Winter 2001

Learning and understanding in abstract algebra

Bradford Reed Findell
University of New Hampshire, Durham

Follow this and additional works at: https://scholars.unh.edu/dissertation

LEARNING AND UNDERSTANDING IN ABSTRACT ALGEBRA

BY

BRADFORD R. FINDELL

BSE, Princeton University, 1985

MA, Boston University, 1990

DISSERTATION

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

in

Mathematics Education

December, 2001

UMI Number: 3030602

Copyright 2001 by
Findell, Bradford Reed

All rights reserved.

UMI®

UMI Microform 3030602
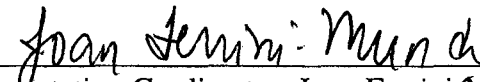
This dissertation has been examined and approved.
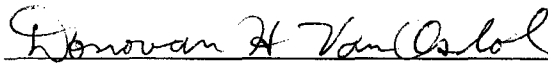
_____
Dissertation Co-director, Karen J. Graham, Professor of Mathematics
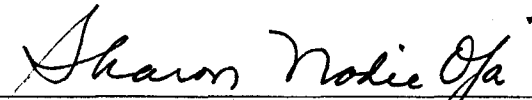
_____
Dissertation Co-director, Joan Ferrini-Mundy, Professor of
Mathematics and Teacher Education, Michigan State University;
Former Professor of Mathematics, University of New Hampshire

_____
Edward K. Hinson, Associate Professor of Mathematics

_____
Donovan H. Van Osdol, Professor of Mathematics

_____
Sharon Nodie Oja, Professor of Education

_____
12/03/2001
Date

DEDICATION

for Jenny, Elizabeth, and Randall Todd,
who exceed all hopes
and make the world grow with possibilities

iv

ACKNOWLEDGMENTS

Perhaps it takes several villages to raise a scholar, for if I do indeed become a scholar, it will be largely because of the support, encouragement, and prodding I have received from individuals of many villages in the world of mathematics education and beyond.

I thank my advisor, Joan Ferrini-Mundy, for helping me craft this study and its account into a coherent whole, for pushing me to articulate and refine many of the premises that were implicitly guiding my thinking, and for providing support and encouragement throughout the process. I am indebted to Joan not only for directing my graduate program and this dissertation but also for providing an abundance of professional opportunities over the past eight years. In particular, I am grateful for the opportunity to join her at the National Research Council, where she again served as my supervisor and mentor.

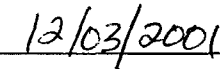Thanks go to Karen Graham, for stepping in as co-director of this study, for helping me sharpen my ideas, for providing an ear to help me think through what needed to be done, for merely smiling at my suggestion that I write an 80-page dissertation, and for playing such a significant and formative role throughout my graduate program.

Thanks also go to the other members of my committee, Nodie Oja, Ed Hinson, and Don Van Osdol, who individually provided keen critiques of drafts and who together urged me to integrate the many disparate ideas that those drafts contained. I also thank Ann Diller for providing helpful insights in the early phases of this study.

I was particularly fortunate in my work at the National Research Council to have an opportunity to work with Jeremy Kilpatrick, when he served masterfully as the chair of the Mathematics Learning Study Committee. I owe him a debt that can never be repaid for the interest he has taken in my career and for encouraging me to come to the University of Georgia to be his colleague. Jeremy graciously responded to drafts of several of the chapters, helped me hone my message, and provided editorial advice at all levels.

I also thank the Mathematics Education Department at the University of Georgia, especially Pat Wilson, for the confidence they have placed in me and the encouragement they all have provided during the completion of this study. It is a precious opportunity to be able to work and learn in such a vital and vibrant department.

Thanks also to my friends and colleagues in mathematics education, including Carole Greenes, Pam Matthews, Al Cuoco, Paul Goldenberg, Ed Dubinsky, Jim Fey, Kathy Heid, Jim Leitzel, Rodger Bybee, Roger Howe, Hung-Hsi Wu, Kevin Miller, Ed Silver, Steve Monk, Ann Shannon, Barbara Edwards, Dawn Berk, Tim Gutmann, and Todd Grundmeier who provided encouragement, support, and guidance and who often told me, in one way or another, not only that I could complete this study but that I should.

Of course, I would be nowhere without the support of my family. I owe a great debt to my mother, Carol, for pointing me in the direction of mathematics education years ago, for showing me what it means to engage with mathematics and to engage students in mathematics, and, more recently, for helping me enter data, for responding to drafts, and for letting three of us live in her house all summer. I also thank my brother Kent and his wife Vickie for providing support and frequent meals. I am grateful to my sisters Karen,

Leanne, and Kirsten for their encouragement and particularly to Kirsten for not gloating that she had completed her dissertation first. A special thanks goes to my daughter Elizabeth, who always amazes and who provides no end of joy, and to my son-to-be Randy, whom I eagerly wait to meet.

Most of all I thank my beloved Jenny, who has single-handedly and unselfishly managed three (almost four) lives for the past nine months, including two moves with the anticipation of a third, and who deserves so much more than I have been able to give. During this ever-expanding study, Jenny also graciously served as a sounding board for my dilemmas, frustrations, and nascent ideas, often helping me formulate those ideas more clearly. There cannot be a more generous and forgiving partner.

Since I began teaching mathematics more than sixteen years ago, there have been many moments at which I believed and acted as though I brought some expertise to the tasks before me, whether I was teaching, writing, editing, or talking about these activities. Not so these past nine months. Several times, however, in moments of singular clarity, I have caught a fleeting glimpse of what it means to be a scholar, of how it feels to be subservient to what the data have to say, and of what is involved in building an idea that seems new. I hope that these glimpses will become more frequent.

TABLE OF CONTENTS

LIST OF TABLES

xi

LIST OF FIGURES

ABSTRACT

LEARNING AND UNDERSTANDING IN ABSTRACT ALGEBRA

by

Bradford R. Findell

University of New Hampshire, December, 2001

Students' learning and understanding in an undergraduate abstract algebra class were described using Tall and Vinner's notion of a concept image, which is the entire cognitive structure associated with a concept, including examples, nonