focuses squarely on the issue of the relationship between two theorems: the student is expected to know which theorem (or property) was instrumental in proving a second theorem. Note, though, that in this exercise the student is expected merely to *identify* the relevant theorem from a list of candidates; the student is not expected to produce a proof, only to remember the idea of it. The third example (Fig. 4.18c) is significant for the present discussion primarily because of the second part of the question, which asks for the student to articulate the distinction between two nearly-identical statements, one of which is an established theorem while the other is only a conjecture.

In contrast to the large number of items coded after *Theory Building*, it is noteworthy that only five items were coded for *Problem Solving*. As was mentioned above, items only received this code if they made *explicit* reference to a particular method of solution. One such item is the "centroid" problem discussed above (Fig. 4.18). Note that this problem takes nearly half of a page to describe a method for locating the centroid of a region by breaking it into simpler parts and calculating a weighted average. From the emphasis the text lays on the generality of the method, it does appear that the teacher is hoping that the students will do more than just solve the problem correctly: he intends for his students to learn some general principles about a particular method of problem-solving.

Another item in the corpus that was coded after *Problem Solving* was a problem that called for a proof that the medians in a triangle are concurrent. This item (Figure 4.19) was coded for *Problem Solving* because the instructions in part (a) stress the need for the student to make intelligent strategic choices about such matters as where to put the axes and what scale to use.

We have proven already that the three medians of a triangle are concurrent. In this problem, you will re-prove the result another way, using a coordinate proof.

a. Draw a triangle in a coordinate system. Make choices about where to put the origins and axes, and choose a scale. Label the three vertices of your triangle with coordinates.
b. Find the coordinates of each of the three midpoints of the triangle.
c. Find the equation of the median from *A* to the midpoint of BC.
d. Find the equation of the median from *B* to the midpoint of AC.
e. Find the equation of the median from *C* to the midpoint of AB.
f. Choose any two of the equations you found in (c)-(e) and solve them.
g. Now take your solution from (f) and plug it into the unused equation from (c)-(e) to verify that it is a solution of all three equations.

Figure 4.19. An item coded after *Problem Solving.*

Formalism and Platonism

The descriptor *Formalism* was assigned to 33 items in the corpus, making it one of the most common codes. Items were assigned this code if they appeared to give a privileged role to reasoning that is based on purely formal properties. For example, the items in Figure 4.20a and 20b both call on students to draw conclusions based solely on notational conventions (neither item was accompanied by a diagram). In 20b, for example, a student would be expected to know that the angles and segments in congruent triangles correspond to one another according to their position within the notation for the triangles. Thus, even though ΔABC and ΔCAB designate the same triangle, the assertion that $\Delta ABC \cong \Delta CAB$ implies that all three angles are congruent to one another, as are all three sides — in other words that the triangle is equilateral. Because this conclusion follows purely from the *formal properties of the notation*, the item expresses the value of Formalism.

The third example (Fig. 4.20c) differs somewhat. The significant point about this item is that a student could not realistically be expected to know the meaning of the theorems referred to — and in fact a footnote printed on the examination explicitly stated that students need not worry about understanding the *content* of the Riemann Conjecture of the Prime Number Theorem. In order to answer the question accurately, a student would need to deliberately ignore the *semantics* of the problem and focus instead on its *syntactic* properties. The conclusion follows by pure logic from its premises, even if one is entirely ignorant of its meaning; and in fact the student must *suppress* any attention to the meaning of the premises in order to solve the problem. Put another way, "the Riemann Conjecture", "Prime Number Theorem" and "grows logarithmically..." are mere placeholders; recalling Hilbert's famous joke about the undefined terms in Geometry, they may as well be "tables", "chairs", and "beer mugs".

What is the intersection of $\overrightarrow{AB}$ and $\overleftrightarrow{BA}$? What is their union?	As you know, when writing a triangle congruence statement, the order of the letters matters. Suppose for three points *A, B,* and *C* is happens to be true that $\Delta ABC \cong \Delta CAB$. What can you conclude about ΔABC in this case? Be as detailed as possible.	The mathematician Riemann proved the following theorem (called the *Prime Number Theorem*) in the 19th century: "If the Riemann Conjecture is true, then the number of prime numbers below *N* grows logarithmically with *N*." Suppose a student, working diligently and in secret for months, manages to prove the Riemann Conjecture to be true. (a) Putting this amazing discovery together with the Prime Number Theorem, what can our student conclude? (b) Which logical law (Syllogism or Detachment) did you use in answering (a)?
a	b	c

Figure 4.20. Three items coded for *Formalism.*

As was noted earlier, no items in the corpus were coded for *Platonism.* This may be because the bar was set too high in the coding process: an item would have been coded for *Platonism* only if it contained some indication that the "thing" under discussion possessed some kind of reality that transcended its formal properties. No such items were found; however, this could be taken as an indication that Platonism is in fact the *default* mode for items in the corpus. Most items in the corpus include a diagram, and refer to the diagram as if it were a picture of a real thing; the "reality" of the figure is taken for granted, not made explicit.

Existence and Converse

In Chapter 1, I briefly mentioned the prominent role given to questions of existence in the corpus of assessment items, and illustrated that role with a list of examples. Figure 4.11 bears out the claim that such questions play a significant role in the corpus; 33 items were coded for *Existence*, making it one of the most frequently-assigned codes over the corpus. Some of those items are shown in Figure 4.21.

It will be recalled that the reason for including a code for *Existence* — which, after all, was not a member of one of the six dialectical pairs described in the previous two chapters as comprising part of the mathematical sensibility — was to attempt to identify those items that place the student in a condition of genuine uncertainty as to what is true (the "encounter with the unknown"). In this connection it is worth noting that, while all of the items in Figure 4.21 ask some variation of the question "Does ____ exist?", four of them, (a, c, f, and g) should be answered "No". And yet there is little in the wording of the items that gives away the fact that the object indicated is impossible;

note in particular that items (a) and (b) are phrased using nearly identical wording, as are (c) and (d), despite the fact in each of those pairs one of the items describes objects that do exist and one describes items that do not. What emerges from these textual strategies is an *intentio operis* of posing questions of existence in a manner sufficiently neutral so as not to render the answers to those questions obvious, and an implied student who is expected to be capable of navigating in a conceptual domain in which the objects under consideration may or may not be possible.

(a) If an angle is bisected into two pieces, can either of the pieces ever be obtuse? Justify your answer.
(b) Is it possible, using only a straightedge and a compass, to construct a 45° angle? Describe how in words, and demonstrate below.
(c) If two lines intersect, can they be non-coplanar? If you answered 'yes', what are lines like this called? (If you answer 'No' to the first part, then no further explanation is necessary.)
(d) If two lines are co-planar, can they be non-intersecting? If so, what are lines like this called? (If you answer 'No', then no further explanation is necessary.)
(e) Can a triangle be both isosceles and obtuse? If so, draw an example and indicate the measure of the three angles. If not, explain why.
(f) Is it possible for a regular polygon to have every angle equal to 155°? If so, how many sides must it have? If not, why not?
(g) We proved that, in any triangle, there is a point that is equidistant from all three sides of the triangle, and that you can construct such a point by bisecting all three angles and seeing where they cross. But we never addressed the "uniqueness problem": Is it possible that there might be a *different* point that is *not* on the intersection of the three angle bisectors, but nevertheless is still equidistant from all three sides of the triangle? Decide whether you think this is possible or not. If you think it is, give me an example. If you think it is not, give me a proof.

Figure 4.21. Seven items coded for *Existence*.

Another code used to locate items that place students in an encounter with the unknown was *Converse*. This code was applied to items that call attention to the fact that the truth of a statement and its converse are logically independent of one another (so that a resolution of the truth status of one of those propositions does not necessarily help resolve the truth status of the other). 24 item classes were coded with *Converse*, making this one of the more common themes in the items corpus. Three of them are reproduced in Figure 4.22. The first of those examples makes use of a parenthetical "warning" to call the student's attention explicitly to a consideration of the meaning of "if an only if"[32]. The second example offers the student a choice of proving either a theorem or its converse — implying, among other things, that a student is accountable for knowing the difference between the two. The third example in that table is more subtle: the proposed implication is true in one logical direction (i.e., a rhombus does have perpendicular diagonals) but not in the other (a quadrilateral with perpendicular diagonals need not be a rhombus).

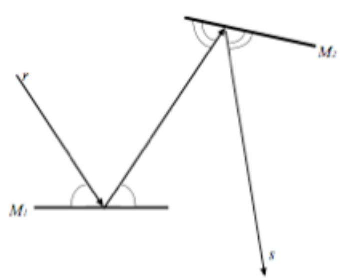
When a ray of light reflects off of a mirror, the angle of incidence is always equal to the angle of reflection. In the figure at right, an incoming light ray *r* reflects off of two mirrors M_1 and M_2. Show that the outgoing light ray *s* is parallel to *r* **if and only if** the two mirrors are parallel to each other. [Before you start working on this problem, make sure you understand what "if and only if" means.] 	State and prove *either* the Angle Bisector Theorem *or* its converse.	True or False: "A quadrilateral is a rhombus if and only if its diagonals are perpendicular." If true, give a proof; if false, give a counterexample, and correct the statement.
(a)	(b)	(c)

Fig. 4.22. Three items coded for *Converse*

[32] The reader may have noticed that the item in Table 22a is reminiscent of the "refracting light" problem reproduced earlier as Table 14a. In fact, these two problems are members of a single item class. I will discuss the evolution of this item class in the next chapter.

Thus in order to answer the question correctly a student must understand that the two directions are independent of one another, be able to identify which implication is true, and replace a (false) biconditional statement with a true conditional one.

The most frequently-assigned code, as Figure 4.11 shows, was *FindCon4Hyp.* This code was used to designate item classes in which a student was expected to find a conclusion that could be deduced from specified hypotheses. As with *Existence* and *Converse*, the purpose of this code was to identify those problems that cast students into a context in which the truth or falsity of the statements at play is not evident. Many of those problems have already been described above. For example, the items in Figure 4.21b (introduced to exemplify the *Formalism* code), and Figure 4.14 (introduced to exemplify *Specialize*) were also coded with *FindCon4Hyp*. Figure 4.23 contains some additional examples. The first of these, 4.23a, is similar in spirit to the problem in Fig. 4.13. The multi-part item in 4.23b covers fairly elementary material (it was taken from a Chapter 1 exam) and yet we can already see emphasis being placed on the interrelationship of contingent possibilities (note also that in the second of those parts, there is *no* conclusion that can be drawn from the given information). The multi-part item in Fig. 4.23c calls for the student to make a conclusion from given hypotheses, and then to reconsider whether the conclusion would still hold under a weakened set of hypotheses — a nearly pure illustration of the generative moves I described in Chapter 2.

$\overline{AC}$ and $\overline{BD}$ intersect each other at *N*. $\overline{AN} \cong \overline{BN}$ and $\overline{CN} \cong \overline{DN}$, but $\overline{AC}$ and $\overline{BD}$ do not bisect each other. Draw $\overline{AC}$ and $\overline{BD}$ and *ABCD*. What special type of quadrilateral is *ABCD*? Prove your answer.	In questions (a)-(d) below, answer the question if possible; if there is not enough information in the question, write "not enough information" (a) Suppose *S* is a set of points that are non-collinear. What is the minimum number of points in *S*? (b) Suppose *T* is a set of points that are collinear. What is the maximum number of points in *T*? (c) Suppose *R* is a set of points that are non-coplanar. What is the minimum number of points in *R*? (d) Suppose *P* is a set of collinear points, *L* is a second set of collinear points, and suppose further that *P* and *L* have one point in common. Now suppose *R* is a third set containing all of the points from both *P* and *L*. Can you say anything about *R* (e.g., is it collinear, coplanar, neither)?	Consider the following theorem: For coplanar lines *m, n,* and *r*, if $m \perp r$ and $n \perp r$, then …………… (a) What should be the conclusion of this theorem? (b) Prove the theorem. You might find it helpful to draw a diagram. (c) Do you believe the theorem would be still true if we deleted the word "coplanar"? Would the proof you wrote in (b) still work? (Note, these are two different questions.) Why or why not?
(a)	(b)	(c)

Figure 4.23. Examples of items coded with *FindCon4Hyp.*

The large number of items coded after *Existence, Converse,* and *FindCon4Hyp* indicates a set of textual strategies that testify to an intention to force students to confront situations of doubt and uncertainty. The "Possible Parallel Lines" problem, introduced as the second motivating example at the beginning of this chapter, further exemplified this: the two distinct strategies described earlier for solving the problem (the "direct" and "indirect" approaches) correspond to different ways of approaching such encounters with the unknown.

Discussion

Throughout this chapter I have argued that a corpus of examination items can, in some cases, be regarded as a single "text", and as such is amenable to description in terms of its "implied" or "model" author and "mock" or "model" reader — or, as I have termed them here, the *implied teacher* and *implied student.* It should be recalled that these implied personages are hypothetical constructs, entities whose characteristics we infer from what Eco called "textual strategies". The value of these constructs is that it provides us with a language for speaking of the intention of instructional texts (that is, the *intentio operis*) that is disconnected from claims about the intention of the actual teacher (*intentio auctoris*). Whether or not the implied teacher and implied student bear any close similarity to their "empirical" counterparts is a question that could in principle be explored through the use of interviews. However, for the purpose of the present analysis I have adopted the stance of the "new criticism" literary school (Ransom, 1941; Wimsatt & Beardsley, 1946/1999), in which the intentions of the *actual* author are regarded as irrelevant and, in any event, generally unknowable.

This stance rests, however, on an important assumption: namely, that a corpus of short examination items exhibits sufficient coherence (thematic and otherwise) to warrant regarding it as a single text. There is, to be sure, some coherence that derives purely from the institution — all items are used for a course called "Honors Geometry" in a single school. But this is akin to regarding an anthology of poems as having coherence simply on the grounds that they are published between a single set of covers; one would not necessarily expect to find a consistent set of textual strategies (a single authorial "voice") under such circumstances. To warrant such an expectation, one would need to document the use of cohesion as a textual strategy to create coherence within the ideational, interpersonal and textual metafunctions as well.

As a first step towards establishing this cohesion, I have introduced the notion of *item classes* (sets of items that repeat in subsequent years of the corpus). The existence of these item classes entitles me to speak of “a course taught over three years”, rather than regard it as three different courses. In the next chapter I document the evolution of items within those classes, and illustrate how that evolution can be used to gain insight into the teacher’s need to adapt his means of assessment to the ecological constraints of his classroom. I also document in detail the existence of *item chains* that give coherence to a one-year “slice” of the corpus (as might be encountered by a student).

Beyond the mere claim that the corpus can be regarded as a text, I have shown that the text can be usefully described by coding item classes after the mathematical dispositions that I identified and discussed in preceding chapters, and that certain of those dispositions are more represented within the corpus than are others. In particular, it is worth noting the prominent role that *Theory Building*, *Formalism*, *Existence*, and *Converse* seem to play in this corpus, especially when we consider the lack of commitment to these dispositions that was found earlier in my analysis of the study group records (Chapter 3), and the near-total absence of the dispositions in the chapter tests from the textbook. In fact, this contrast is part of what gives the corpus coherence, and helps to generate a picture of teaching practice which, idiosyncratically, attempts to teach students what it is to “be like” a mathematician.

Chapter 5

Coherence and adaptation in the assessment items corpus

Introduction

In the preceding chapter I examined a corpus of examination items written and used by a geometry teacher over a 3-year period, and showed that those items allocate a significant role to the mathematical dispositions — the elements of the mathematical sensibility that was documented in Chapter 2. In this chapter, I examine that same corpus of items from a different perspective and for different aims: rather than look at individual items (or even item classes), I look at the way the items in the corpus are structured over *time*. In particular, I look at the corpus over multiple distinct time-scales, with the goal of documenting distinct phenomena. These scales, which were introduced in Chapter 4, and are represented schematically in Fig. 5.1, will now be elaborated on:

- The time-scale of the *single year* allows us to read the corpus as it might be encountered by a hypothetical student, with the goal of finding evidence of *cohesion* within that corpus. Finding such cohesion provides an empirical warrant for the theoretical position staked out in the last chapter: namely, that the corpus can legitimately be regarded as a *single text* (rather than a large anthology of independent, smaller texts), which in turn entitles us to inquire after the text's "model author" and "model reader" (which I have referred to previously as the "implied teacher" and "implied student"). This cohesion is represented in Fig. 5.1 by the thin horizontal lines that link items together within a single year.

- In contrast, the time-scale of the *three-year corpus* allows us to track the evolution of items from year to year. We will see that as certain items drop out of the corpus, others enter it to take their place; simultaneously other items remain present but undergo change form year to year. I will argue below that these changes in the corpus attest to a process of *adaptation,* as the (implied) teacher must negotiate new norms of assessment to correspond to his expanded instructional goals. These chains are represented in Fig. 5.1 as the vertical lines that link corresponding items in subsequent years.
- A third manner in which the corpus can be analyzed is at the scale of the *sequence of whole assessments*. From the teacher's point of view, each assessment is a sequel to the previous one, and it is not unreasonable to expect that changes in the structural features of assessments (length, duration, etc.) might occur in response to events on the ground. Much of this chapter will be devoted to narrating the evolution of the assessment corpus at this scale, represented in Fig. 5.1 by the dark horizontal and diagonal lines linking each assessment to the next.

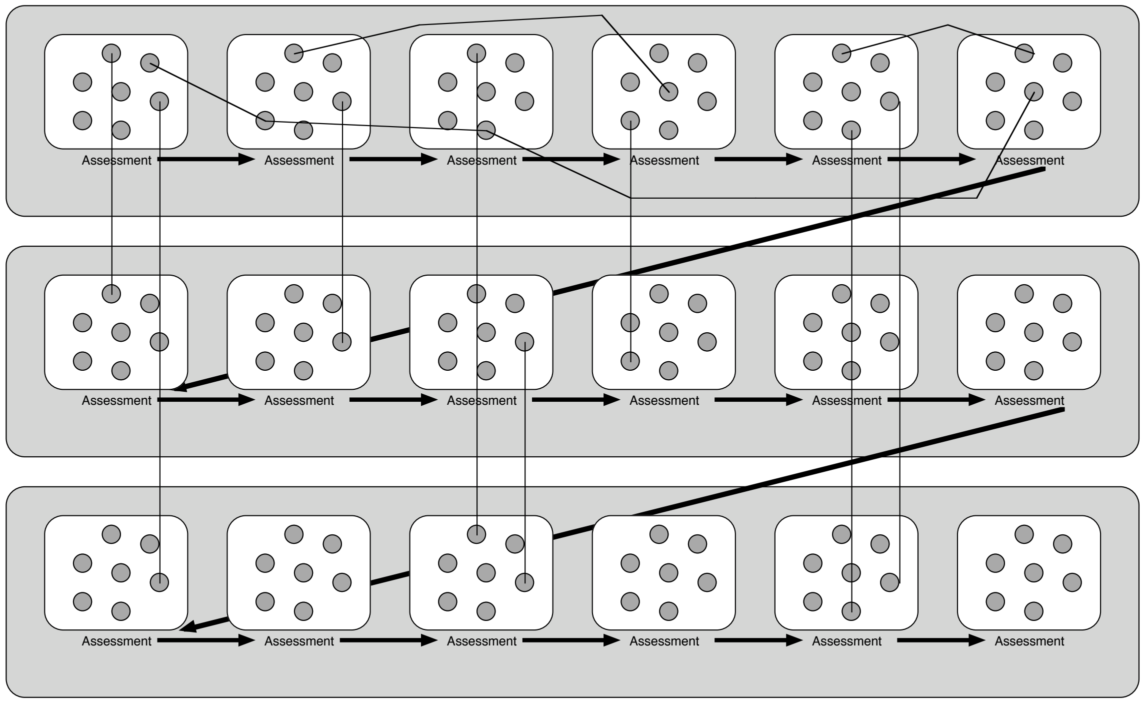


Figure 5.1. The corpus at many timescales

Cohesion with a single year: Item threads in Y2

One of the goals of this chapter is to understand the extent to which cohesion in the corpus allows us to regard it as a single text, with a coherent implied teacher and student. Such cohesion additionally may allow for the deployment of multiple dispositions around a single content area, over an extended time frame on the order of the school year — for example, a diagram representing a particular geometric configuration may be re-used throughout the year, each time for a different purpose and calling for the deployment of different dispositions. Seen from this perspective the unit of analysis is not the individual item, but rather strings or chains of items that create cohesion. Something somewhat analogous (albeit on a smaller scale) is found in Nathan & Long (2002), which analyzed the organization of categories of problem-solving activities in textbooks. In Nathan & Long's methodology, individual exercises were first coded as either S (symbolic) or V (verbal). The unit of analysis, however, was not the individual exercise, but rather *sections of exercises*. Each section was thus coded as either SS (symbol-to-symbol), SV (symbol-to-verbal), VS (verbal-to-symbol), or VV (verbal-to-verbal), according to the way in which the section was organized as a whole. For the present discussion, I am interested not in the scale of the section or even the chapter, but rather the entire year; still, the notion of tracking the *development* of content over time is central to the attention given here to cohesion in the corpus.

To operationalize this form of cohesion, I identify items that either look backward to previous problems or anticipate future interrelations can be manifested explicitly: for example, one question on a chapter exam from the second semester in Year 2 posed a proof exercise very similar to one that had appeared on a previous midterm exam, but included the statement "You may use any theorem in Chapters 1-5... This is much easier than the problem on the midterm." In other cases the interrelationship between items might be less obvious, but could be inferred from the re-use of a diagram, or the posing of several problems that build on one another. The result of such an identification is an *item thread*.

Recall that the linguistic theory of *systemic functional linguistics,* or SFL (Halliday 1994; Martin & Rose 2003), identifies three distinct "metafunctions" of language: the *ideational* (how language represents what is being talked about), *interpersonal* (how language creates positions and relationships between "speaker(s)" and "audience"), and *textual* (how language accounts for the mode and organization of a text). The key notion of SFL is that any text can be analyzed for how it uses language in the service of these three metafunctions. *Cohesion*, meanwhile, refers to the use of lexical and grammatical relationships to maintain coherence in a text and give it meaning (Halliday & Hasan 1976; Hasan 1984). As discussed in Chapter 4, such coherence can be created out of and through any or all of the three metafunctions. That is, cohesion can operate not only at the ideational level (e.g. repeated references to the same thematic content) and at the textual level (e.g. comments such as “As we saw in the last chapter...”), but also the interpersonal level, as in Thompson & Zhou’s (2001) analysis of the role of evaluative statements as cohesive signals.

All three forms of coherence exist in the examinations corpus. To clarify, I am here concerned principally with coherence across and within a *single year*; this coherence helps to create a single text *from the perspective of the implied student*, who of course does not have access to examinations from previous years. There is also coherence, from the teacher’s perspective, across multiple years; this coherence is manifested in part by changes in the structural features of examinations over time, as I will show later. For the present, I restrict my attention to the exams in the second year of the corpus.

Interpersonal cohesion

Scattered throughout the data from Y2, there a number of items that make explicit reference to the teacher and/or student *as individuals*. By this I mean more than just the presence of a first or second-person pronoun, as in “What can we say about...?” or “Show your work”. Stock phrases such as this are part of the normal discourse of teachers and students; they serve in part to establish and reinforce customary power relationships among teachers and students (Pimm 1987; Rowland 1992). I am interested here in the presence of items that testify to some shared experience on the part of teacher and students, or that create an impression of the teacher and of the implied student.

Such content was heavily present in the second exam of Y2. This exam, which covered Chapter 2 of the course textbook (Larson, Boswell & Stiff 2000), focused on propositional logic, including the laws of syllogism, detachment (*modus ponens*), and contrapositives. At the top of the first page, in a block of text offset from the remainder of the text by a solid border, appeared a set of general instructions to the student, including the following comment:

> This may be the first test you've ever taken that has footnotes on it (it's certainly the first one I've ever written). The footnotes are there to provide you with some additional context for the examples on the test, if you're interested, and in one case to give you a hint; but if you find that the footnotes just confuse you, feel free to ignore them.

I note the presence of seven personal pronouns (six 2nd-person, one 1st-person) in this short passage. And indeed the presence of the footnotes, as elements that stand outside of the actual question-and-answer portion of the exam, is one of the more curious elements of this exam. As we will see, the teacher[33] uses the footnotes to engage in a kind of running commentary on the items — a commentary that is less about authority or assessment, and more about intellectual curiosity. This dialogue runs parallel to, and independently of, the main text (as indicated by its spatial separation from the main body of the page).

---

[33] Recall that all claims about the "teacher" in this chapter are meant to refer to the "implied teacher", i.e. the *intentio operis* as revealed through a set of textual strategies.

One question on the exam, and its corresponding footnote, are reproduced in Fig. 5.2. This question is one of several in which students are asked to identify the hypothesis, conclusion, converse, inverse, and contrapositive of a particular conditional statement. What makes this particular statement noteworthy is its allusion to mathematics far outside the scope of a high school Geometry course, and the footnote indicating to students what they are, and are not, required to worry about in answering the question:

The *Riemann Conjecture*[1] : If z is a solution of the equation $\zeta(z) = 0$, then $\mathrm{Re}(z) = \frac{1}{2}$.

Hypothesis: ______________________________

Conclusion: ______________________________

Converse: ______________________________

Inverse: ______________________________

Contrapositive: ______________________________

[1] The symbol ζ is the Greek letter *zeta*, and represents a function called (naturally enough) "the Riemann *zeta* function". It's probably a good idea to just leave it at that. Don't worry about copying the Greek letter beautifully; just make a *C* with a squiggle on the top and another on the bottom.

Fig. 5.2. A question (with its footnote) from the Y2, Chapter 2 exam.

What is even more striking about this example is the way the commentary continue onto the next page of the exam (Fig. 5.3)

The next couple of questions are also about the Riemann Conjecture, but don't worry — you don't need to have gotten the previous questions right to answer the next ones.

The mathematician Riemann proved the following theorem (called the *Prime Number Theorem*) in the 19th century:

"If the Riemann Conjecture is true, then the number of prime numbers below *N* grows logarithmically with *N*."[3]

Figure 5.3. (continued on next page)

7. (*6 points*) Suppose a *«name of school omitted»* student, working diligently and in secret for months, manages to prove the Riemann Conjecture to be true.

(a) Putting this amazing discovery together with the Prime Number Theorem, what can our student conclude?
(b) Which logical law (Syllogism or Detachment) did you use in answering (a)?

8. (*9 points*) Now suppose that your Geometry teacher has managed to prove the following statement, henceforth to be known as "[redacted]'s Theorem": "If the number of primes numbers below *N* grows logarithmically with *N*, then the Artin-Stafford Conjecture is true."[4]

(a) Is it legitimate to conclude that "If the Artin-Stafford Conjecture is true, then the Riemann Conjecture is true"? If so, what logical law (Syllogism or Detachment) justifies this?
(b) Is it legitimate to conclude that "If the Riemann Conjecture is true, then the Artin-Stafford Conjecture is true"? If so, what logical law (Syllogism or Detachment) justifies this?
(c) Is it legitimate to conclude that "The Riemann Conjecture is true if and only if the Artin-Stafford Conjecture is true"? If so, what logical law (Syllogism or Detachment) justifies this?

[3] The question here is, how many primes are there below any given number *N*? For example, there are 4 prime numbers below 10, 8 prime numbers below 20, 10 prime numbers below 30, 12 prime numbers below 40, etc. As you look ever higher, the primes become more and more scarce, so the rate at which the number of primes grows is slowing down. *Logarithmically* is a technical word that describes how fast the number of primes grows. You'll learn about logarithms in Algebra 2 (next year).

[4] Oh, how I wish it were that simple.

Fig. 5.3. More questions and footnotes from the Y2, Ch 2 exam.

The discourse in this passage is complex. The opening sentence addresses the student directly and reassuringly ("You don't need to have gotten the previous questions right"). The statement of the Prime Number Theorem is accompanied by a footnote that explains in capsule form the substance of the theorem, while simultaneously drawing a connection between the rather abstract mathematics of the PNT and the knowledge that students will encounter in their subsequent mathematics courses. Note again the presence and role of personal pronouns and the direct address to the reader here; it is as if the implied student is *on the stage* along with the mathematical content being assessed. This is further reinforced by the hypothetical situations presented in problems 7 and 8, which place (respectively) an imaginary student in the school, and the teacher-author himself, as characters in a historical-mathematical drama — as does the attempt at self-deprecating humor in footnote 4.

The examples above are not alone; among the 116 items in the Y2 portion of the corpus, I have identified 11 instances in which the teacher makes use of language in service of the interpersonal metafunction. Most of these instances, like the ones we have seen, take the form of paratextual comments: footnotes, asides, and other commentary not directly necessary for the student to know what he or she is to do. The cumulative effect of these language choices is to create an impression of the teacher, or rather the image of the teacher that one encounters through the text (cf. Booth's "second self", cited on p. 180 above), and of the student that the teacher addresses. The constancy of these impressions over time is one of the sources of cohesion in the corpus: that is, a reader sense that the several assessments and dozens of items contained within it are all the product of a single persona.

In this context it is worth noting that the assessment items included within the textbook (Larson et al 2000) — the "Chapter Exams" printed at the end of each chapter — avoid the use of personal pronouns and other forms of address nearly entirely. Across the entire 229-item corpus, not a single item was found that creates any sense of an author or a reader; consistently, the elided second-person imperative ("Prove that...", "Solve for the variable...") and passive voice ("Which theorem can be used...") are used to depersonalize the text. This is hardly surprising, of course; the authors of a textbook are many steps further removed from the eventual student-reader than is the teacher who constructs his own exam items for use with students in his own class. The language choices of the textbook serve the interpersonal metafunction in their own way: they create an impression of an impersonal author, one above and removed from the lives of the students.

The significance of the findings above lies in the fact that the teacher who authored this corpus of assessments *could have emulated* the impersonal, textbook style, *but did not do so.* It is in part through this departure from the usual norms of discourse, the customary way in which the interpersonal metafunction is deployed in written assessment items, that this collection of assessment items functions as a source of textual cohesion, creating a coherent and stable sense of an implied teacher.

Ideational and textual cohesion

In addition to establishing and reinforcing a relationship between teacher and student, the items in the examinations corpus also establish and reinforce an ongoing relationship with particular *content areas*. This is of course something one would likely expect in any collection of math exams: to the extent that mathematics is a cumulative subject, material from any point in the year is likely to be invoked at subsequent moments. And to the extent that Geometry is about a certain category of "things" (points, lines, planes, angles, etc.) and the relationships among them (incidence, congruence, parallelism, etc.), one expects those things and relationships to reappear throughout the course of the year. All of the above stands as examples of ideational cohesion, and as such it is neither very surprising nor very interesting to find it throughout the corpus. What *is* both surprising and interesting is the presence of certain recurring themes or motifs that are not "necessary", insofar as one could easily imagine a corpus that omits them without shortchanging the curriculum.

One motif of this sort in the Y2 assessment items is located in a collection of five problems that I refer to as the "duals thread". Each of these five problems asks students to consider one or more properties of a figure that is formed by joining together the midpoints of adjacent sides of a polygon. The first problem in the duals thread appeared on the exam for Chapter 4 (Fig. 5.4a). In this problem, the student is asked to prove (using a "coordinates proof") that any two of the triangles formed by joining together the midpoints of a triangle are congruent.

Subsequently, two related problems (the "fractals" problem and the "antwalk" problem) appeared on the Midterm exam. These two problems are interesting in their own right for a variety of reasons, and I discuss them both in detail below (see pp. 287-292). For the present it suffices to reproduce the diagrams that accompanied those exercises (Fig. 5.4b, 5.4c). The reader will note at once the similarity of these to the diagram in Fig. 5.4a.

The next assessment, the Chapter 5 Exam, included yet another problem that made use of essentially the same diagram, and called for another proof of essentially the same property as did the previous items in the thread (Fig. 5.4d). The principal difference between the problems in Fig. 5.4a and Fig. 5.4d are the operations and resources available to students. The problem in Fig. 5.4d comes after Chapter 5, which introduces a particular theorem (the "midsegment theorem"[34]) that makes the problem fairly simple — a theorem that was not available to students when taking the earlier assessments; hence the parenthetical comments that follow the statement of the task.

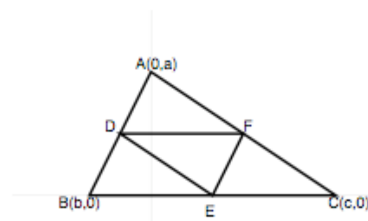
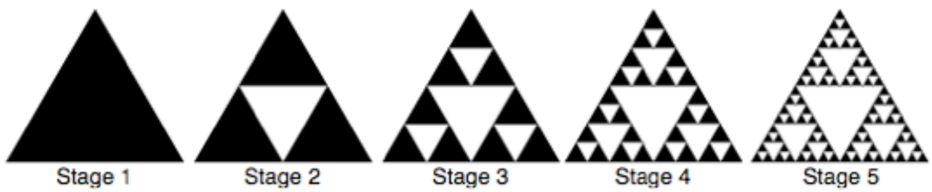
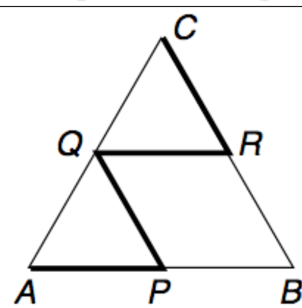
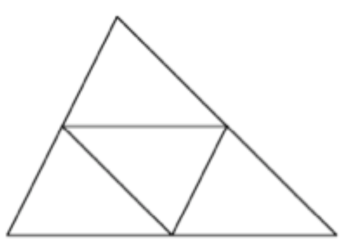
(a)	6. (15 points) In the diagram at right, D is the midpoint of $\overline{AB}$, E is the midpoint of $\overline{BC}$, and F is the midpoint of $\overline{AC}$. Choose any two of the four "small" triangles in the diagram, and prove they are congruent. **Hint:** First prove that $\overline{DE} \parallel \overline{AC}$, $\overline{EF} \parallel \overline{AB}$, and $\overline{DF} \parallel \overline{BC}$. *Note: I strongly recommend using a coordinates proof for this problem.* 
(b)	
(c)	
(d)	3. (*10 points*) The midpoints of the three sides of a triangle are found, and are used as vertices for a new triangle. Prove that ***all four small triangles*** formed in this fashion are congruent. (You will want to label the points on the diagam. You may use any theorem in Chapter 1-5.) [Note: this is much easier than the problem on the midterm.] 

Fig. 5.4. Four problems from the Y2 "duals thread".

[34] The midsegment theorem states that the segment formed by joining the midpoints of two adjacent sides of a triangle is parallel to, and half the length of, the third side of the triangle.

The duals thread reappears for a fifth and last time on the Final Exam, where it forms the ideational backbone of a take-home research project: students are charged with investigating the properties of duals of polygons. As with the problems on the Midterm exam, I will give a full treatment of the Y2 Final Exam in the pages below. Here it is just worth mentioning that the instructions on that Final Exam make explicit reference to the fact that students are expected to be able to say something about the dual of a triangle, and frame the larger research project as an extension and generalization of that.

The existence of the duals thread is significant partly because it does not exist in the textbook items corpus. That is not to say that the textbook does not contain problems like those in Fig. 5.4; on the contrary, it does. But all of them are located in the exam for Chapter 5 — that is, the chapter in which the midsegment theorem, that theorem that was marked as instrumental in Fig. 5.4d, was introduced. The exam items corpus is different not for the mere presence of this content, but for its *repetition throughout the course of the year*. The theme is introduced first on the Chapter 4 Exam, before its "proper place" (as defined by the textbook), where it is difficult to prove. It reappears (in somewhat masked form) in two questions on the midterm. It appears again on the Chapter 5 Exam (its "normal" location). The student is then reminded of it at the end of the year, and asked to review and extend what has been learned. This is what is meant by my claim that the corpus of assessment items is marked by a degree of ideational cohesion that ties it together, transforming it from a collection of independent texts to a single, extended work.

The examples above also illustrate the way the textual metafunction of language is employed to produce coherence. The textual metafunction is that use of language that calls explicit attention to the organization of a text: phrases like "as we saw in the previous chapter", "as I will show below", and "the examples above" all illustrate the textual metafunction. In this connection notice that the problem in Fig. 5.4d makes explicit reference to an earlier member of the duals thread ("this is much easier than the problem on the midterm"). Examples of this sort are rare in the assessment items corpus, but they do exist; additional examples will be described below. Usages like this show that the implied teacher expects the implied students to have a shared memory of problems from a prior exam. That is, the exams (and the items on them) do not stand alone: the discourse within them establishes each question as an installments in an ongoing, continuous interaction among teacher, student, and content.

Adaptation in the corpus

Fundamental to the analysis in this chapter is a perspective on learning and knowing that proceeds from a practice-based point of view. Following Piaget (1975) and von Glasersfeld (1995), I take the position that individuals learn by adapting to feedback provided by the environment in response to their actions. Taking this perspective as a starting point, Brousseau (1997) has added the critical notion of the *milieu* in which learning takes place. Brousseau uses the word "milieu" in a technical sense, to refer to the system counterpart of the learner: the learner acts on and in the milieu, and the milieu in turn provides feedback to the learner.

In the current context, I wish to argue that changes in the corpus over time attest to adaptation in the above sense *on the part of the teacher.* That is to say, I take as a starting point the presumption that, as a teacher interacts with students, administrators, other teachers, parents, and the artifacts of teaching (written homework assignments and tests, lesson plans, report cards), there is the possibility for the teacher to learn in response to the feedback received from the educational milieu. A lesson plan that does not come off how the teacher envisioned it may be modified the next time it is used; a homework assignment that takes too long to grade may lead to shorter assignments, or more cursory grading, in the future; complaints from students, parents and/or administrators about too many failing students may lead to gradual (and even unrecognized) grade inflation. Of course from an outsider's perspective, many of these changes might, to the extent that they take a teacher further away from some ideal of what practice ought to be like, be viewed as maladaptive. But insofar as they serve the teacher's interest by modifying his existing practices to better accommodating feedback from the milieu, I regard them here as a form of learning.

At any given point in the course of a learning trajectory, there exist certain systems of interactions between the learner (here, the teacher) and the milieu. Balacheff (see Balacheff & Gaudin 2002) refers to those systems of practices as *conceptions*, where this term is understood to refer not to unobservable cognitive operations but rather to those interactions that can be observed. In Balacheff's *cK¢* model, a conception is characterized by a set of *problems* that the conception handles, *operators* and *representations* that the learner employs to deal with those problems, and *controls* that are used to evaluate whether the task has been correctly solved, and if not what should be done next. In the current context, consider “Writing an exam to assess students’ learning” as a problem that a teacher must solve periodically. Some of the operations a teacher might use in the process of solving this task are: selecting questions from a textbook or a question bank provided by the textbook publisher; amending those questions in both non-substantive ways (e.g. changing the equations slightly in a problem that uses algebra) and substantive ways (adding or removing parts of a multi-part problem, providing more or less scaffolding, etc.); reusing questions that were discussed in class; or creating new questions from scratch. Among the choices a teacher has available as *representations* of the problem are: the number of items on the test, the number of pages, the amount of white space between items, the amount of time students have to work on the assessment, the presence or absence of diagrams, and so forth. Finally, as *controls* the teacher has various ways of determining whether a test is “going well”, both during and after the fact: Are students asking for clarification? Do too many students finish the test early, or run out of time without completing it? Are there too many failing students, which could indicate that the test was harder than appropriate? Or too many students with perfect grades, which could indicate the opposite? All of these provide some measure of feedback to the teacher that can, at least in principle, result in changes on the next assessment.

Evolution within item classes

The preceding discussion reveals another reason why it is important to document the existence of cohesion within the corpus: it is only in the context of a cohesive body of assessments that changes over time can be interpreted as evidence of teacher learning, i.e.

of adaptation in response to feedback from a milieu[35]. That is to say: if a collection of exam items were to be examined and found to lack significant cohesion, then variations in that collection would not be significant. It is only the existence of a more-or-less stable background that makes changes meaningful.

In this section I turn to an analysis of how individual items evolved over the three-year corpus. In particular, I examine the *item classes* in the corpus (represented schematically in Fig. 5.1 by the vertical lines linking corresponding items in different years) and track the different kinds of changes evident in those item classes. (Recall that an item was said to belong to the same *item class* as an item from a previous year if it was a repetition or an obvious adaptation of the earlier item.) Before turning to the data, some preliminary considerations are in order.

What kind of changes might one expect to find in an item class, and what might such changes testify to? First, there is the possibility that an item might disappear entirely from one year to the next — that is, that an item present in the corpus in Year n might not reappear in the Year ($n + 1$). Such a disappearance may or may not be associated with an overall reduction in the number of problems present from year to year. If the total number of problems present on each exam or within each year does *not* drop, then the disappearance of any single particular question implies that it was *replaced* with another question. A comparison of the dropped questions in one year to the new questions introduced in the next might then suggest why the change took place: perhaps the new problems are shorter, easier, etc. This might then testify to a lowering of

---

[35] To be clear, “learning” here is not intended to refer to something that is necessarily positive according to some theory of what good teaching is like, but only to adaptation. Bad habits are learned just as much as good ones are.

expectations on the teacher's part. On the other hand if the new problems appear more difficult or complex than the ones they replace, we might infer that the teacher is raising expectations of his students. If, alternatively, a number of problems were to disappear *without* being replaced by new ones in subsequent years, this could be taken as a reflection not of the individual items but of the test as a whole: perhaps the teacher came to recognize that it had too many questions, that students could not finish it in the allotted time, etc. Of course even in this question we would still be entitled to ask, Why eliminate *these* questions rather than others? And again there are a number of possibilities that could be suggested by comparing the questions that disappear with those that survive: perhaps the problems eliminated were too difficult, or not difficult enough, or sufficiently similar to the surviving problems that their disappearance could be understood as the elimination of redundancy.

But disappearance is only one of many things that could happen to a problem over time. Another possibility is that a problem might survive, but in modified form. These modifications correspond to the categories that I used in the previous chapter to model items:

- Changes in the *goal* to be produced can take the form of changes in wording (e.g. "show that" becomes "prove that"), or in multi-part problems gaining or losing parts (i.e. an increase or decrease in the number of distinct goals);
- Changes in the *resources* provided to the student in and through the text include the addition or elimination of *scaffolding* (hints, a diagram, breaking a complex problem down into multiple parts, etc.);

- Changes in the *stakes* of a problem could take the form of a reclassification of a required item as an extra credit item, or vice versa.

Moreover the coding of items after mathematical dispositions provides another means of describing changes in an item class. For example, if the version of the item used in Y1 was coded after three dispositions, but the version of the item used in the following year was coded after only one, then we can say that there has been a decrease in the number of mathematical dispositions represented in the item. Similarly a problem can experience an increase in the number of dispositions represented in it.

In any given problem, it may impossible to explain *why* a change in one of the above ways might have occurred. But if the aggregation of all item classes in the corpus displays consistent patterns of change (e.g. if the problems consistently gain additional scaffolding, exhibit reductions in the number of dispositions, and become shorter) then we would be justified in ascribing those changes to the presence of some form of feedback on the teacher.

In fact all of the above kinds of changes are present in the corpus. Fig. 5.5 summarizes the kinds of changes that items underwent over the three-year period of the corpus. (Items can undergo more than one kind of change.) Notice, first, the large number of problems introduced in Y1 that were eventually dropped: fewer than half survived one year, and less than 30% of survived to Y3. The decrease in the number of items over time is part of a more complex trend towards shorter tests in Y1 and Y2; this trend will be discussed in more detail below, when I narrate the changes in the structure of examinations over time.

If we restrict our attention only to those items that repeat for at least one year, and focus on the nature of the changes those items underwent, we find that among the 82 problems that were introduced in Y1 and survived for at least another year, more than 1/3 experienced some form of rewording.

Changes to items introduced in Y1	Number of items (total *N* = 193)	Percent
Drop out after Y1	101	n/a
Drop out after Y2	35	42.7%
Rewording/clarification	29	35.6%
Scaffolding reduced	2	2.4%
Scaffolding added	7	8.5%
Problem lengthened	8	9.8%
Problem shortened	5	6.1%
Dispositions increased	3	3.7%
Dispositions decreased	3	3.7%
Stakes increased	0	0
Stakes decreased	2	2.4%

Figure 5.5. Evolution in the item classes from Y1 to Y2.
Note. Percentages are calculated relative to the 82 problems that survived into Y2.

In addition there were 17 changes that could be regarded as a lowering of the teacher's expectations (increased scaffolding, a reduction in the number of tasks or dispositions, or a repackaging of a problem as lower-stakes, e.g. extra credit). Interestingly, though, these are nearly balanced by 13 changes that could be viewed as an *increase* in the teacher's expectations: decreased scaffolding, an increase in the number of tasks or dispositions, or a repackaging of a problem as higher-stakes. Moreover, while it is conceivable that a single problem might evolve simultaneously in opposite directions — for example, having the stakes raised, but with additional scaffolding provided — there were no problems that underwent change of this sort from Y1 to Y2. Overall there is a slight trend towards making the problems clearer, simpler, and lower-stakes.

If we shift focus to the problems that survive from Y2 to Y3, we see that the situation is somewhat more stable (Fig. 5.6). The surviving items from Y1 still experience a high attrition rate (42%) in the passage from Y2 to Y3, but among those 57 that survive into the third year there are relatively few modifications. The 24 new problems introduced in Y2 fared even better than the 101 problems they replaced: of the newer group, just over 70% were re-used in Y3, most with no modification.

Changes to items after Y2	Number of items	Percent
Items repeated from Y1	*82*	
Dropout after Y2	35	
Rewording/clarification	10	15.6%
Scaffolding added	4	6.3%
Scaffolding reduced	0	0%
Problem lengthened	3	4.7%
Problem shortened	1	1.6%
Dispositions increased	2	3.1%
Dispositions decreased	0	0%
Stakes increased	0	0%
Stakes decreased	2	3.1%
Items introduced in Y2	*24*	
Dropout after Y2	7	
Rewording/clarification	7	10.9%
Scaffolding added	3	4.7%
Scaffolding reduced	0	0%
Problem lengthened	0	0%
Problem shortened	0	0%
Dispositions increased	0	0%
Dispositions decreased	0	0%
Stakes increased	1	1.6%
Stakes decreased	0	0%

Figure 5.6. Evolution in the item classes from Y2 to Y3.
Note. Percentages are calculated relative to the 64 items that survive into Y3.

What does all this mean? If we understand the high rate of attrition and modification of the problems introduced in Y1 as evidence of some indication of a mismatch between what the teacher originally set out to accomplish, and what actually transpired — i.e. as reflecting some sort of negative feedback on the teacher — then the relatively low rate of modification of those problems that remained in the corpus after Y2 stands as an indicator that those items were *well-adapted* to the milieu. Additionally the fact that a larger proportion of the items introduced in Y2 survived into Y3 with little or no change stands as an indication of teacher learning: that is, the teacher's item-writing skills have adapted, in some fashion, to produce items that are more viable in the classroom.

What is the nature of this learning? To answer this question, I now provide selected examples of the various kinds of item evolution in the corpus, and discuss what they may say about the interaction of the teacher with the milieu of teaching.

Examples of item evolution

Figures 5.5 and 5.6 both show that many items survived from one year to the next with modifications in wording. Figure 5.7 illustrates this phenomenon with three pairs of items that exhibit such changes. Although the changes in wording are minor in each pair, the details of the changes are highly suggestive. Consider the change from 5.7a to 5.7b: The addition of the sentence "if there is more than one answer, give both" not only signals to students that they ought to be on the lookout for more than one possibility; it also leads us to consider the possibility that the first problem might have "failed" by eliciting only a single answer from too many students. If this were so, then the rewording can be understood as evidence that the teacher learned to become more explicit with his expectations.

Or consider the change from Fig. 5.7c to Fig. 5.7d. Why might a teacher change "Give an example" to "Make up your own example"? Perhaps with the original version of the problem, a significant fraction of the class provided examples that had been previously discussed in class. Certainly a student might have thought such a response was appropriate. If so, then the teacher's decision to change the problem might have been motivated by a desire to test whether students understood the principle behind those examples, rather than simply their ability to recall the example.

On a number line, point *S* has coordinate 1. If *ST* = 5, then what could be the coordinate of point *T*? (a)	On a number line, point *S* has coordinate 1. If *ST* = 5, then what could be the coordinate of point *T*? If there is more than one answer, give both. (b)
Give an example of a sentence which is false, but has a true converse. (If this is not possible, explain why.) (c)	Make up your own example of a sentence which is false, but has a true converse. (If this is not possible, explain why.) (d)
Are the points *A*(4, 5), *B*(-3, 3), *C*(-6, -13) and *D*(6, -2) the vertices of a kite? Explain your answer (e)	Are the points *A*(4, 5), *B*(-3, 3), *C*(-6, -13) and *D*(6, -2) the vertices of a kite? Justify your answer. (f)

Figure 5.7. Items that underwent rewording from one year to the next.

On the other hand it is hard to know what to make of the change from Fig. 5.7e to Fig. 5.7f. These two items differ only in a single word — "explain" is replaced with "justify" — but it is not completely clear what, if anything, that change signifies. If the teacher had changed the word "explain" to the word "prove", or even a phrase such as "write a paragraph proof", then the teacher's intentions would be more transparent. We might speculate that "justify" had a shared meaning within the discourse of the classroom, one that was more specific than "explain". But the evidence in Fig. 5.7 is not really sufficient to support such a conclusion.

There is one case of a problem that underwent successive rewording in each year of the corpus. That problem (Fig. 5.8) asks students to derive some consequences about a triangle from a hypothesis stated in the form of a congruence. Recall that the final version (Fig. 5.8c) was used in Chapter 4 (Fig. 4.20, p. 232) to illustrate the code *Formalism*. It is also worth noting that a proof of the base angles theorem along similar lines was denounced by a teacher in the ThEMaT Study Group as "stupid" (ITH121505, interval 59). Here we see the problem evolve in a very consistent manner: More text is added to the statement of the task to both motivate the problem and to clarify what is expected. One can almost see the teacher struggling to find a way to make this problem viable, despite what (it seems) must have been strong feedback to the contrary.

What does $\triangle ABC \cong \triangle BCA$ tell you? (a)	If $\triangle ABC \cong \triangle BCA$, what can you conclude about the triangle? Justify your answer. (b)	As you know, when writing a triangle congruence statement, the order of the letters matters. Suppose for three points *A, B,* and *C* it happens to be true that $\triangle ABC \cong \triangle CAB$. What can you conclude about $\triangle ABC$ in this case? Be as detailed as possible. (c)

Figure 5.8. Three versions of the same item in successive years.

Intriguingly, another problem in the Y1 corpus — one quite similar to that in Fig. 5.8a — failed to adapt in this fashion, and dropped out of the corpus after one year. That problem (Fig. 5.9) seems intended to target approximately the same mathematical ideas (i.e., that the notational convention for similarity and congruence indicate correspondences between the constituent parts of the geometric objects). Consider the possible solutions that a student might produce: Students who have understood the formalist point — that the correspondence between congruent parts is indicated by the notational convention — will ignore the diagram, and read off from the statement of the congruent pentagons that $\overline{AB} \simeq \overline{RQ}$, $\overline{BC} \simeq \overline{QP}$, $\angle C \cong \angle P$, etc. On the other hand a student who has failed to learn that point is likely to be guided by the *visual* correspondence between the two pentagons in the diagram, in which *A* appears to match up with *P*, *E* with *T*, and so forth, leading to incorrect conclusions such as $\overline{AB} \simeq \overline{PQ}$, $\overline{BC} \simeq \overline{QR}$, $\angle C \cong \angle R$, etc. In terms of what students are asked to do, then, this problem "means" that a student is supposed to know to follow the notational correspondence and ignore (even actively suppress information coming from) the diagrams.

In the diagram below, $ABCDE \cong RQPTS$. List all the angle and segment congruences that you can based on this information.

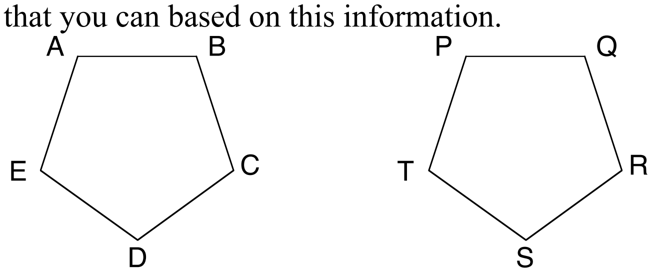


Figure 5.9. A problem that dropped out after Y1.

The problem (in its various incarnations) of Fig. 5.8 can be described in the same way. But in addition, that problem could be regarded as standing for: an (implicit) statement of the theorem that "equilateral" and "equiangular" are one and the same property for triangles, an (also implicit) connection between the formal properties of an object and the symmetries of its diagram, and (more explicitly) the notion that a single figure can be regarded as more than one formal object. Nothing analogous to these can be said to be present in the problem of Fig. 5.9. In this light we can perhaps understand why the teacher kept trying to find a way to make the triangle version of the problem “work”, but was willing to give up after only one try with the pentagon problem: only the former problem had a sufficiently high mathematical value to justify an ongoing attempt to make it viable in the face of what (we can only presume) must have been consistent negative feedback.

Figs. 5.5 and 5.6 also show that a (smaller) set of problems underwent changes in the number of parts students were required to complete, and/or the number of dispositions after which they were coded. Fig. 5.10 shows two illustrations of this. The item in Fig. 5.10a (from Y1) was reduced both in number of parts (the “formal logic” portion was eliminated) and in dispositions (the version in Fig. 5.10a was coded for *Formalism*, the version in Fig. 5.10b was not). In contrast, the item in Fig. 5.10c experienced an *increase* in both length and dispositions: the version in Fig. 5.10d (from Y2) calls for a proof of a biconditional statement, and includes a diagram that shows a condition opposite to those which students are to prove equivalent, stressing the contingent nature of the hypothesis and the conclusion, and the link between them. I note in passing that the refracting light item (Fig. 5.10c) was cited in the previous chapter as an illustration of the *Utility* code, and the double-reflection item (Fig. 5.10d) was cited in that same chapter as an illustration of *Converse*.

Below are several examples of "logical" reasoning. Some of the examples use *legitimate* (i.e. valid) logic, and some of the examples use *incorrect* logic. For each example, circle the symbolic expression that best represents the kind of logic used. Then indicate whether it is valid or not; and if it is, why?

Given: If you want to sing the blues, you've got to pay your dues.
Given: You don't want to sing the blues.
Conclude: You don't want to pay your dues.

(a) Given $a \Rightarrow b$ and given $\sim b$, conclude $\sim a$.
(b) Given $a \Rightarrow b$ and given a, conclude b.
(c) Given $a \Rightarrow b$ and given $\sim a$, conclude $\sim b$.
(d) Given $a \Rightarrow b$ and given $b \Rightarrow c$, conclude $\sim c \Rightarrow \sim b$.

Is it valid logic?
If you answered "yes", what logical rule(s) does this use?

(a)

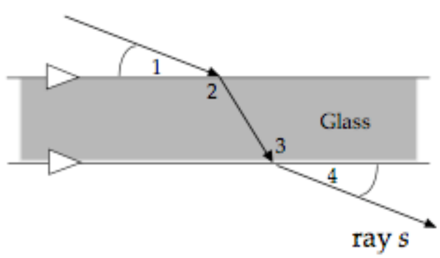
Below are several examples of "logical" reasoning. Some of the examples use *legitimate* (i.e. valid) logic, and some of the examples use *incorrect* logic. For each example, indicate whether the reasoning is valid or not; and if it is, why?

Given: If you want to sing the blues, you've got to pay your dues.
Given: You don't want to sing the blues.
Conclude: You don't have to pay your dues.

Is it valid logic?
If you answered "yes", what logical rule(s) does this use?

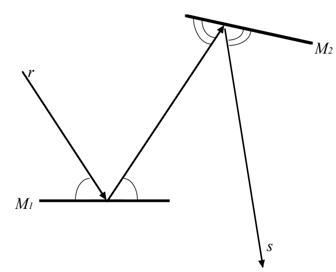
(b)

When light enters glass, the light bends. When it leaves glass, it bends again. If both sides of a pane of glass are parallel, light will leave the pane at the same angle at which it entered. (See diagram) Prove that the path of the exiting light is parallel to the path of the entering light.



(c)

When a ray of light reflects off of a mirror, the angle of incidence is always equal to the angle of reflection. In the figure at right, an incoming light ray *r* reflects off of two mirrors M_1 and M_2. Show that the outgoing light ray *s* is parallel to *r* **if and only if** the two mirrors are parallel to each other. [Before you start working on this problem, make sure you understand what "if and only if" means.]



(d)

Figure 5.10. Items that gained or lost parts or dispositions over time.

Summary of item class analysis.

In the preceding pages I have shown, on the basis of the data (Fig. 5.5 and Fig. 5.6) and illustrated with specific examples, that an analysis of change within item classes over the three-year period of the corpus reveals evidence of some kind of adaptation on the part of the teacher. While it is impossible to make any definitive claims as to what the teacher adapts *to*, there does seem to be a clear trend towards increased scaffolding and more explanatory text over time, which suggests that students' performance was other than what the teacher was hoping for — for example that grades were too poor, or were distributed too unevenly across the class, or too many students had difficulty finishing the class in the allotted time, or too many students needed to ask clarifying questions during the exam. On the other hand there does *not* seem to be a corresponding trend towards shorter problems, or fewer mathematical dispositions in those problems that survive. On the contrary there are some indications to suggest that the teacher may have provided additional scaffolding or explanation to precisely those items that were longer and most representative of the mathematical sensibility, perhaps in an effort to keep them viable despite some negative feedback from the milieu.

In the remainder of this chapter, I will step away from the scale of the individual items, the item classes, and the item threads, and look at the structure of each assessment taken as a whole. We will see that we find largely the same phenomenon at this larger grain size: although there is a long-term trend toward fewer and shorter assessments, there are indications that the teacher resists the pressure toward reducing the complexity of the assessments, instead providing additional resources to the students to compensate for whatever negative feedback may be present.

A narrative of examinations

In this section I describe the corpus at the scale of the examinations, each taken as a *single unit*. At this scale, the features that emerge are descriptive not of individual questions, but rather of the whole assessment: how many questions appear on it, how many points the entire exam is worth, how much time is allowed to students to complete the exam, how much time elapses between assessments, what material resources (e.g. calculator, note card, etc.) students were allowed to make use of, and so on. When examined from this perspective, the corpus looks like a single sequence of units, with the first unit in Year 2 a direct successor to the last in Year 1, and so on: that is, we are following the dark horizontal and diagonal arrows in Figure 5.1. The fundamental question of interests here is, *What changes and what remains constant across this sequence of assessments?*

This question is significant for two distinct reasons. The first reason has to do with the stance I have adopted with respect to the "implied teacher" of the corpus. The contention that this corpus implies a *single* teacher (rather than a multiplicity of implied teachers, one for each assessment) rests on the premise that the corpus can fruitfully be regarded as a single extended *text;* and this in turn relies on the presumption that there may exist cohesion across the corpus. That is, if the basic structure of examinations remains more or less steady from one exam to the next, this consistency enables us to infer a consistent authorial presence behind the corpus.

The second reason to care about this question is, in a sense, complementary to the first: if consistency across the corpus supports the inference of an implied teacher, then *change* across the timeline of the corpus can provide evidence of *adaptation* on the part of that implied teacher. To take a simple example: if the number of questions on each exam were to steadily decline over the first half of the corpus, and thereafter remain stable, it would be entirely reasonable to infer that the teacher is responding to some kind of feedback from the instructional milieu: perhaps too many students are unable to complete the exams in the time allotted, or the average grade in the class is too low, or the exams take too long to grade, etc. In terms of the conceptions of teaching discussed earlier in this chapter, the decline in the number of questions would then be evidence of a process of the teacher's adaptation to an instructional "problem", while the eventual stabilization would be understood as the emergence of an instructionally viable "solution" to that problem.

The first year of the corpus: Negotiating the measurement of achievement

As can be seen from Figure 5.11 (below), the eleven exams in Year 1 corresponded to Chapters 1-11 of the textbook in an essentially one-to-one fashion: although there were some exceptions, such as the absence of an exam dedicated exclusively to Chapter 5 (the content of which was included in the Midterm Exam) and a combined exam for Chapters 7-8, the assessment schedule in the classroom conformed to the default chronology of the chapter divisions in the textbook. Note that the exams were spaced roughly at 3-4 week intervals.

Certain features of the exams in this portion of the corpus deserve special mention. First, four of the first five examinations lacked any explicit designation of the "point values" for individual items or for the assessment as a whole; for those assessments, the data in Figure 5.11 are taken from the teacher's hand-written grading guide, where those were available. Note also that the total number of points possible on those early exams does not conform to the common practice of normalizing assessments to a 100-point scale. In contrast, after the midterm exam, all items on every assessment are accompanied by an explicit designation of how many points they are worth; at the same time, although the 100-point scale never becomes quite universal, it seems to emerge in the second half of the year as a default.

Assessment	Date	# of items[a]	Points Possible
Chapter 1 Exam	9/25/2001	17 (22)	n/a[b]
Chapter 2 Exam	11/01/2001	18 (18)	89
Chapter 3 Exam	11/20/2001	17 (17)	92[b]
Chapter 4 Exam	12/13/2001	12 (12)	57[b]
Midterm Exam (inc. Ch. 5)	1/23/2002	29 (31)	135[b]
Chapter 6 Exam	2/21/2002	15 (15)	100
Chapter 7-8 Exam	3/21/2002	15 (15)	85
Chapter 9 Exam	4/24/2002	8 (11)	101
Chapter 10 Exam	5/14/2002	8 (18)	100
Chapter 11 Exam	5/30/2002	7 (7)	85
Final Exam	6/13/2002	27 (27)	27

Notes. [a] The first number reported is the number of items as designated on the exam itself; the second number is the number of items as entered into the corpus database. [b] Point values were not printed on the exam questions given to students, but were found on the teacher's grading guide (not extant for Chapter 1).

Figure 5.11. Assessments in Year 1.

What can one say about the (implied) teacher, and his adaptation to the teaching milieu, based on these few observations? In the first place, it must be said that the values of questions on an exam do not "just happen" to total 100 points; when such a phenomenon occurs, it indicates a deliberate effort on the part of the teacher. Why do teachers make such an effort? Several possibilities come to mind. Recall that Kulm (1990) includes *feedback to students* and *communication of expectations* as two of the functions of assessment. Examinations with point values that sum to 100 serve both of these functions: On the one hand, a standard 100 point scale makes it possible (for the teacher, but also for the student, parents, administrators, etc.) to easily compare a students' grades across assessments and tell at a glance whether he or she is improving, struggling, etc. Moreover, the 100 point scale creates a default correspondence with letter grades: 90-100 = A, 80-90 = B, and so on. In referring to these 10-point letter bands as the "default", I do not mean not to say that the 10-point letter bands are universal, but merely that they are common, and that they (like the 100-point scale) create a natural way to interpret a students' performance on a test as a measure of achievement.

Thus the 100 point scale serves the teacher's interest in tracking students' ability, and also provides the student (and parents, administrators, etc.) with clear and immediately understandable feedback on his or her performance. Furthermore, the inclusion of point values on the exam itself communicates the relative weight of different questions; such information discloses to the time-pressed student the costs and benefits of abandoning a difficult problem to work on a more tractable one, enabling the student to perform a kind of "triage" during the exam.

The *absence* of a 100-point scale and of designated point values on the exam, then, suggests a teacher who is somewhat insensitive to the communicative functions of assessment, perhaps a teacher more focused more on the content of the individual questions themselves that on letting the student know what is and is not most important to do on the test. The teacher we see in the first half of Year 1 appears unaware of, or at least unconcerned with, the need to let students know ahead of time how important questions are relative to one another, or what their score "means" in terms of a standard A-F letter scale. That this state of affairs changes following the Midterm Exam may be taken as indicating some "learning" on the part of the teacher — that is, a response to some feedback from the milieu. The timing of this change may or may not be coincidental, but at the very least it strongly suggests that the need to mark students' semester grades on their report cards, and the attendant communications with parents and administrators concerning student progress, may have contributed to this "learning".

Another clear trend visible in Fig. 5.11 is the gradual reduction in the number of problems per assessment. With the exception of the Midterm and Final Exam (for which students had extra time available), students encountered an average of 16 problems on each exam in the first semester, but only an average of 10.6 problems on each exam in the second semester. There are many ways this phenomenon might be interpreted. On the one hand, it could be understood as indicating a *lowering of the teacher's expectations of his students* — possibly in response to poor student performance in the first semester. But there are other possibilities: for example, it is conceivable that the assessments in the second semester might be *fewer in number* but also *richer in content*, which would indicate an increase in the teacher's deliberateness in selecting and authoring items.

To gain greater understanding of the meaning of this reduction, I examined each assessment to see how many of its items received one or more code for one of the mathematical dispositions (see previous chapter). Figure 5.12 tabulates the number and "density" of dispositions that are present in each assessment. It will be immediately noticed that both of these drop precipitously in the second half of the year: the four assessments prior to the midterm had an aggregated disposition density of 0.67, compared with only 0.27 after the midterm. This, taken together with the overall decline in the number of items, suggests that the decline in the number of assessment items can be substantially accounted for by the reduction in the number of items that attend to one or more of the mathematical dispositions.

Assessment	# of items[a]	Dispositions[b]	Disposition density
Chapter 1 Exam	17 (22)	9 (10)	0.45
Chapter 2 Exam	18 (18)	7 (9)	0.50
Chapter 3 Exam	17 (17)	7 (13)	0.76
Chapter 4 Exam	12 (12)	7 (14)	1.17
Midterm Exam (inc. Ch. 5)	29 (31)	12 (13)	0.42
Chapter 6 Exam	15 (15)	7 (12)	0.80
Chapter 7-8 Exam	15 (15)	5 (7)	0.47
Chapter 9 Exam	8 (11)	1 (2)	0.18
Chapter 10 Exam	8 (18)	0 (0)	0.00
Chapter 11 Exam	7 (7)	2 (2)	0.29
Final Exam	27 (27)	2 (2)	0.07

Notes. [a] The first number reported is the number of items as designated on the exam itself; the second number is the number of items as entered into the corpus database. [b] The first number reported is the number of items that were coded for *at least* one disposition; the second number is the total number of dispositions coded for in the assessment. [c] The disposition density is the ratio of the (total) number of dispositions on the assessment to the number of items as entered into the database.

Figure 5.12. The declining role of the dispositions in Year 1.

It is tempting to regard this reduction as evidence for some kind of adaptation to the milieu of teaching — that is, to infer that the teacher experienced some kind of negative feedback in the first semester, and responded to that feedback by severely reducing his expectations of students vis-à-vis the dispositions. But there are other possibilities that should be considered; for example, it may be that the mathematical content of the latter chapters of the textbook (polygons, circles, and solids) is not as well-suited for "disposition questions" as is the content of the earlier chapters (parallelism, triangles, and quadrilaterals).[36] Or it may be that some other set of expectations emerges

[36] In this connection it is worth noting that the teachers in the ThEMaT study groups (see Chapter 3) were in general agreement that the latter half of the year contains fewer proofs than does the first half. While "doing proofs" and "learning the mathematical dispositions" are certainly not synonymous, there is undoubtedly some close correspondence between the two. On the other hand, the reduced prominence of proof in the second half of the year is not an intrinsic property of the mathematical content, but rather an instructional phenomenon that has yet to be fully explained.

in the second semester and crowds out the mathematical dispositions. In order to fully understand the phenomenon, we would need to look for trends *across* the 3-year corpus. If such an examination were to show a consistent reduction in the presence of the dispositions from one year to the next, we would be more justified in ascribing this decline to adaptation. As we will see below, the full story is more complex than that.

Renegotiating norms of assessment: Year 2

Following a one-year gap (during which the teacher taught other courses), the teacher returned to teaching geometry. Figure 5.13 summarizes the assessment instruments used in this second year of the corpus. Some trends are worth noting here.

Assessment	**Date**	**# of items**[a]	**Points Possible**
Chapter 1 Exam	9/16/2003	14 (17)	100
Chapter 2 Exam	10/21/2003	13 (17)	90[b]
Chapter 3 Exam	11/14/2003	14 (14)	100
Chapter 4 Exam	12/18/2003	8 (8)	100
Midterm Exam	1/9/2004	10 (10)	100
Chapter 5 Exam	2/26/3005	7 (11)	83[c]
Chapter 6 Exam	4/1/2004	11 (11)	88[d]
Chapter 7-8 Exam	5/6/2004	10 (10)	80
Final	6/3/2004	18 (18)	34[e]

Notes. [a] The first number reported is the number of items as designated on the exam itself; the second number is the number of items as entered into the corpus database. [b] Includes one question worth 5 points "extra credit". [c] Includes 5 points for completing top portion of page 1, and one question worth 5 points "extra credit". [d] Includes one question worth 8 points "extra credit". [e] Includes 17 questions worth 1 point each, and one take-home problem worth the remaining 50%.

Figure 5.13. Assessments in Year 2

One noteworthy trend visible in Figure 5.13 concerns the amount of time between exams. Whereas the exams in Year 1 were spaced roughly 3-4 weeks apart, the exams in Year 2 are spaced 4-5 weeks apart. This additional time spent on each chapter has a cumulative effect: by the middle of May in Year 2, the class is nearly two full chapters "behind schedule" (if we take the Year 1 exams as defining a normative timeline of the year), ground that is never fully recovered: note the complete absence of assessments focused on Chapters 9, 10, and 11 in Year 2.

A second trend concerns the default 100-point scale, discussed earlier. In light of the discussion above, it is surprising to note that the teacher in Year 2 appears to be following a trajectory that is precisely the reverse of that which was visible in Year 1: although four of the first five exams conform to the "100 points possible" standard, after the midterm exam the standard is completely abandoned. This requires explanation.

Indeed the midterm exam marks a significant transition moment, not only in Year 2 but across the entire 3-year corpus. The abandonment of the 100-point scale is not the only significant change that occurs here: the Year 2 midterm exam also stands as the first assessment in the corpus to consist in part or entirely of a *take-home problem set*. To be precise, the midterm exam consisted of 10 questions, each worth 10 points, to be completed outside of scheduled class meetings. Students were given one week to complete the problems, were given specific instructions regarding the resources they were permitted to make use of, and were asked to sign an honor statement. The first page of the midterm exam, containing these instructions, is reproduced in Fig. 5.14.

Honors Geometry
Midterm Exam 2004

Instructions

This is an open-book, open-note, take-home midterm exam. You may not work with another student, teacher, parent, or tutor on these questions, although you *may* ask Mr. [redacted] for help with concepts and methods. Please make sure you sign the honor statement below.

Write all your answers on separate sheets, and attach them to this test *in order*. All answers must be written clearly and legibly. Typing is optional. If you have messy handwriting, or do the problems out of order, you may want to recopy all of your final answers neatly and turn in a "final draft".

Every question on this test requires work. You must show all work. Solutions without work will not be given any credit.

There are 10 questions on this exam; each is worth 10 points, for a total of 100 points possible.

The test is due Thursday, January 15, 2004, no later than 1:00 PM.

HONOR STATEMENT: I attest that I completed this exam without any assistance from any other student, teacher, parent, or tutor.

____ (signature)

Figure 5.14. The first page of the take home midterm exam (Year 2).

The content of the midterm exam will be described below in detail, when I describe the emergence of open-ended problems as a phenomenon of the corpus. For the moment our attention is on the attributes of the *individual exams* (rather than the attributes of the *items* on the exam), and the ways in which those attributes change over time. From this point of view there are several salient points to be observed. The first is that the decision to use a take-home problem set in place of a traditional, in-school midterm exam poses significant risks for the teacher, in that it makes it impossible for a teacher to monitor students as they work and ensure that the work they submit is actually their own. That is to say, the risk of *cheating* becomes a significant problem for the teacher. The fact that the teacher's instructions to the students (Fig. 5.14) spell out explicitly that students may not seek the assistance of others (students, teachers, parents, or tutors) indicates that the teacher was fully aware of this risk. In this context the honor statement can be understood as an attempt to articulate why a student *should not* cheat: not because a student is likely to get caught, but because it would be *dishonorable* to do so. In effect the teacher is telling students, "I know you could cheat, but I trust that you will not do so." The student tempted to seek unauthorized help is being told that to do so might gain them a higher grade, but only at the cost of their own integrity. Further, the teacher's offer to provide help and assistance can be understood as an attempt to forestall any illicit student activity: it says that a student who needs help can seek it directly from the teacher, and thus has no need to make use of unsanctioned resources.

What was the result of this "experiment" on the teacher's part? To answer that question we can look forward into the remainder of the corpus. Although the remainder of the chapter assessments for the year were traditional, in-school examinations, the Final Exam at the end of Year 2 consisted of two parts: an in-school portion (consisting of 17 multiple-choice questions) and a take-home portion (consisting of a single, open-ended problem, referred to as a "research project"), each worth 50% of the exam grade. I will describe the content of the open-ended problem in detail below; here I once again restrict my attention to the *form* of the exam. In the present discussion the teacher's instructions to the students (Fig. 5.15) are noteworthy in part because of the wide range the teacher indicates in what will be expected: anywhere from 2 to 15 pages worth of typed, illustrated text. Notice also that the teacher gives no indication of how the exam will be graded, and indeed we will see later that the very openness of the problem would make the creation of a rubric very problematic. All of this suggests that the take-home portion of the final will necessarily be graded using somewhat subjective (or at least tacit) criteria. Balancing this is an almost obsessive degree of precision in the policy for projects turned in late. This policy — and the complementary "early bird bonus" incentive for students to turn in their work early— might be interpreted as a preemptive attempt to manage the challenge of grading all of those research projects, some of which might be expected to run as long as 15 pages, in time to mark final report cards.

The fact that the teacher chose to once again include a significant take-home component in the assessment provides indirect evidence that the midterm experiment had been, from the teacher's perspective, a success worth repeating. (Note the reappearance of the honor statement.) On the other hand, the balancing of this take-home portion with a more traditional, in-class multiple choice exam could be interpreted variously. One possibility is that the teacher felt a need to balance the somewhat subjective nature of the take-home portion with a clear-cut, objective measure of students' attainment. Conducting the multiple-choice portion of the exam *in class* (rather than making it part of the take-home exam) could be viewed as a means for the teacher to regain some of the control that is lost when students do their work out of the teacher's watchful eye. And it should not go unmentioned that a multiple choice exam can be graded quickly, a perhaps welcome counterpoint to the time-consuming nature of grading research papers.

I assume that your research report will be typed, illustrated with lots of diagrams, and proofread for grammar and spelling. I anticipate most of you will produce reports that will be in the neighborhood of 5 pages long (double-spaced) – I can't imagine doing this well in fewer than 2 full pages, but some of you *might* end up writing 10-15 pages or more.

I hope to take the best research reports and combine them to produce something that we can submit for publication in journals of teaching mathematics. With hard work and some imagination, you may end with your work in print in a nationwide publication!

Honor Statement
The first page of your research report must contain the following honor statement, written in your own handwriting and signed:

> HONOR STATEMENT: I attest that I completed this exam without any assistance from any other student, teacher, parent, or tutor.
>
> ________________________________ (signature)

Research reports that do not contain the honor statement *will not be accepted*.

Due Date:
This portion of the final exam is due **Monday, June 7th**, no later than **12:00 noon.** Papers turned in late will face a 0.4% reduction in your grade for every hour it is late. So, for example, papers turned in Tuesday at 1:00 PM (25 hours late) will lose 10%. (If you email it to me Monday evening, I will look at the time the paper was sent, not the time it was received.)

I will be glad to receive your work early. Turn your work in on or before Friday, June 4th at noon, and you will receive a 5% "early bird" bonus on your work.

Figure 5.15. A portion of the instructions for the Year 2 final exam.

Looking further ahead, this model of assessment (take-home portion + in-class portion) reappears in the third year of the corpus. We will see below that once again the teacher used a take-home Midterm. But even more significantly, in Year 3 *every assessment after the Midterm* contained a take-home component. In this sense we can say that the use of take-home assessments, which was first implemented in the Midterm Exam of Year 2, eventually stabilized as a normal part of assessment in the corpus.

Take-home exams both entitle and oblige students to make decisions about how much time they will devote to the work, in a way that in-class exams typically would not be expected to. But in fact one of the most interesting changes in Y2 concerns the timing and duration of in-school assessments. The Chapter 2 exam contains a new element at the top of the first page: a pair of checkboxes, just below the assessment title and above the blank line for students' names, with which students were to indicate whether they were done with the exam or needed additional time (see Figure 5.16). This new feature occurs, without modification, at the top of the first page of every subsequent assessment (except the Final) in the remainder of the year.

Honors Geometry
Chapter 2 Test — 10/22/03

❑ *Don't grade yet, I want extra time.* ❑ *Grade it now, I'm done.*

Name ____________________________

Figure 5.16. The top portion of the Year 2, Chapter 2 exam.

Although this "extra time" checkbox survives the year, it undergoes an interesting modification at the beginning of Year 3. In Exam 1 of the first year, students once again are given the option of checking a box to request extra time — but this time, an explicit *cost* is attached to exercising that option (Fig. 5.17): students are informed that taking extra time will lower their total exam grade by the equivalent of 5%. It also informs students that they may, after the exam is graded, choose to retake the exam, at the cost of a 10% reduction in their total possible score.

There are 80 points possible on this test + 8 points available extra credit.

If you do not finish the test in the time available, you may "purchase" extra time at a cost of 4 points (and you forfeit the right to do the extra credit). ***Please check one of the two boxes below when you turn the test in***.

❑ Don't grade yet, I want extra time. ❑ Grade it now, I'm done.

After the test is graded, you can decide to do a retake at a cost of 8 points (plus you forfeit the right to do the extra credit.) You don't need to decide now.

Figure 5.17. The top portion of the Year 3, Chapter 1 exam.

The emergence and evolution of the "extra time" checkbox illustrates the way in which an examination of documents can shed light on the teacher's interaction with the instructional milieu. What inferences can we draw from this phenomenon? It certainly seems natural to interpret these changes as attesting to an *adaptation* on the teacher's part; that is, to propose that the teacher was responding to an instructional problem. The precise nature of that problem cannot be identified with certainty, but one possibility that seems plausible is that perhaps in the first half of the year a substantial fraction of the students were having difficulty completing the exams within the allotted timeframes. If so, the result would be poor grades for those students. This in turn could lead to negative feedback on the teacher, who might have to justify his assessment procedures to students, parents, and administrators. Providing extra time to any student who requests it would "solve" that "problem" for the teacher. *Requiring* students to explicitly indicate whether or not they are finished helps to forestall the possibility of a student asking for extra time after his or her exam has already been graded. It also shifts the locus of responsibility for determining when an exam is over from the teacher (as is customary) to the student.

On the other hand, giving all students the right to extend the duration of the exam at will could open a Pandora's Box of new problems for the teacher: If a teacher grades and returns exams to those students who have indicated that they are finished, there is the risk of correct solutions "leaking" to students who are not yet done, and this could potentially compromise the meaningfulness of the results. On the other hand if the teacher does not return the exams until every student has finished, the communicative function of the assessment (giving meaningful feedback to students and other stakeholders on their performance) is undermined (feedback given to students long after they complete their work may come too late to be useful). Moreover it is easy to imagine the logistical difficulties of scheduling additional time (after school, at lunch, during study hall, etc.) for students to work on their exams. And the open-ended timeframe for completing the exam competes with the decidedly non-negotiable schedule of semesters and quarters, parent-teacher conferences and report card marking.

These considerations suggest that if the “extra time” option signals an attempt to solve a teaching problem, the introduction of a “cost” for exercising that option signals a kind of second-order correction: it creates a disincentive for students who might otherwise check the box in order to increase their score by a small increment. A student who is far from ready to have his exam graded might sensibly decide that the potential gain exceeds the 5% cost of admission, but a student who anticipates getting a score in the high 80% or low 90% range has much less to gain and would therefore probably be less likely to exercise the option.

Similar considerations apply to the 10% “purchase price” for retaking an exam. Prior to Year 3, Exam 1, there is no evidence in the documents to tell us whether students were, or were not, allowed to retake exams. One might imagine that students were allowed to request retakes on an ad hoc basis, but that the number of such requests eventually became too large for the teacher to manage without instituting some form of disincentive. Alternatively one might imagine that students were not previously allowed to retake exams, and that the institution of a retake policy in Year 3 constitutes a further development of the “extra time” policy. Regardless of which is the case, it seems clear that the institution of the retake policy serves dual functions: it provides a regularized way of handling students' requests for retakes, while simultaneously creating a disincentive to discourage frivolous requests and to keep the volume of such requests at a manageable level.

We thus see that the structure of examination in Year 2 undergoes changes in the use of time and other resources, by both teacher and students, in three distinct ways: more time is devoted to the content of each chapter; and more autonomy is devolved to individual students in determining how much time to spend on their exam; and a substantial portion of the assessments are transformed into open-book, open-note take-home problem sets. We also see that the examination structure *takes away* from students a resource that had previously been introduced, namely the 100-point scale which, as I argued above, plays a supporting role in providing feedback to students on the relative weight of different problems and on their performance in the class over time.

I wish to propose that these changes are not entirely unrelated to one another, but rather that the new resources provided may in some sense *enable* the elimination of the 100-point scale. That is to say: by spending more time on each chapter, the teacher may be making it possible for some students to learn the material more thoroughly; by allowing for extended time, the teacher may be making it possible for some students to earn higher grades; by assigning students to work on take-home problems, the teacher shifts emphasis away from rapid and accurate recall of facts and procedures, and allows students to consult with textbooks, notes, and the teacher himself. The gain students derive from these additional resources may be sufficient to compensate for the cost of using a nonstandard point scale.

Admittedly this is speculative. Without looking at student work or at grades, we have no way of knowing whether the introduction of these novel features had any effect on student achievement. On the other hand, let us approach the problem from the other side. *Suppose* that a significant proportion of students were having difficulty completing their exams in the time allotted. Certainly this would lead to some negative feedback on the teacher. What other kind of adaptation to this feedback might one expect, other than the creation of a mechanism by which students requesting additional time?

One possibility is that a teacher might reduce the number of questions on each exam. And indeed this does appear to have happened: a comparison of the Chapter 4, Chapter 6, and Chapter 7-8 exams in Years 1 and 2 (note that there was no Chapter 5 exam in Year 1) shows that exams in the second year contain nearly 1/3 fewer questions. A second possibility, compatible with the first, is that a teacher would change the *content* of the questions on the exams, to make them easier to complete in the limited time. This could happen in several ways: one or more difficult questions might disappear, possibly replaced by easier ones; or a question could survive but with some scaffolding, such as an explicit hint; or a multiple-part problem could reappear with fewer parts to it.

In fact a closer examination of the changes from Year 1 to Year 2 shows that all of these occurred (Fig. 5.18). In total, 19 of the 42 items on the three Y1 exams disappear entirely (with only 4 new items appearing on Y2 exams to replace them); 10 survive unchanged; and 12 survive with some modifications. This latter group was discussed in the previous section, in which it was shown that, in general, the trend was towards increased scaffolding and more explanatory text.

Assessment	# items (Y1)	# items (Y2)	Changes
Chapter 04	12	8	3 items dropped 1 item moved to Y2 Ch. 2 exam 8 items survive (5 unchanged, 3 modified)
Chapter 06	15	11	8 items dropped 7 items survive (4 unchanged, 3 modified) 1 new item 3 items moved from Y1 Ch 11 exam
Chapter 07-08	15	10	8 items dropped 7 items survive (1 unchanged, 6 modified) 3 new items

Figure 5.18. Changes from Year 2 to Year 3

Two measures that are relevant here are the number and density of the mathematical dispositions that are present (in the sense discussed in the previous chapter) in the assessment items. It will be recalled that the data shows a sharp decline in both of these measures over the course of Year 1 (Fig. 5.12), and that the reduction in length of the assessments over the course of that year can be largely accounted for by the elimination of items that attend to one or more of the mathematical dispositions. In light of the foregoing discussion, it is appropriate to see whether this trend continued in Year 2. Fig. 5.19 contains the details.

A comparison of Fig. 5.12 to Fig. 5.19 paints a somewhat mixed picture (Fig. 5.20). There is a consistent decrease in both the number of items and the number of dispositions on 5 of the 6 exams that are directly comparable between the two years. For the most part it appears that the reduction in the number of items is accounted for by reducing the number of items that contain evidence of the dispositions. The Chapter 2 exam is a notable exception in that the number of items drops only slightly but the number of dispositions *increases* slightly.

Assessment	# of items[a]	Dispositions[b]	Disposition density[c]
Chapter 1 Exam	14 (17)	8 (10)	0.59
Chapter 2 Exam	13 (17)	9 (16)	0.94
Chapter 3 Exam	14 (14)	4 (8)	0.57
Chapter 4 Exam	8 (8)	4 (8)	1.00
Midterm Exam	10 (10)	6 (17)	1.70
Chapter 5 Exam	7 (11)	10 (11)	1.00
Chapter 6 Exam	11 (11)	3 (5)	0.45
Chapter 7-8 Exam	10 (10)	5 (8)	0.80
Final	18 (18)	4 (12)	0.67

Notes. [a] The first number reported is the number of items as designated on the exam itself; the second number is the number of items as entered into the corpus database. [b] The first number reported is the number of items that were coded for *at least* one disposition; the second number is the total number of dispositions coded for in the assessment. [c] The disposition density is the ratio of the (total) number of dispositions on the assessment to the number of items as entered into the database.

Figure 5.19. The dispositions in Year 2.

Based on all of the foregoing, a clear picture emerges of teaching in which, over time, students' accountability decreases (as measured both by the number of items on assessments, and the number of assessment items that contain evidence of one or more dispositions), while simultaneously the amount of classroom time devoted to covering each chapter increases (at the expense of coverage of the curriculum), and students gain the option to take additional time to complete exams.

Assessment	Items		Items with Dispositions	
	Y1, Y2	chg	Y1, Y2	chg
Chapter 1	22, 17	-5	9, 8	-1
Chapter 2	18, 17	-1	7, 9	+2
Chapter 3	17, 14	-3	7, 4	-3
Chapter 4	12, 8	-4	7, 4	-3
Chapter 6	15, 11	-4	7, 3	-4
Chapter 7-8	15, 10	-5	5, 5	0

Fig. 5.20. A comparison of Year 1 to Year 2.
Note: Y2 assessments with take-home components are omitted, as are assessments that do not have analogues in both years.

But this is only part of the story. So far very little has been said about the content of the take-home Midterm and Final Exam. As we will see below, these take-home assessments are remarkable for the prominent role they reserve for problems that contain the mathematical dispositions. The analysis of those assessments suggests, contrary to the above, that the teacher's expectations of his students *vis-à-vis* the mathematical dispositions are not *lessening*, but rather are *migrating* from the in-school exams to the take-home exams. I now turn to a more full description and discussion of this phenomenon.

The take-home Midterm and Final (Year 2)

The Midterm exam in Year 2 consisted of 10 problems. Of those problems, exactly half were coded after one or more of the mathematical dispositions; in total, 16 dispositions were assigned to those five problems. Fig. 5.21 shows the details.

Question #	Dispositions present
1	Generalize, Abstraction
2	Utility
3	Existence, Converse, Generalize, Specialize
4	n/a
5	Formalism, FindCon4Hyp
6	n/a
7	Utility
8	n/a
9	n/a
10	FindHyp4Con, FindCon4Hyp, Existence, Converse, Generalize, Specialize, Complexity

Figure 5.21. Coding the problems in Y2's Midterm Exam for the dispositions

It is immediately obvious that the density of dispositions in the Midterm exam was significantly higher than on any of the in-class exams in Y1 or Y2. Across the whole exam, the density of dispositions in the exam was 1.7; if we restrict our attention only to the six problems that had at least one disposition, there were an average of 2.83 dispositions per problem. Even if we disregard the clear outlier (Problem 10, with 7 dispositions) the remaining five problems have an average of 2.0 dispositions per problem. It is also worth noting the alternating pattern evident in Fig. 5.21: it appears as though the teacher has attempted to balance the disposition-rich problems with more "normal" ones.

At this point is seems appropriate to look into the content of some of the questions. Figure 5.22 contains two of the disposition-rich problems from the exam. The first of these, Question 1 (Fig. 5.22a), has several distinctive features. Most immediately obvious is the *quantity of text* present in the problem; on the original exam, Question 1 filled an entire page. Second, there is an introductory paragraph in which new mathematical content is presented: it seems from the text that students are not expected to have previously encountered the Cantor set or Sierpinski gasket, or indeed the notion of a fractal at all. In this respect the problem resembles the "simplices" problem described in the previous chapter (Fig. 4.15). Note, however, how different this problem is from that prior example are in terms of what they call on students to do. Previously it was seen that in order to solve the "simplices" problem a student needed to be able to draw a few pictures, count to 6, and make some predictions based on pattern recognition (some of which would only have been plausible for a student who had some prior exposure to Pascal's Triangle). In contrast, the "fractals" problem requires students to:

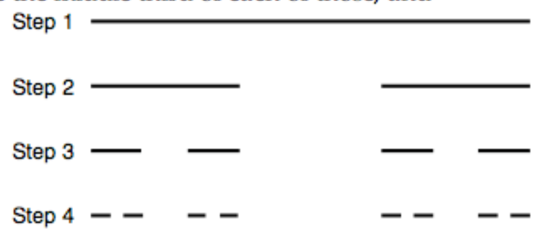
a. recognize that each of the small segments in any given stage in the construction of the Cantor set is one-third the length of the segments in the prior stage
b. recognize also that there are twice as many segments in each successive stage

c. integrate those two facts into the recognition that the total length of each stage is 2/3 that of the previous stage
d. represent this symbolically using exponential notation, as $(2/3)^n$
e. extrapolate forwards until the value is less than 0.001
f. and adapt all of the preceding to the analogous case of the Sierpinski gasket.

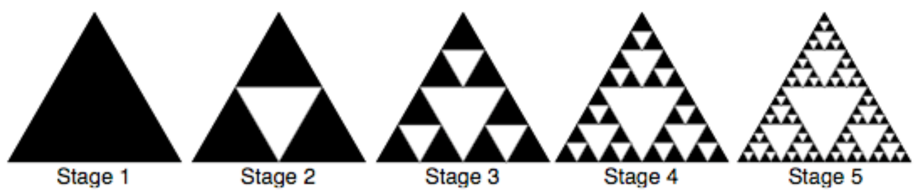
It should be pointed out that the various parts of the “fractals” problem are not really *open-ended*, in that the solutions are fully specified by the data in the problem, and there is really only one correct solution for each part. Despite this the problem is rich in mathematical dispositions: it situates the examples of cases of an abstract concept (“fractal”), calls on students to generalize a pattern of numbers, and presents visually complex diagrams (the fifth stage in the construction of the Sierpinski gasket, for example, contains too much detail at too small a scale for a student to easily “read” it).

(a)

1. An interesting geometric figure called the *Cantor set* is constructed in the following fashion. Start with a line segment of length 1. Delete from it the middle third; this produces two disconnected segments, with a total combined length of 2/3. Then remove the middle third of each of those, and continue in this fashion forever. (See the figure at right). This is an example of a *fractal*, which means a figure in which each part of the figure contains a copy of the entire whole (i.e., if you zoom in on any part of the Cantor set, you find a copy of the entire Cantor set).



   (a) Find the total length of each of the second, third, and fourth stages in this pattern.
   (b) Complete the conjecture: the total length of the n^{th} stage in this pattern is __________.
   (c) At what stage is the total length of the Cantor set less than 0.001?
   (d) A similar fractal, called the *Sierpenski gasket*, is constructed as follows: start with a triangle of area 1, and delete the middle from it, forming three identical (smaller) triangles. Continue removing the middle from each triangle. (The process is illustrated below.) Write a conjecture about the area of the n^{th} figure in the pattern.



(b)

3. Consider the following (true') statement: "If a quadrilateral is a kite, then the angle bisectors of the quadrilateral all meet in a common point." (See the diagram below.)

   (a) State the converse of this statement.
   (b) State the contrapositive of this statement.
   (c) State the inverse of this statement.
   (d) Which, if any, of the above do you know must be true?
   (e) Give a counterexample to the following (false) statement: "For any quadrilateral, the angle bisectors all meet in a common point."

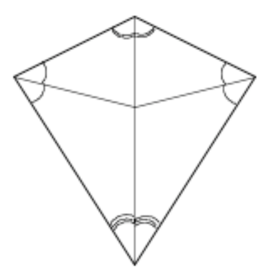


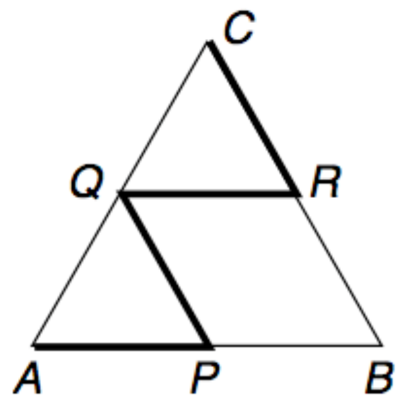
Figure 5.22. Two questions from the Y2 Midterm Exam.

The second disposition-rich problem (Fig. 5.22b) is quite different: it contains much less text, and deals with geometric objects (quadrilaterals, kites, angle bisectors) that a student would be expected to have seen before. Like the "fractals" problem, the "kite" problem begins by presenting the student with some new mathematics: namely the fact that the angle bisectors of a kite meet at a point. The problem explicitly identifies this as a true statement; at the bottom of the page (not reproduced in Fig. 5.22), a footnote on the word “true” comments to students “You can just take my word for it, although the proof is not that hard”, from which it may be inferred that students are not expected to have prior knowledge of this fact. The first three parts of this multi-part problem call for little more than an understanding of the words “converse”, “inverse”, and “contrapositive”. To answer the fourth part, a student could make use of the fact that any statement and its contrapositive are logically equivalent — something they would have been expected to learn when studying Chapter 2 of the textbook — and hence, even in the absence of a proof, the contrapositive must be true (since the original statement was sanctioned as true). Finally, the fifth part of the problem presents the student with a (false) generalization, and calls for the student to disprove it by showing that a counterexample exists — which a student might do by, e.g. drawing a rectangle with its angle bisectors, or any other non-rhombus parallelogram. For this last part, many different solutions could be judged equally correct.

The most interesting problem from the Y2 Midterm is, as might be expected, the outlier of the set — Problem 10, which alone was coded for 7 dispositions. This problem is reproduced in its entirety in Fig. 5.23 below. Once again the sheer quantity of the text is immediately noticeable, as is the way in which it addresses the reader directly (“...what you have to do...”, “If you give up one or more of these assumptions....”).

10. Two ants, Arthur and Angela, are walking on the triangle shown below. Arthur walks from *A* to *B* to *C*. Angela, on the other hand, follows the bold-faced zigzag pattern shown on the diagram (from *A* to *P* to *Q* to *R* to *C*).

Here's the question: Is one of these paths longer or shorter than the other, or are they the same length? Here's the tricky thing: *There is not enough information in the problem to answer this question definitively.* So what <u>you</u> have to do is make some extra assumptions in order to make this problem solvable.



Your answer to this question (which should be written in complete sentences, and might resemble a short essay) should include the following information: Under what assumptions can you be certain that the paths are the same length? If you give up one or more of these assumptions, how does it effect the conclusion? Can you describe a situation in which Arthur's walk would be longer? What about the reverse?

Figure 5.23. The tenth and final problem from the Y2 Midterm.

The "antwalk" problem in Fig. 5.23 is almost identical to a problem that was used in Herbst & Brach (2006, p. 84) as a probe for exploring what share of work students accept responsibility for when doing proofs. In that study it was shown that in any situation in which a proof is called for, students expect to be told explicitly what "givens" they are to use, and what they are to prove. Both the "givens" and the "prove" are normally fixed parameters of the problem, specified by the teacher or a teacher's surrogate (e.g. textbook); students' share of the labor is to create a chain of logical reasoning that leads from the former to the latter. In light of that prior research the problem in Fig. 5.23 appears exceptionally unusual: it openly subverts those normative expectations, flouting the under-specified nature of the problem ("Here's the tricky thing: *There is not enough information in the problem...*") and calling for the students to provide both a set of assumptions and a set of conclusions that follow from those assumptions.

Moreover it also asks students to *vary* those assumptions and explore how the conclusions change in response. It asks for students to try to find counterexamples, to consider the converse, to weaken hypotheses — in short it calls on students to perform essentially all of the generative moves that were enumerated in Chapter 2 as comprising part of mathematicians' categories of perception.

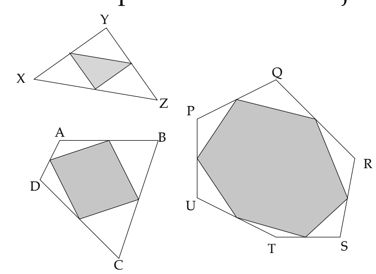
What kind of response could a teacher expect from such a task? And how might those responses influence future the decisions the teacher might make when writing future exams? Certainly, a question like the "antwalk" problem seems like a significant departure from the norm. If students' performance on the task were to fail to live up to the teacher's expectations, we might expect to find no further experiments of this sort. But in fact the opposite is true: when we turn to the Final Exam from Y2, we find that the abnormal features of this problem reappear in an even more magnified form.

The take-home portion of the Y2 Final consisted of a single problem, the instructions for which were covered three full pages on the original exam. These instructions are reproduced in their entirety below (Fig. 5.24):

Honors Geometry Final Exam
Part 1 (Take-Home Portion)

This portion of the Final Exam is worth half of the entire final (so, 10% of the semester grade).
Information on Part 2 of the Final is at the end of this document.

Begin with any polygon. If you join the midpoint of each side to the midpoint of the two adjacent sides, the result will be another polygon, called the *dual* of the original polygon.



For example, the figures at right show the dual of triangle *XYZ*, the dual of quadrilateral *ABCD*, and the dual of hexagon *PQRSTU*. (Each dual is shown shaded in gray.)

Your task is to conduct original mathematical research in the subject of duals of polygons. What are their properties? What are the relationships between the original figure (called the "predual") and its dual?

"Research" does not mean "look it up in a book or on the Internet" (I doubt you'll be able to find anything anyway). Instead, "research" means:

1. To pose interesting questions
2. To investigate your questions, by drawing lots of pictures, making measurements, and looking for patterns
3. To formulate conjectures, and present evidence to support your conjectures
4. When possible, to answer your questions, and *prove* (or *disprove*) your conjectures (at which point they stop being conjectures, and become *theorems*).

Please note that you don't have to answer *all* of the questions you raise; nor do you have to *prove* all of the conjectures you propose. Some questions might be just too hard. I want to know about the questions you *couldn't* answer, as well as the ones you could. (Of course, I expect that you'll be able to answer *some* of your questions – I'll be very disappointed if you can't produce *any* theorems.)

To get you started, and to illustrate what I mean, here are some examples of the kinds of questions you might be able to answer.

Question 1. What can you say about the dual of a triangle?
Question 2. What can you say about the dual of a quadrilateral?
Question 3. Under what circumstances will the dual of a quadrilateral be a *square*?

Question 1 is easy to answer (and I expect you to answer it in the course of your research).

Question 2 is not so easy to answer. Part of the answer is:
Theorem. For any quadrilateral, the dual will be a *parallelogram.*
Proof. You can find this on page 364 (Example 2) in the textbook.
There's much more that could be asked about duals of quadrilaterals, however.

Figure 5.24 (continues on next 2 pages)

Question 3 is very hard to answer. Here's a conjecture:

Conjecture. The dual of a quadrilateral will be a square if, and only if, the predual is also a square.

Is this conjecture true? Maybe yes and maybe no. You ought to investigate it and try to find out.

Besides these three questions, here are some words and phrases that might spark some ideas in your mind. Not everything on this list is necessarily going to be useful to everyone, but you might get some interesting questions by pondering this list.

Lengths • angles • bisectors • concurrency • parallel • perpendicular • congruence • similarity • triangles • isosceles • equilateral • incenter circumcenter • median • orthocenter • circumscribed circle • inscribed circle • midsegment • special quadrilaterals • convex • nonconvex • equiangular polygons • equilateral polygons • area • perimeter • symmetry • transformations • proportions • regular polygons • circles • chords • secants • tangents • polyhedra • Platonic solids • duals of duals

What do you turn in?

This assignment is different from what you are probably accustomed to, because you have to *generate the questions* -- not just answer them.

The process of doing research is often messy – diagrams get scribbled in the margins of notebooks, on the backs of envelopes, etc. And the ideas can come at odd times – while you're in the shower, sitting in history class, watching TV. Sometimes you even get ideas in your sleep!

I don't expect (or want) you to turn in every single scrap of paper that you wrote on in the course of doing this research. Instead, I want you to take your results, and write up a *research report*. Your research report should tell me what you did in an organized and coherent fashion. You should include descriptions of your investigations, questions you asked (both the ones you were, and were not, able to answer), conjectures you considered (even if you subsequently found out they were wrong), theorems you proved – anything.

I assume that your research report will be typed, illustrated with lots of diagrams, and proofread for grammar and spelling. I anticipate most of you will produce reports that will be in the neighborhood of 5 pages long (double-spaced) – I can't imagine doing this well in fewer than 2 full pages, but some of you *might* end up writing 10-15 pages or more.

I hope to take the best research reports and combine them to produce something that we can submit for publication in journals of teaching mathematics. With hard work and some imagination, you may end with your work in print in a nationwide publication!

Honor Statement

The first page of your research report must contain the following honor statement, written in your own handwriting and signed:

> HONOR STATEMENT: I attest that I completed this exam without any assistance from any other student, teacher, parent, or tutor.
>
> ______________________________ (signature)

Figure 5.24 (continued)

Research reports that do not contain the honor statement *will not be accepted*.

Due Date:
This portion of the final exam is due **Monday, June 7th**, no later than **12:00 noon**. Papers turned in late will face a 0.4% reduction in your grade for every hour it is late. So, for example, papers turned in Tuesday at 1:00 PM (25 hours late) will lose 10%. (If you email it to me Monday evening, I will look at the time the paper was sent, not the time it was received.)

I will be glad to receive your work early. Turn your work in on or before Friday, June 4th at noon, and you will receive a 5% "early bird" bonus on your work.

Lastly, the other portion of your final exam – the in-school portion – will be held at the normal time during exam week.

Figure 5.24. The take-home portion of the Y2 Final Exam

There is so much to say about this open-ended problem that it is difficult to know where to begin. First, note that the text repeatedly stresses the novelty of the task, and the originality of the work that students are expected to do: words and phrases such as "original mathematical research" are scattered throughout it. The instructions make explicit that students are expected to pose their own questions, to investigate them, to make conjectures, and (if possible) to either prove or refute those conjectures. They also make clear that students are expected to *organize their work* in a coherent report — that is, not to produce a miscellany of disconnected findings but rather to build a rudimentary *theory* of duals. And the instructions openly acknowledge that it can be valuable to ask a good question, even if no answer is forthcoming. (In this connection it is worth noting that the twin dispositions of *Surprise* and *Confirmation* both depend on the possibility that a problem may exist without a clear resolution for an extended period of time, something that I have shown in Chapter 3 is uncommon in teachers' practice.) All of this is part of the specification of the *goals* and *framing* of the problem.

At the same time, the openness of the work students are asked to do is bounded by the limiting of the conceptual domain: this is to be a research project about duals, and work that is not connected to that general theme in some way has no place in it. For students who do not know where to begin, three “starter questions” are provided, and for those who or are stumped for what to do next, a long “list of words and phrases that might spark some ideas” is provided; however students are not obliged to make use of any or all of these words and phrases. These (along with the definition, the figure, reference to a particular page in the textbook) are part of the *resources* provided by the teacher.

The problem also specifies the sorts of *operations* students are expected to deploy in their work on the problem. All of the “generative moves” discussed previously are explicitly present in these instructions. Students are asked to find conclusions that follow from certain hypotheses, to find conditions under which certain conclusions follow, to inquire whether the converse of a statement is true, to generalize, and to specialize. They are invited to look at concepts from the “list of words and phrases” and to inquire after connections between them. For example, a student might read through that list and pose any of the following questions:

a. Is the *incenter* of a triangle the *circumcenter* of its dual? If not always, when does this happen?
b. Is the *centroid* of a triangle also the centroid of its dual?
c. Do nonconvex polygons have duals?
d. Can a dual ever be nonconvex?
e. If a polygon is symmetric, is its dual also symmetric? What about the converse?

f. How is the area of a triangle / quadrilateral / general polygon related to the area of its dual? What about the perimeter?

g. Do circles have duals? What would those be?

h. Do solids have duals? What would those be?

Some of these questions can be given clear and unambiguous answers. Others are more difficult, yielding only in special cases. Some call for students to make new definitions. Some of the questions are relatively easy to answer, while others are quite subtle and difficult. Students are thus free to make this research project as challenging as they want it to be.

The existence of the take-home portion of the Year 2 Midterm and Final Exam thus counters the impression that was produced by the earlier analysis of the Year 2 in-school assessments. While the in-school assessments show evidence of a *decrease* in the teacher's expectations of his students (as measured by number of items per assessment and number of dispositions present in each exam), the out-of-school assessments show that those expectations are not disappearing, but are rather finding their expression in another form. In summary, what we have here is the beginning of a *bifurcation* of the corpus into two distinct subcorpora: a set of in-class assessments that more-or-less resembles a "normal" assessment, and a parallel set of out-of-school assessments that provide students with disposition-rich, open-ended problem contexts. As we shall see below, in the third year of the corpus this trend continued, and the gap between the two subcorpora of assessments grew even wider.

Achieving a new equilibrium: Year 3

The assessments from the third year show a continuation of the trends that began to emerge in Year 2. I have already discussed the manner in which the "I need extra time checkbox", introduced in Y2, was modified at the beginning of Y3 by the introduction of both a "purchase price", amounting to 5% of the total exam grade, and an explicit retake policy. It was also mentioned previously that, beginning with the Midterm exam in Y3, every unit assessment consisted of both a take-home problem set and an in-class portion (designated as a "Quiz" on the documents). The full details for the Y3 assessments can be seen in Fig. 5.25.

Assessment	**Date**	**# of items**[a]	**Points Possible**
Unit 1 Exam	9/27/2004	10 (14)	88[b]
Unit 2 Exam	11/3/2004[c]	15 (17)	93[b]
Unit 3 Exam	12/2/2004	14 (15)	100
Midterm	12/20/2004	9 (9)	100
Unit 4 Take-Home Problems	1/24/2005	3 (4)	60[d]
Unit 4 In-School Quiz	1/24/2005	4 (4)	n/a[d]
Unit 5 Take-Home Problems	3/4/2005	5 (5)	60
Unit 5 In-School Quiz	3/4/2005	5 (7)	30
Final Research Project	6/7/2005	1 (1)	40
Final Exam (In-School)	6/7/2005	20 (20)	40

Notes. [a] The first number reported is the number of items as designated on the exam itself; the second number is the number of items as entered into the corpus database. [b]Includes one question worth 8 points "extra credit". [c]Exam was spread over two consecutive class meetings (11/3 and 11/5). [d]Point values not printed on exam; reconstructed from graded student work or grading guides, where extant.

Figure 5.25. Assessments in Year 3.

One remarkable characteristic of the Y3 assessment data is the continuation of a trend noticed in Y2: the spacing between assessments continues to grow, seemingly at the expense of coverage of the latter chapters. We note that the spacing between the first four exams was on the order of four weeks, comparable to Y2; after the Midterm exam, however, six full weeks elapse between the Unit 4 and Unit 5 assessments, and following

Unit 5, formal unit assessments appear to disappear *entirely* — notice the three-month gap between the Unit 5 assessment and the Final Exam.

The absence of formal assessments for the subsequent chapters of the textbook does not necessarily mean that the content of those chapters was not covered. On the contrary, the Y3 Final Exam includes a number of questions on precisely that content, and a Review Sheet distributed prior to the Final makes it clear that students were accountable for knowing that material. Still, the absence of formal assessments appears in context as the ultimate extension of the trend that began in Y2.

It is also noteworthy that, as in both previous years, the mid-year break appears to have been a time of transition. Previously we saw (in Y1) that from the midterm exam onward, assessments began to conform to a 100-point scale, and became shorter in length and less focused on the mathematical dispositions; we also saw (in Y2) the disappearance of the 100-point scale and the first use of take-home problem sets. Now we see that following the Y3 Midterm, both the Unit 4 and Unit 5 assessments were formally divided in two parts, one designated "Take-Home Problems" and the other designated a "Quiz". The bifurcation described at the end of the previous section is now institutionalized: rather than one corpus of assessments, we can now recognize the existence of two distinct subcorpora.

We find corroboration for this if we look at the two subcorpora through the lens of the mathematical dispositions (Fig. 5.26). A comparison of Fig. 5.26 with Fig. 5.20 (above) shows that the length of the first three exams remained relatively stable from Year 2 to Year 3, as did the number of items containing dispositions on each of those assessments. Beginning with the midterm, however, there is a clear change: the number and density of dispositions on the in-school assessments drops sharply, while the number and density of dispositions on the take-home assessments climbs dramatically.

Assessment	# of items[a]	Dispositions[b]	Disposition density[c]
Unit 1 Exam	10 (14)	6 (7)	0.50
Unit 2 Exam	15 (17)	9 (16)	0.94
Unit 3 Exam	14 (15)	5 (12)	0.80
Midterm	9 (9)	6 (18)	2.00
Unit 4 Take-Home Problems	3 (4)	2 (6)	1.50
Unit 4 In-School Quiz	4 (4)	2 (3)	0.75
Unit 5 Take-Home Problems	5 (5)	5 (8)	1.60
Unit 5 In-School Quiz	5 (7)	5 (5)	0.71
Final Research Project	1 (1)	1 (10)	10.00
Final Exam In-School	20 (20)	2 (2)	0.10

Notes. [a] The first number reported is the number of items as designated on the exam itself; the second number is the number of items as entered into the corpus database. [b] The first number reported is the number of items that were coded for *at least* one disposition; the second number is the total number of dispositions coded for in the assessment. [c] The disposition density is the ratio of the (total) number of dispositions on the assessment to the number of items as entered into the database.

Figure 5.26. The role of the dispositions in Year 3.

Earlier I argued that the gradual decline over time in the role allocated for the mathematical dispositions on in-school assessments can be taken as an adaptation on the part of the teacher, who (I theorize) is responding to some form of feedback from his milieu. I noted that the midpoint of the year is a natural time to expect for such an adaptation: the administrative need to record students' grades on report cards, and the consequent communication with, parents, administrators, counselors, etc., all potentially serve as sources of feedback on the teacher. At this point in the narrative it appears that the teacher has found a way to respond to that feedback and still preserve a role for the mathematical dispositions. He has done this by creating two parallel forms of assessment, one of which (the in-school quizzes) conforms more closely to what one might expect of a “normal” assessment, while the other (the take-home problems) carves out a niche for more mathematically authentic investigation.

The difference between the two forms of assessment can best be seen through a direct comparison of the questions on the two Unit 4 assessments (Fig. 5.27). The four questions in the left-hand column of Fig. 5.27 constituted the Unit 4 in-school quiz; the three questions in the right-hand column were the Unit 4 take-home problem set. First, it should be noted that the problems from the in-school quiz are not trivial, nor are the mathematical dispositions entirely absent. On the contrary, 3 out of 4 problems call explicitly for a proof; two of the problems (#1 and #2) call for the student to determine what conclusion follows as a consequence of the provided “givens”; and one of the questions (#2) is one of the exemplars of the *Formalism* code (see Figure 4.20b).

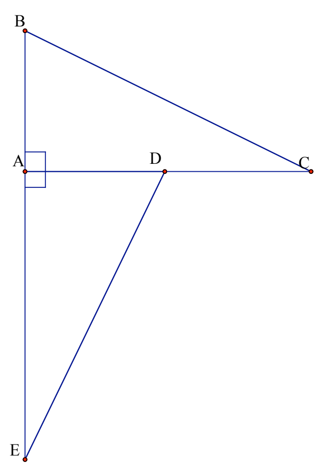
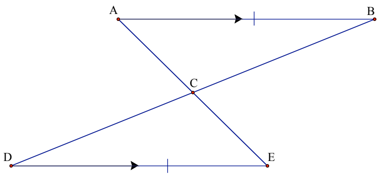
1. In the diagram on the right, $\overline{AB} \cong \overline{AD}$ and $\overline{ED} \cong \overline{BC}$. Write a true triangle congruence statement, and state what rule you used. 	1. Line *l* is defined by the equation $y = 4x - 2$. a. Find the equation of the line that is parallel to line l and passes through the point (4,9). b. Find the equation of the line that is perpendicular to line l and passes through (0,3). c. Find the coordinates of the point where **the original line *l*** crosses the line you found in (b). d. Find the distance between the point (0,3) and your answer from (c).
2. As you know, when writing a triangle congruence statement, the order of the letters matters. Suppose for three points A, B, and C it happens to be true that $\triangle ABC \cong \triangle CAB$. What can you conclude about $\triangle ABC$ in this case? Be as detailed as possible.	2.a. Prove that if P is a point on the perpendicular bisector of $\overline{QR}$, then P is equidistant from Q and R (this means that PQ = PR). b. Conversely, prove that if then *P* is equidistant from *Q* and *R*, then *P* is on the perpendicular bisector of $\overline{QR}$.

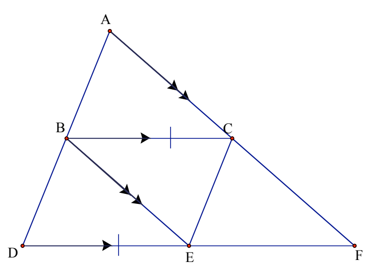
Figure 5.27 (continues on next page).

3. In the diagram at right, $\overline{AB} \parallel \overline{DE}$ and $\overline{AB} \cong \overline{DE}$. Write a true congruence statement, and give a proof.



4.

5. In the diagram at right, you are given that $\overline{BC} \cong \overline{DE}$, $\overline{BC} \parallel \overline{DE}$, and $\overline{BE} \parallel \overline{AC}$. Using *only this information* (don't make any other assumptions), prove $\triangle ABC \cong \triangle BDE$.



3. There are a number of properties of quadrilaterals that we have discovered almost accidentally in the course of the last few weeks. The purpose of Problem #3 (which is really a set of problems) is to use what you know about congruent triangles to organize and prove those properties in some logical fashion.

Below is a list of ten conjectures about quadrilaterals. Some of them are true, and some are false. You are to do the following:

(a) Figure out which are which
(b) Provide clear counterexamples to the false ones
(c) Provide proofs for the true ones

Note that you can rearrange the conjectures in any order you want; once you've proven one, you can use it to help you prove the subsequent ones. So part of the task here is to figure out what would be a "good" order to prove these in.

The conjectures:

- In a parallelogram, the two diagonals bisect each other.
- If a quadrilateral's two diagonals are congruent, then it is a parallelogram.
- In a parallelogram, there are two pairs of congruent opposite sides.
- If a quadrilateral has two pairs of congruent opposite angles, then it is a parallelogram.
- In a parallelogram, there are two pairs of congruent opposite angles.
- If a quadrilateral's two diagonals bisect each other, then it is a parallelogram.
- In a parallelogram, the two diagonals are also angle bisectors.
- If a quadrilateral has two pairs of congruent opposite sides, then it is a parallelogram.
- In a parallelogram, the two diagonals are congruent.
- If a quadrilateral's two diagonals are also both angle bisectors, then it is a parallelogram.

You will probably need the following definitions:

Definition. A *quadrilateral* is a set of four line segments that intersect only at their endpoints, and in which each endpoint is shared by exactly two segments.

Definition. A *parallelogram* is a quadrilateral with two pairs of parallel sides.

(a)	(b)

Figure 5.27. The (a) in-school and (b) take-home assessments for Y3, Unit 4.

Nor are the three take-home problems all paradigmatic examples of contexts for open-ended, mathematically authentic problem-solving. On the contrary the first question is a fairly pedestrian (but still challenging) multi-part problem, the kind that might just as well be expected to appear on an in-class assessment — and indeed, the exact same question *had* appeared in both Y1 and Y2 on the Chapter 3 exams. On the other hand, the second and third problems on this assessment are quite unusual, and I turn now to an analysis of their content.

The first thing that should be noted about Problem 2 is that the mathematical propositions the student is asked to prove have a normative role in the curriculum; in the textbook that students used for the course in question (Larson, Boswell, & Stiff, 2001), they are referred to as the "Perpendicular Bisector Theorem" and its converse, which appear in the first page of Section 5.1 of the textbook[37]. It is thus particularly interesting that the teacher includes the proof here on the assessment for the *previous* chapter: it indicates that the teacher is using the take-home problem set not only to assess what students have *already* learned, but also to introduce new material that will be relevant for future work (cf. my discussion in Chapter 4 of the *prospective* function of assessment).

Another noteworthy feature of problem #2 on the take-home problem set is that no diagram is provided to students. This is in sharp contrast to the three proof problems on the in-class quiz, all of which included complete diagrams. Elsewhere (Weiss & Herbst, 2007) I have shown that in the secondary Geometry course, proof problems are customarily presented using a "diagrammatic register", that theorems are customarily introduced using a blend of a the "conceptual" and "generic" registers, and that teachers

---

[37] The rationale for including this theorem at the beginning of Chapter 5 ("Properties of Triangles") appears to be that it is instrumental in proving that the three angle bisectors of any triangle are concurrent. This fact appears (and is labeled a "theorem") in section 5.2 of the text (p. 273); interestingly, however, the proof is buried deep in an appendix at the end of the textbook (p. 835). The implication seems to be that the textbook's authors feel obliged to ensure that a proof is provided, but do not expect students to see it, or for teachers to hold students accountable for understanding it. In this connection, it is particularly interesting that one of the questions on the Y3 Chapter 5 Take-Home problem set (the "two parabolas" problem, not discussed in this dissertation) seems to have been designed precisely to test whether students understand the structure of that proof, and are capable of applying the reasoning to another context.

normally shield students from the responsibility to perform translations between those registers. Problem #2 defies that custom: it makes use exclusively of conceptual (verbal) tokens (e.g. “perpendicular bisector”, “equidistant”) and generic signs ($P, Q, R, \overline{QR}$); it is the student who must produce a diagram that represents this abstract situation. Moreover the explicit separation of the biconditional proposition into two distinct conditional statements, and the explicit labeling of the second part as a converse, calls attention to the mathematical disposition of considering the converse of a proposition.

By far the most provocative question on the take-home problem set is Problem 3. Many of the attributes that have previously been noted about take-home problem sets in the Y2 Midterm and Final are visible here, as well:

Like the Antwalk problem on the Y2 Midterm, and the Duals research project on the Y2 Final, Problem 3 is extremely text-heavy. Notice the framing of the problem: the first paragraph serves not only to introduce the theme (properties of quadrilaterals) but also to establish a shared memory of prior mathematical work (“we have discovered almost accidentally in the course of the last few weeks”) and to motivate the task. The goal is articulated as “to organize and prove those properties in some logical fashion”; that is, the student is here charged with responsibility not only for proving theorems, but for the *organization of those theorems* into a theory. This is reinforced by the explicit comments that “you can rearrange the conjectures” and “part of the task here is to figure out what would be a ‘good’ order to prove these in”. This is the mathematical disposition of *Theory Building*, embodied in a task for students.

Notice as well the prominent role of *personal pronouns* in this problem: “we” (the class) have discovered these properties, “you” (the student) are going to organize and prove (or disprove) them. In fact "you" appears six times in the text of this problem. This is an instance of the more general phenomenon noted above, in my discussion of how personal pronouns are one of the key resources the implied author employs to create coherence in the corpus. Here we see that the pronouns establish the roles and demarcate the zones of accountability for the author and reader of the text: "you" are responsible for deciding what is true, providing proofs and counterexamples, and organizing the theory. On the other hand "you" are not (yet) responsible for knowing or producing definitions for the words "quadrilateral" or "parallelogram" (they are provided because "you will probably need" them).

Other noteworthy features of this problem include the fact that it avoids entirely both the diagrammatic register and the generic register, instead remaining entirely within the conceptual register (Weiss & Herbst, 2007). This includes the use of conditional statements in *unparsed* form (cf. p. 47). Students are thus responsible for extracting the hypothesis and conclusion from the statements, for instantiating the abstract properties into generic representatives, and for producing a diagram that illustrates the properties of the generic objects — all of this prior to the actual determination of which properties are true and which are false. And neither the parsing, nor the determination of which propositions are true, is trivial: for each proposition in the list, the converse also appears in the list (but the pairs are camouflaged both by the sort order and by variations in grammar); moreover among these "converse pairs" are examples in which both are true, as well as examples in which both are false, and one pair in which only one is true[38].

---

[38] Specifically, the seventh tenth property is false, but its converse, the tenth property, is true: if each diagonals of a quadrilateral bisects both of the angles at its endpoints, then the quadrilateral must be a rhombus, and hence a parallelogram. It should be noted that the wording

A final feature of this problem is that, like Problem 2 on the same problem set, it holds the student accountable for proving and organizing content that, from the point of view of the textbook's own chronology, is premature: the various families of quadrilaterals are not even supposed to be defined until Chapter 6, which is when the properties of special quadrilaterals are normally lodged. This is thus another illustration of the prospective, rather than retrospective, function of assessments. In fact comparing the content of the textbook's Chapter 6 to the propositions listed in Problem 3, we can see that there is one true proposition among the latter that is not included in the textbook (namely, the tenth property in the list).

All of this provides us with another point of view on the phenomenon of diminishing coverage, noted above in connection with Year 3. While it is true that the Y3 corpus does not include an assessment devoted explicitly to Chapter 6, the properties of quadrilaterals, this absence does not mean that the content was not covered. On the contrary, the appearance of this question two chapters "early" partly explains *why* no assessment for Chapter 6 exists: as the teacher's adherence to the curriculum as defined by the textbook's table of content gradually erodes, assessments corresponding to individual chapters become impractical (in this connection note also the teacher's change in nomenclature, from "Chapter N" in Y1 and Y2 to "Unit N" in Y3).

Consider also the difficulty of assigning a grade to students' solutions to Problem 3. In particular, the fact that the problem makes students accountable for creating a logical organization of the theory (on the grounds that any proposition, once proved, might be useful in proving subsequent propositions) implies that the teacher should take

---

of the tenth property is sufficiently imprecise that a serious argument could be put forth that it is false. Even the translation between registers is nontrivial in this problem.

that organization into account when deciding how many points to assign to a solution. But on what objective basis can a teacher decide that one such organization is better than another — and how can a numerical value be attached to that difference? Certainly there is much in Problem 3 that can be assessed objectively — the truth or falsity of each proposition, the suitability of any given proof or counterexample — but the problem calls for more than that. An organization of the theory cannot be evaluated as purely right or wrong, but rather as elegant or inelegant, efficient or inefficient, aesthetic or clumsy — categories of value that do not translate well into numerical scores. It is difficult to imagine any grading rubric for this problem that does not have a subjective component to it.

The foregoing discussion illustrates the claim that I have made above — that the bifurcation of the corpus into two distinct assessment subcorpora (take-home and in-class) serves to create a mechanism by which the teacher can continue to hold students accountable for the mathematical dispositions — and even increase those expectations of students — while also preserving a role for more normal assessments. The four problems on the in-class problem set are not trivial, but they are relatively *short* (both in terms of the amount of text and time it takes to specify them, and the amount of text and time it takes to solve them), have *well-defined, unique solutions* (in contrast to the open-ended nature of the take-home problems), can be graded using more or less objective criteria, play a purely retrospective function, and are to be done under the teacher's direct supervision. The problems on the take-home set, in contrast, call for the student to take part in a wide range of mathematical activities (conjecturing, proving, disproving, translating between registers, theory building), include a mix of open-ended and fully-specified problems, and allow for some degree of creativity on the student's part — but call for a more extended timeframe, and require a measure of trust on the teacher's part.

If the above analysis is correct — if the bifurcation of the corpus serves to carve out to distinct spheres of assessment, one more "normal" and the other more "authentic" — then one might expect that the "extra time" checkbox would no longer be necessary on the in-school quizzes. That is, if the function of the take-home assessments is in part to create a way to hold students accountable for mathematical work that is more demanding of time and other resources than an in-school assessment allows, the emergence of those assessments would seem to obviate the necessity for allowing students additional time on their *in-school* assessments, which by design were intended to be not as time-consuming. And in fact this is precisely what happens in the corpus: neither the Unit 4 nor Unit 5 assessments in Y3 (the assessment after the midterm, the very two that were formally divided into two components) includes the by now familiar "extra time" checkbox. It no longer appears, presumably because the teacher no longer considered it necessary.

Discussion

Looking back over the entire 3-year corpus, we see a continuous process of teacher learning, i.e. adaptation to some kind of feedback. Although we have no direct evidence of the nature of this feedback, I have speculated above on some of its sources, and I review here the way in which such feedback might have led to some of the changes described above.

- *Students' grades* constitute a measure not only of their own work; indirectly they serve as a commentary on the teacher's competence. If grades are too low, a teacher risks appearing ineffective, overly strict, or unreasonable in his expectations; if grades are too high, a teacher risks the perception of having low standards. This is not to suggest that a teacher is primarily concerned with what others think; the above applies to a teacher's self-evaluation, as well. If a teacher sees that none of the students in his class were able to earn above 70% on a test, it is entirely reasonable to conclude "I must have made the test too hard", and to consciously write an easier test next time. Alternatively, he might preserve the difficulty of the exams, but allow students who fail an exam to retake it. On the other hand, a teacher might conclude that he needs to spend more time teaching the students the content, so that they will be better prepared for the next test.

- *Students' ability to complete an assessment in the allotted time* provides a second, complementary source of feedback to the teacher. If a teacher comes to the realization that the tests he is writing cannot easily be completed within the customary time (which is not *quite* the same as saying that they are too hard), a teacher might do any of the following:
    - plan to devote two class periods for each future exam
    - make the exams shorter
    - allow students to choose whether they need additional time
    - repackage some of the more time-consuming problems as take-home problem sets
- But all of the above options entail further complications. Spending more time teaching students the content of each chapter means sacrificing the latter chapters. Devoting two class periods to each exam cuts into the amount of time available

> for teaching, further compromising the coverage of the curriculum as defined by the textbook. Allowing students to choose whether they need additional time or retake a test creates logistical challenges around grading and returning exams to students, marking report cards, etc. And migrating time-consuming problems into take-home assessments removes the student from the teacher's observation, opening up the possibility of cheating.

The assessments corpus provides indirect evidence of the above cycles of feedback and adaptation. Some of the structural features of the assessments — the honor statements attached to take-home problems, the "purchase price" for extra time and retakes — can be understood as second-order corrections, adaptations to the feedback produced in response to adaptations to feedback.

The trends I have noted in the number and density of mathematical dispositions across the corpus provide an independent corroboration of this evidence. Initially (up through the middle of Y2) both measures gradually but consistently decrease. Following the introduction of the take-home midterm in Y2, we see the number and density of dispositions continue to decrease on *in-school* assessments, but this is balanced by a strong role for the dispositions in the *out of school* assessments. All of the above paints a clear and strong picture of teacher learning, or (we might say) of instructional problem-solving. In terms of the model of conception described at the beginning of this chapter, we could say that the teacher has learned to solve a particular task ("Produce an assessment for students that includes a significant role for the mathematical dispositions") by means of novel representations and operations (take-home problem sets containing a mix of retrospective and prospective items, using a blend of registers, etc.) with controls (can students complete the work in the allotted time, is there still some measure of supervision of students at work) to determine whether an assessment stands as a good "solution" to the task.

The above considerations lead to a somewhat mixed set of conclusions. On the one hand, the continued existence of a role for the mathematical dispositions in the assessments corpus seems to provide evidence that it is, indeed, possible for a high school Geometry course to hold students accountable for learning the mathematical sensibility. On the other hand, the adaptations described in this chapter, and the various sorts of feedback that I have argued might help explain those adaptations, strongly suggest that for a teacher to take on the responsibility for cultivating the mathematical sensibility in students takes more than just knowledge and the desire to do so: it also requires a commitment of time that competes with the default timeline of the curriculum, and calls for the development of novel forms of assessment — forms that compete with the (also significant) imperative for standardized, uniform testing. I take up these matters in the next chapter.

Chapter 6

Conclusions

Looking back

In this final chapter, I look back over the work done in Chapters 1-5. I review some of the key theoretical and methodological contributions of this dissertation, and discuss how they help shed light on the questions introduced in Chapter 1. I also attempt to develop a synthesis of the findings reported above, and propose some directions for further research.

In Chapter 1, I stated the main research question as: *What is the capacity for the high school Geometry course to serve as an opportunity for students to learn the elements of a "mathematical sensibility"?* The relevance of this question stems from the (not uncommon) position that the mathematics classroom can and should be conceptualized as more than just a delivery system for content knowledge; it also can function as an environment in which to cultivate in students some of the values and habits of mind of mathematicians. But this perspective frequently treats the notion of "mathematical sensibility" as a rather nebulous construct. In order to sharpen the question and to transform the intuitive notion of "mathematical sensibility" into something researchable, I undertook in Chapter 2 to produce a more detailed articulation of the elements of the mathematical sensibility.

I began my analysis by borrowing, from the sociology of Bourdieu (1998), the notion of *practical rationality*: the network of dispositions — shared categories of perception and appreciation — that characterize a practice. Throughout the remainder of the dissertation, I have taken "practical rationality of mathematicians" as a formal analogue of the informal notion of "mathematical sensibility". I also narrowed the scope of the question by examining in particular that portion of the mathematical sensibility that is concerned with appraising the value of *mathematical results,* i.e. implications of the form "if P, then Q". This portion of the mathematical sensibility includes within it the posing of problems, which may be regarded as partial anticipations of future results (e.g. "Does P imply Q?", "Which P implies Q?", etc.). This restriction of the conceptual domain excludes other important aspects of mathematical practice, such as problem-solving, proof, and conviction (both the process of becoming convinced of the truth of an assertion, and the process of convincing others). It excludes other important mathematical practices such as symbolization and translation across semiotic registers. Further research is necessary to articulate the mathematical sensibility as it applies to these other zones of mathematical practice.

My method for discovering the dispositions that comprise this restricted portion of the mathematical sensibility was rooted in the notion that practitioners' knowledge is usually tacit, and encoded in the form of narratives. Thus, I turned to a collection of narratives of mathematicians and their practice to find two kinds of dispositions:

(a) examples of the generative moves that mathematicians may use to move from one problem or result to the posing of a second problem;

(b) examples of the categories of value that mathematicians may use to describe the salient characteristics of a result or problem.

This analysis yielded a network of disposition that were organized in dialectical pairs, consistent with what philosophers of mathematics have suggested regarding the role of dialectic in providing a motive force within mathematics. These dispositions are reproduced below (Figure 6.1). These eight dialectical pairs played multiple roles in this dissertation: they were, on the one hand, the *findings* of the analysis in Chapter 2; at the same time they functioned as theoretical and methodological resources in the empirical studies of Chapter 3-5.

Categories of perception (generative moves):

(strengthen/weaken) hypothesis	⟷	(strengthen/weaken) conclusion
generalize	⟷	specialize
consider converse	⟷	consider converse

Categories of appreciation

utility	⟷	abstraction
surprise	⟷	confirmation
theory-building	⟷	problem-solving
simplicity	⟷	complexity
formalism	⟷	Platonism

Figure 6.1. The mathematical sensibility.

In the first of these empirical studies (Chapter 3) I analyzed a collection of conversations among groups of experienced teachers gathered in study groups around representations of teaching. This study made use of many of the same theoretical and methodological considerations of the analysis of narratives in Chapter 2, but in a different configuration. In Chapter 3, narratives of practice functioned not as a source of data, but rather as a collection of *probes* into the practical rationality of teachers; the dispositions comprising the mathematical sensibility found in the previous chapter were then deployed as an instrument for retrieving and aggregating relevant intervals from the large data corpus produced by Project ThEMaT. My analysis of these intervals showed that experienced teachers of high school Geometry generally disclaim responsibility for incorporating the various mathematical dispositions into their practice. To be sure, some of those dispositions have a greater affinity to teachers' practice than do others. In particular the data shows that teachers tend to view Geometry as being organized by various problem-solving methods, rather than by connections among the elements of the theory; that they apologize for the abstraction of the course, while paying (what appears to be) lip service to its utilitarian value; and that they reject mathematical formalism while endorsing a kind of tacit Platonism. Teachers reported that, because they feel obliged not to leave questions unresolved at the end of a lesson, their practice provides little or no opportunity for appraising results as either surprising or confirmatory. As for the generative moves, I found little evidence that teachers in the study groups regard problem-posing as something that they are supposed to teach, or that their students are supposed to learn: rather there is a clear division of labor in which posing problems is the responsibility of the teacher, and solving problems as posed is the responsibility of students. Teachers acknowledged that there may be benefits in creating contexts in which a student would be genuinely unsure of what is true ("healthy confusion") and that this uncertainty could help to promote a problem-posing orientation among students, but they also argued that those benefits are outweighed by the costs of alienating students, or of being perceived by students as dishonest or ignorant.

When teachers were asked to consider teaching scenarios that foregrounded one or more of those dispositions that are not commonly represented in practice, teachers pointed to a range of institutional factors as warrants for why those scenarios would not be viable. These institutional factors include: time constraints, an over-crowded curriculum, student mobility and inattentiveness, the need to prepare students for standardized exams, and a heterogeneous student body.

In Chapter 4, I turned to a collection of assessment items written and used by a teacher in Honors Geometry classes over a three-year period. In this chapter, I borrowed from literary criticism the important notions of "implied author" and "implied reader". These two constructs are anthropomorphisms, the former for the textual strategies deployed in a written document, the latter for the capacities that a text both presumes of and helps to create in its reader. I adapted these into the constructs of *implied teacher* and *implied student*. It is the implied teacher (not the empirical teacher) that students encounter and must adapt to when taking an exam, and it is the implied student (not the empirical student) that tells us what an assessment item holds students accountable for learning. Thus it is these two hypothetical figures, not their real-world counterparts, that are the parties bound together by the didactical contract.

I analyzed the individual assessment items by repurposing the dispositions of the mathematical sensibility as codes, and showed that those dispositions are accorded a significant role within the corpus. I illustrated this role with examples of items coded for each of the dispositions. While not all dispositions were as well-represented in the corpus, as a group they were found in more than 38% of all items, and there were an average of roughly 2 dispositions for every 3 items in the corpus. In contrast an examination of an analogous corpus of assessment items taken from the textbook used in the course found that the dispositions were found in about 16.5% of all items, that only five of the dispositions were represented at all, and that no items were coded with more than one disposition.

In Chapter 5, I examined the assessment items at multiple timescales for evidence of both coherence and adaptation across the three-year corpus. I showed that the items make use of all three metafunctions of language (ideational, textual, and interpersonal) to create a coherent background. Against this coherent background, changes in both the norms of assessment and in the formulation of individual items can be understood as a kind of teacher "learning" (i.e., adaptation). I documented the bifurcation of the assessment items corpus into two distinct subcorpora, one consisting of take-home problem sets and one consisting of more normal in-school quizzes. I argued that this adaptation can be understood as an accommodation between, on the one hand, the intention of the implied teacher to hold students accountable for learning the elements of a mathematical sensibility — an intention represented through the use of explicit textual strategies — and on the other hand some negative feedback on the teacher from the milieu of teaching. I turn now to a discussion of this negative feedback, and work towards a synthesis of the findings from Chapter 3 and Chapter 5.

What makes teaching the mathematical sensibility difficult?

It is important to recognize that there are some striking correspondences between the findings of Chapter 3 and those of Chapter 5. In particular, many of the warrants that were pointed to by the ThEMaT study group teachers for why they do not hold students accountable for learning the mathematical sensibility have analogues in the analysis of the assessment items corpus.

For example, teachers in the ThEMaT study groups frequently cited the heterogeneous nature of the students they teach as a constraint. They argued that if they had classes composed exclusively of bright students they could carve out a role for the mathematical sensibility — but since their classes include both honors students and special education students, such a role would not be workable. In this connection, it is perhaps not insignificant that the corpus of examination questions analyzed in Chapters 4-5 come from an *Honors* Geometry course. I point this out not to reify or endorse the teachers' claim, but rather to raise what seems to be a pertinent question: To what extent does the designation of a course as "honors" enable a teacher to pursue goals that would otherwise be too risky? Note that a student who struggles in an Honors course may be urged to transfer into a different, non-Honors section, an option that does not exist for a student already in a section not designated as "honors". What would have happened if the author of these items would have tried to use them in a non-Honors course? What kind of adaptation would have been generated in response to the feedback from that different milieu of teaching?

A second institutional constraint to which the ThEMaT teachers referred was their obligation to the intended curriculum, as represented by the textbook. As a general rule teachers do not choose their textbooks; textbooks are chosen for them, either by a department chair or by district-level selection committees. Moreover nearly all of the teachers in the study groups worked in schools where there was more than one section of Geometry, often with more than one teacher, and this produces a need (or at least a perceived need) for consistency across all sections. In marked contrast, the author of the assessment items taught in a small private school, in which there was only one section of the course, and was entirely free of district or state level mandates. This unusually high degree of autonomy almost certainly was a crucial affordance for the teacher. Again I raise this point not to endorse the study group teachers' point of view, but rather to flag an important, and still unanswered, question: Could assessments like the ones studied in Chapters 4-5, and the teaching that they imply, be viable in a more ordinary school context? That experiment has yet to be conducted.

A third issue that the ThEMaT study group teachers raised contrasts the limited number of contact hours available to them, with the amount of time it would take to create opportunities for students to practice and learn the mathematical dispositions. Teachers argued that there simply is not enough time in each day, or in the school year, to create such opportunities. The evidence from the assessment items corpus strongly suggests that there may be something to this: my analysis showed that, over the length of the three year corpus, the amount of time (as measured in weeks) devoted to each chapter consistently lengthens, with the result that by the end of Year 3 there was no time left for assessments devoted to the latter chapters of the textbook. On the other hand, it is too facile to conclude on this basis alone that the curriculum has not been adequately covered. As I showed in Chapter 5, much of the content that would normally be covered in the latter chapters of the textbook appears "prematurely" in the assessments of the earlier chapters. But while my analysis does not provide a basis for determining whether or not the curriculum has been adequately covered (and indeed, answering such a question would require a carful consideration of what "covering" the curriculum even means), it certainly seems to be the case that the intention to teach the mathematical sensibility forces a change in curricular focus, and may be incompatible with the default, "official" timeline of the year as determined by the textbook.

Looking forward

The theoretical perspectives and methods deployed in this dissertation, and the findings reported in it, suggest a number of directions for future research. In this final section I anticipate some of those directions, and discuss how the present work might prove instrumental in investigating them.

One direction concerns the application of the methodology of Chapters 3-5 to study other sorts of data. In particular, my articulation of the dispositions shows in Figure 6.1 could be used to look for evidence of the mathematical sensibility in, for example, field records, student work, teachers' lesson plans, and so forth. Moreover the scope of such investigations could be broadened to look beyond the Geometry course, and into other areas of K-12 mathematics. The dispositions could also be useful in a study of post-secondary mathematics education, and in particular of how mathematics majors and graduate students become enculturated into the rationality of mathematical practice.

A second direction concerns the generality of the findings reported here. As discussed above, it is far from clear to what extent the teaching described in Chapters 4-5 (as evidenced by the assessments) depended on a particular institutional context: namely, a small, private high school with a single section of Honors geometry. What would happen if similar assessments were used in larger schools, or with more heterogeneous groups of students, or by more than one teacher? Put another way, what different sorts of feedback on the teacher would be produced by a different milieu of teaching, and what kinds of adaptation might result from that feedback?

A third direction concerns a broader application of the constructs developed in Chapter 4, namely the implied teacher and the implied student. I believe these constructs provide a disciplined way of disentangling *teaching* from the individual *teacher*, by providing a language for describing the *intentio operis* as distinct from the *intentio auctoris*. There is, at present, a rather well-developed body of educational research that focuses on teachers' beliefs and values; without discounting the great value and importance of such research, it is also worth considering whether this focus on the teacher's intention has led to the establishment of a version of the "intentional fallacy" in the field of education research (see footnote, p. 183 above). Certainly from the practitioner's point of view intentions matter, and insofar as the education research community is concerned with preparing future practitioners the focus on intention is appropriate and vital. But as observers of practice, we also need a way to describe teaching without requiring knowledge of the teacher's intent.

A fourth direction concerns the potential use of the findings reported here in teacher education. To what extent can teacher education take responsibility for cultivating a mathematical sensibility in *preservice teachers*? This question is in some sense the analogue of the overall research problem of the dissertation, transposed into a different context. For example, could some of the assessment items from the corpus be adapted for use in a "Geometry for teachers" course? What kind of feedback would *that* milieu produce, and how would the items evolve in response to that feedback?

One final issue that should be raised here concerns the teacher's own mathematical knowledge, and how that knowledge becomes useful in crafting opportunities for students to learn. Throughout this dissertation I have deliberately avoided raising these questions, as my intent throughout was to focus not on the *individual teacher* but on the *teaching*. But they are nonetheless questions that should be addressed. At the time that I wrote the assessment items studied here[39], I had completed most of the requirements for a doctorate in pure mathematics. It would be naïve to pretend that this played no role whatsoever in my writing the items. On the contrary, creating problems for students is one of the key activities of teaching in which mathematical knowledge for teaching (MKT), and in particular the specialized content knowledge of teachers, can be brought into play (Ball, Thames & Phelps 2008). It certainly seems reasonable to at least conjecture that my own mathematical knowledge was in some way instrumental here. But in *what* way, precisely? Surely not everything that I knew about mathematics was equally valuable in my teaching practice. What kind of knowledge, specifically, played a role?

In terms of the position I outlined in Chapter 2 — that knowledge of a practice is encoded in narratives of practice — it seems reasonable to conjecture that the answer to this last question might best be represented by narratives of my own mathematical practice. And indeed, when asked about my own mathematical sensibility, I (like all practitioners) do tend to respond with stories and anecdotes. In studying the range of questions associated with teachers' mathematical sensibility, it may be worth considering

[39] Here I drop, at last, the third-person mask I have worn since p. 167. At issue now is not the hypothesized implied teacher, but the actual, flesh-and-blood empirical teacher.

the proposition that, as scholars and teacher educators, we need to attend to teachers' own stories of mathematical investigation.

In connection with this last point, in the Fall of 2006 I began a pilot study of a "Geometry for teachers" course. As part of that study I interviewed four preservice teachers enrolled in the course. The final interview questions were, "Have you ever made a mathematical discovery of your own? If so, what was it, and how did you come to discover it?" These questions were designed to elicit stories of mathematical wondering, of the student as a problem-poser in an encounter with the unknown. The results, although only preliminary, were nevertheless very suggestive: three of the four students had no stories of their own to tell. This suggests that one of the missing links in current conceptualizations of MKT may be the storied nature of professional knowledge. It further suggests that one of the goals of teacher education might be the creation of opportunities for preservice teachers to acquire their own narratives of practice.

There is something admittedly paradoxical about this proposal: I seem to be suggesting that preservice teachers should pose their own problems and pursue their own investigations, and at the same time that these freely-chosen, autonomous investigations should be carried out within the institutional boundaries of a mandatory teacher education course. The paradox is essentially the same as that discussed in Chapter 1 in connection with the NCTM's call for all students to learn to become problem-posers. Recall that in my critique of the "Lei" anecdote from the *Standards* (see p. 3 of this dissertation) I pointed out there that it is easier to create a single illustration of an individual student engaged in serious mathematical inquiry than it is to describe the *work of teaching* that makes such inquiry possible in a classroom. The take-home assessment items described in the preceding chapters —in particular, the “antwalk” problem from the Y2 midterm and the “duals” research project from the Y2 and Y3 final exam — provide an image of what such work might look like. Those problems both create a context in which students are obliged to pose problems, while at the same time bounding the conceptual space in which such problem-posing is to be done, and providing resources to support such work.

It is my hope that further development along the lines described above may help to push the goal of cultivating of a mathematical sensibility into the foreground not only of teacher education, but also of K-12 education more generally.

References

Arnol'd (2000). On A.N. Kolmogorov. In *Kolmogorov in Perspective* (pp. 89-108). Providence: American Mathematical Society.

Askew, M. (2008). Mathematical discipline knowledge requirements for prospective primary teachers, and the structure and teaching approaches of programs designed to develop that knowledge. In P. Sullivan and T. Wood (eds.), *Knowledge and Beliefs in Mathematics Teaching and Teaching Development* (pp. 13–35). Rotterdam: Sense Publishers.

Balacheff N. & Gaudin N. (2002). Students conceptions: an introduction to a formal Characterization. In *Les cahiers du laboratoire Leibniz*, No. 65. (http://www-leibniz.imag.fr/LesCahiers/)

Ball, D.L. (1988). Knowledge and reasoning in mathematical pedagogy: Examining what prospective teachers bring to teacher education. Unpublished doctoral dissertation, Michigan State University, East Lansing.

Ball, D.L. (2000). Working on the inside: Using one's practice as a site for studying mathematics teaching and learning. In Kelly, A. & Lesh, R. (Eds.). *Handbook of research design in mathematics and science education* (pp. 365-402). Mahwah, NJ: Lawrence Erlbaum Associates.

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on the teaching and learning of mathematics* (pp. 83-104). Westport, CT: Ablex.

Ball, D.L. & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, W.G. Martin, and D. Schifter (Eds.) *A Research Companion to Principals and Standards for School Mathematics* (pp. 27-44). Reston, VA: National Council of Teachers of Mathematics.

Ball, D. L., Thames, M., and Phelps, G. (2008). Content Knowledge for Teaching: What Makes It Special? *Journal of Teacher Education*, 59(5), 389-407.

Bass, H. (2005). Mathematics, mathematicians and mathematics education. *Bulletin of the American Mathematical Society* 42(4), 417-430.

Bell, A., Burkhardt, H., & Swan, M. (1992a). Balanced Assessment of Mathematical Performance. In R.A. Lesh and S.J. Lamon (Eds.), *Assessment of Authentic Performance in School Mathematics* (pp. 119-144). Washington, DC: American Association for the Advancement of Science Press.

Bell, A., Burkhardt, H., & Swan, M. (1992b). Assessment of Extended Tasks. In R.A. Lesh and S.J. Lamon (Eds.), *Assessment of Authentic Performance in School Mathematics* (pp. 145-176). Washington, DC: American Association for the Advancement of Science Press.

Bell, E.T. (1965). *Men of mathematics.* New York, NY: Simon and Schuster.

Berthoff, A. (1987). The teacher as RE-searcher. In D. Goswami & P. Stillman (Eds.), *Reclaiming the classroom: Teacher research as an agency for change* (pp. 28-38). Upper Montclair, NJ: Boynton/Cook.

Bleakley, A. (2000). Writing with invisible ink: narrative, confessionalism and reflective practice. *Reflective Practice* 1(1), 11-24.

Booth, Wayne C. (1961). *The Rhetoric of Fiction*. Chicago: University of Chicago Press.

Bourdieu, P. (1998). *Practical reason*. Stanford, CA: Stanford University Press.

Brousseau, G. (1997). *Theory of didactical situations in mathematics*. The Netherlands: Kluwer Academic Publishers.

Brown, J.S., (2005). "Narrative as a Knowledge Medium in Organizations". In J.S. Brown, S. Denning, K. Groh, & L. Prusak (Eds.), *Storytelling in organizations: Why storytelling is transforming 21st century organizations and management* (pp. 53-96). Burlington, MA: Elsevier Butterworth-Heinemann.

Bruner, J. (1960). On learning mathematics. *The Mathematics Teacher*, 53, 610-619.

Caldwell, J.H., & Goldin, G.A. (1987). Variables affecting word problem difficulty in secondary school mathematics. *Journal for Research in Mathematics Education,* 18 (3), 187-196.

Carpenter, T. P., Moser, J. M., & Romberg, T. A. (1982). *Addition and subtraction: A cognitive perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Carpenter, T. P., & Moser, J. M. (1984). The acquisition of addition and subtraction concepts in grades one through three. *Journal of Research in Mathematics Education*, 15, 179-202.

Chase, W.G. & Simon, H.A. (1973). Perception in chess. *Cognitive Psychology,* 4, 55-81.

Chazan, D., Callis, C., & Lehman, M. (2008). *Embracing Reason: Egalitarian Ideals and the Teaching of High School Mathematics.* New York: Routledge.

Chazan, D. and Yerushalmy, M. (1992). Research and classroom assessment of students' verifying, conjecturing, and generalizing in Geometry. In R.A. Lesh and S.J. Lamon (Eds.), *Assessment of Authentic Performance in School Mathematics* (pp. 89-115). Washington, DC: American Association for the Advancement of Science Press.

Christofferson, H.C. (1930). A fallacy in geometry reasoning. *Mathematics Teacher*, 23, 19-22.

Cochran-Smith, M., & Lytle, S. (1990). Research on teaching and teacher research: The issues that divide. *Educational Researcher,* 19(2), 2-11.

Cochran-Smith, M., & Lytle, S. (1999). The teacher research movement: A decade later. *Educational Researcher,* 28(7), 15-25.

Committee on Prospering in the Global Economy of the 21st Century: An Agenda for American Science and Technology and Committee on Science, Engineering, and Public Policy (2007). RISING ABOVE THE GATHERING STORM: *Energizing and Employing America for a Brighter Economic Future.* Washington, DC: National Academies Press.

Corfield, D. (2001). The importance of mathematical conceptualisation. *Studies in the History and Philosophy of Science,* 32(3), 507-533.

Curtis, C.W., Daus, P.H., and Walker, R.J. (1961). *Euclidean geometry based on ruler and protractor axioms.* New Haven: School Mathematics Study Group.

Cusick, P. (1983). *The egalitarian ideal and the American high school: Studies of three schools.* New York: Longman.

Davis, P. (2000). *The education of a mathematician.* Natick, MA: A.K. Peters.

de Groot, A.D. (1978). *Thought and Choice in Chess.* The Netherlands: The Hague.

de Lange, J. (1992). Assessing mathematical skills, understanding, and thinking. In R.A. Lesh and S.J. Lamon (Eds.), *Assessment of Authentic Performance in School Mathematics* (pp. 195-214). Washington, DC: American Association for the Advancement of Science Press.

Dewey, J. (1902). *The child and the curriculum.* Chicago: The University of Chicago Press.

Diez, M.E. & Raths, J. (2007). *Dispositions in teacher education* (eds.) Charlotte, NC: Information Age Publishing.

Dottin, E.S. (2009). Professional judgment and dispositions in teacher education. *Teaching and Teacher Education* 25(1), 83-88.

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist*, 23, 167-180.

Eco, U. (1979). *The role of the reader: Explorations in the semiotics of texts*. Bloomington: Indiana University Press.

Eco, U. (1992). Between author and text. In Collini, S., (Ed.), *Interpretation and Overinterpretation* (pp. 67-68). Cambridge: Cambridge University Press.

Eco, U. (1994). *The Limits of Interpretation.* Bloomington: Indiana University Press.

Feiman-Nemser, S. & Remillard, J. (1992). Perspectives on Learning to Teach. In F. Murray (Ed.), *The Teacher Educator's Handbook* (pp. 63-91). San Francisco: Jossey-Bass.

Goldin, G. (1992). Toward and assessment framework for school mathematics. In R.A. Lesh and S.J. Lamon (Eds.), *Assessment of Authentic Performance in School Mathematics* (pp. 63-88). Washington, DC: American Association for the Advancement of Science Press.

González, G. (2009). *Mathematical Tasks and the Collective Memory: How Do Teachers Manage Students' Prior Knowledge When Teaching Geometry with Problems?* Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

González, G. & Herbst, P. (2006). Competing arguments for the Geometry course: Why were American high school students supposed to study Geometry in the twentieth century? *The International Journal for the History of Mathematics Education* 1(1), 7-33.

Goodwin, C. (1994). Professional vision. *American Anthropologist* 96(3), 606-633.

Gowers, W.T. (2000). The two cultures of mathematics. In V.I. Arnold, M. Atiyah, & B.W. Mazur (Eds.), *Mathematics: Frontiers and Perspectives* (pp. 65-78). Providence, RI: American Mathematical Society.

Guba, E.G., and Lincoln, Y. (1994). Competing Paradigms in Qualitative Research. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 105-117). Thousand Oaks, CA: Sage.

Grabiner, J. (1974). Is mathematical truth time-dependent? *The American Mathematical Monthly,* 81(4), 354-365.

Greenberg, M. (1980). *Euclidean and non-Euclidean geometries: development and history* (2nd ed.). San Francisco: Freeman.

Gutstein, E. (2003). Teaching and learning mathematics for social justice in an urban, Latino school. *Journal for Research in Mathematics, Education,* 34(1), 37-73.

Gutstein, E. (2006). *Reading and writing the world with mathematics.* New York: Routledge.

Halliday, M.A.K. (1994). *An introduction to functional grammar* (2nd edition). London: Edward Arnold.

Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Hardy, G.H. (1992). *A mathematician's apology.* New York: Cambridge University Press. (Originally published 1941).

Hasan, R. (1984). Coherence and cohesive harmony. In J. Flood (Ed.), *Understanding reading comprehension: Cognition, language and the structure of prose* (pp. 181-219). Newark, Delaware: International Reading Association.

Heaton, R. M. (1994). Creating and studying a practice of teaching elementary mathematics for understanding. Unpublished doctoral dissertation, Michigan State University, East Lansing.

Hegel, G.W.F. (1807 / 1977). *Phenomenology of the Spirit.* (A.V. Miller, Trans.) Oxford: Clarendon Press.

Herbst, P. (2004). Interactions with diagrams and the making of reasoned conjectures in geometry. *Zentralblatt für Didaktik der Mathematik*, 36, 129-139.

Herbst, P. (2006). Teaching Geometry with problems: Negotiating instructional situations and mathematical tasks. *Journal for Research in Mathematics Education*, 37(4), 313-347.

Herbst, P. (2006-2009). Internal ThEMaT Project memoranda.

Herbst, P. & Brach, C. (2006). Proving and doing proofs in high school geometry classes: What is it that is going on for students? *Cognition and Instruction* 24(1), 73-122.

Herbst, P. & Chazan, D. (2003a). Exploring the practical rationality of mathematics teaching through conversations about videotaped episodes: the case of engaging students in proving. *For the Learning of Mathematics,* 23(1), 2-14.

Herbst, P. and Chazan, D. (2003b). ThEMaT: Thought Experiments in Mathematics Teaching. *Proposal to the National Science Foundation, Education and Human Resources Directorate, Division of Elementary, Secondary, and Informal Education, Teachers' Professional Continuum Program.*

Herbst, P. & Chazan, D. (2006). Producing a viable story of geometry instruction: What kind of representation calls forth teachers' practical rationality? In Alatorre, S., et al. (Eds.), *Proc. of the 28th PME-NA Conference* (Vol. 2, pp. 213-220). Mérida, México: UPN.

Herbst, P. & Miyakawa, T. (2008). When, how, and why prove theorems? A methodology for studying the perspective of geometry teachers. *Zentralblatt für Didaktik der Mathematik*, 40(3), 469-486.

Herbst, P. & Nachlieli, T, & Chazan, D. (submitted). Studying the practical rationality of mathematics teaching: What goes into "installing" a theorem in Geometry?

Hofstadter, D. (1979). *Gödel, Escher, Bach: An eternal golden braid.* New York: Basic Books.

Huberman, M. (1996). Focus on research moving mainstream: Taking a closer look at teacher research. *Language Arts,* 73(2), 124-140.

Jurgensen, R., Brown, R., and Jurgensen, J. (1990). *Geometry.* Austin: Holt, Rinehart & Winston.

Kanigel, R. (1991). *The man who knew infinity: a life of the genius, Ramanujan.* New York: Maxwell Macmillan International.

Kiss, O. (2006). Heuristic, methodology, or logic of discovery? Lakatos on patterns of thinking. *Perspectives on Science* 14(3), 302-317.

Kline, M. (1953). *Mathematics in western culture.* London: Oxford University Press.

Kulm, G. (1990). *Assessing higher order thinking in mathematics.* Lawrence Erlbaum Associates.

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery* (J. Worrall & E. Zahar, Eds.). Cambridge: Cambridge University Press.

Lakatos, I. (1978) History of science and its rational reconstructions. In J. Worrall & G. Currie (Eds.), *The methodology of scientific research programmes: Imre Lakatos' Philosophical Papers* (Vol. 1, pp. 102-138). Cambridge: Cambridge University Press.

Lampert, M. (2001). *Teaching problems and the problems of teaching.* New Haven: Yale University Press.

Lang, S. & Murrow, G. (1988). *Geometry* (2nd edition). New York: Springer-Verlag.

Larson, R., Boswell, L, & Stiff, L. (2001). *Geometry*. McDougal Littell: Evanston, IL.

Larvor, B. (1999). Lakatos' mathematical Hegelianism. *The Owl of Minerva* 31(1). (http://hdl.handle.net/2299/460)

Larvor, B. (2001). What is dialectical philosophy of mathematics? *Philosophia Mathematica* 3(9), 212-229.

Le Lionnais, F. (1948/1986). La beauté en mathématiques. In F. Le Lionnais (Ed.), *Les grands courants de la pensée mathématique* (pp. 437-465). Paris: Editions Rivages.

Leinhardt, G. & Ohlsson, S. (1990). Tutorials on the structure of tutoring from teachers. *Journal of Artificial Intelligence in Education* 2(1), 21-46.

Leinhardt, G. & Steele, M. (2005). Seeing the complexity of standing to the side: Instructional dialogues. *Cognition and Instruction,* 23(1) 87-163.

Lemke, J.L. (2000). Across the scales of time: Artifacts, activities, and meanings in eco-social systems. *Mind, Culture, and Activity*, 7, 273-290.

Lemke, J.L. (2002). "Mathematics in the Middle: Measure, Picture, Gesture, Sign, and Word." In Anderson, M., Saenz-Ludlow, A., Zellweger, S. & Cifarelli, V., (Eds.). *Educational Perspectives on Mathematics as Semiosis: From Thinking to Interpreting to Knowing.* pp. 215-234. Ottawa: Legas Publishing.

Lesh, R.A. and Lamon, S.J. (1992). *Assessment of Authentic Performance in School Mathematics.* Routledge.

Li, Yeping. (2000). A comparison of problems that follow selected content presentations in American and Chinese mathematics textbooks. *Journal for Research in Mathematics Education,* 31(2), 234-241.

Martin, J.R. & Rose, D. (2003). *Working with discourse: Meaning beyond the clause.* London: Continuum.

Mehan, H. & Wood, H. (1975). Five features of reality. In *The reality of ethnomethodology* (pp. 8-33). New York: John Wiley & Sons.

Mesa, V. (2004). Characterizing practices associated with functions in middle school textbooks: An empirical approach. *Educational Studies in Mathematics,* 56, 255-286.

Misco, T. & Shiveley, J. (2007). Making sense of dispositions in teacher education: arriving at democratic aims and experiences. *Journal of Educational Controversy* 2(2). Retrieved online on 6/01/2009 from http://www.wce.wwu.edu/Resources/CEP/ejournal/v002n002/a012.shtml.

Miyakawa, T. & Herbst, P. (2007a). The nature and role of proof when installing theorems: the perspective of geometry teachers. In J. H. Woo et al. (Eds.), *Proc. of the 31st PME* (Vol.3, pp. 281-288). Seoul: PME.

Miyakawa T. & Herbst P. (2007b). Geometry teachers' perspectives on convincing and proving when installing a theorem in class. In T. Lamberg et al. (Eds.), *Proc. of the 29th annual meetings of PME-NA* (pp. 366-373). Lake Tahoe: University of Nevada, Reno.

Moise, E. & Downs, F. (1991). *Geometry*. Menlo Park, CA: Addison-Wesley.

Moses, R. P. (2001). *Radical equations: Math literacy and civil rights*. Boston: Beacon Press.

Munkres, J.R. (1975). *Topology: A first course.* Englewood Cliffs, NJ: Prentice-Hall.

Nathan, M.J., & Long, S.D. (2002). The symbol precedence view of mathematical development: A corpus analysis of the rhetorical structure of textbooks. *Discourse Processes,* 33(1), 1-21.

NCTM. (1991). *Professional standards for teaching mathematics.* Reston VA: National Council of Teachers of Mathematics.

NCTM (2000). *Principles and standards for school mathematics.* Reston, VA: National Council of Teachers of Mathematics.

Newcomb, S. et al. (1893). Reports of the conferences: Mathematics. In National Education Association, *Report of the Committee on secondary school studies* (pp. 104-116). New York: Arno Press.

Niss, M. (Ed.) (1993). *Investigations into Assessment in Mathematics Education.* Boston: Kluwer Academic.

Oakley, C.O. & Baker, J.C. The Morley trisector theorem. *American Mathematical Monthly* 85(9), 737-745.

Pais, A. (1982). *Subtle is the Lord: The science and the life of Albert Einstein*. New York: Oxford University Press.

Parikh, C. (1991). *The unreal life of Oscar Zariski.* Boston: Academic Press.

Parker, J. (2005). *R.L. Moore: mathematician and teacher.* Mathematical Association of America.

Peirce, B. (1882). *Linear associative algebra.* New York: D. Van Nostrand.

Piaget, J. (1975). *Equilibration of cognitive structures*. Chicago: University of Chicago Press.

Pimm, D. (1984). Who is "we"? *Mathematics Teaching,* 107, 39-42.

Pimm, D. (1987). *Speaking mathematically: Communication in mathematics classrooms.* London: Routledge and Kegan Paul.

Poincaré, H. (2001). Intuition and logic in mathematics. In H. Poincaré, *The value of science* (S. J. Gould, Ed; pp. 197-209). New York: The Modern Library. (Original work published 1913)

Polanyi, M. (1997). Tacit knowledge. In L. Prusak (Ed.), *Knowledge in Organizations* (pp. 135-146). Newton, MA: Butterworth-Heinemann.

Polya, G. (1957). *How to solve it: A new aspect of mathematical method* (2nd ed.). Garden City, NY: Doubleday.

Popper, Karl. (1940). What is dialectic? *Mind* 49, 403-26.

Ransom, J.C. (1941). *The new criticism.* Norfolk, CT: New Directions.

Rothstein, E. (1995). *Emblems of mind: The inner life of music and mathematics.* Chicago: University of Chicago Press.

Rowland, T. (1992). Pointing with pronouns. *For the Learning of Mathematics,* 12(2), 44-48.

Rowland, T. (1999). Pronouns in mathematics talk: Power, vagueness and generalisation. *For the Learning of Mathematics,* 19(2), 19-26.

Schön, D. (1983). The reflective practitioner: How professionals think in action. New York: Basic Books.

Schultz, J., Hollowell, K., Ellis, W., & Kennedy, P. (2001). *Geometry*. Austin: Holt, Rinehart & Winston.

Schwab, J. (1961/1974). Education and the structure of the disciplines. In I. Westbury & N. Wilkof (Eds.), *Science, curriculum, and liberal education* (pp. 229-272). Chicago: University of Chicago Press.

Schwab, J. (l976). Education and the state: Learning community. In *Great Ideas Today* (pp. 234-71). Chicago: Encyclopedia Britannica.

Searle, J. (1975). Indirect speech acts. In P. Cole & J.L. Morgan (Eds.), *Syntax & Semantics, 3: Speech Acts.* New York: Academic Press.

Serra, M. (1997). *Discovering Geometry*: *An inductive approach.* Berkeley, CA: Key Curriculum Press.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher,* 27(2), 4-13.

Shiryaev, A.N. (2000). Andrei Nikolaevich Kolmogorov (April 25, 1903 to October 20, 1987). A biographical sketch of his life and creative paths. In *Kolmogorov in Perspective* (pp. 1-88). Providence: American Mathematical Society.

Simon, H.A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, 7, 268-288.

Sinclair, N. (2002). *Mindful of Beauty: The roles of the aesthetic in the doing and learning of mathematics.* Unpublished doctoral dissertation, Queen's University, Kingston, Ontario, Canada.

Singh, S. (1998). *Fermat's Enigma: The Epic Quest to Solve the World's Greatest Mathematical Problem.* New York: Walker.

Steffe, L. (2004). On the construction of learning trajectories of children: The case of commensurate fractions. *Mathematical Thinking and Learning*, 6(2), 129-162.

Stein, M.K., Smith, M., Henningsen, M., & Silver, E.A. (2000). *Exploring cognitively challenging mathematical tasks: A casebook for teacher professional development.* New York: Teachers College Press.

Stigler, J.W., Fuson, K.C., Ham, H., & Kim, M.S. (1986). An analysis of addition and subtraction word problems in American and Soviet elementary mathematics textbooks. *Cognition and Instruction*, 3 (3), pp. 153-171.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.

Thompson, G. & Zhou, J. (2001). Evaluation and organization in text: The structuring role of evaluative disjuncts. In S. Hunston & G. Thompson (Eds.), *Evaluation in Text: Authorial stance and the construction of discourse* (pp. 121-141). Oxford: Oxford University Press.

Thurston, W. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society,* 30, 161-177

Tsoukas, H. & Hatch, M.J. (2005) Complex thinking, complex practice: The case for a narrative approach to organizational complexity. In H. Tsoukas (Ed.), *Complex Knowledge: Studies in Organizational Epistemology* (pp. 230-262). London: Oxford University Press.

Tyack, D., & Tobin, W. (1994). The "grammar" of schooling: Why has it been so hard to change? *American Educational Research Journal,* 31, 453-79.

Ulam, S. (1976). *Adventures of a mathematician.* New York: Scribner.

Usiskin, Z. (1980). What should not be in the algebra and geometry curricula of average college-bound students? *The Mathematics Teacher*, 73, 413-424.

von Glasersfeld, E. (1995). *Radical Constructivism: A Way of Knowing and Learning*. London: Falmer Press.

Weiss, M. & Herbst, P. (2007). "Every single little proof they do, you could call it a theorem": Translation between abstract concepts and concrete objects in the Geometry classroom. Paper presented at 2007 meeting of the American Education Research Association, Chicago, IL.

Weiss, M., Herbst, P., and Chen, C. (2009). Teachers' perspectives on "authentic mathematics" and the two-column proof form. *Educational Studies in Mathematics* 70(3), 275-293.

Wenger, E. (1997). *Communities of Practice: Learning, Meaning, and Identity*. New York: Cambridge University Press.

Wiener, N. (1956). *I am a mathematician.* London: Gollancz.

Wilson, S. (1995). Not tension but intention: A response to Wong's analysis of the researcher/teacher. *Educational Researcher,* 24(8), 19-22.

Wimsatt, W.K. & Beardsley, M.C. (1946/1999). The intentional fallacy. In N. Warburton (Ed.), *Philosophy: Basic Readings* (pp. 480-492). New York: Routledge.

Wong, E.D. (1995). Challenges confronting the researcher/teacher: conflicts of purpose and conduct. *Educational Researcher,* 24(3), pp. 22-28.

Wu, H. (1999). *Some remarks on the teaching of fractions in elementary school*. Retrieved electronically from http://math.berkeley.edu/~wu/fractions2.pdf on November 8, 2007.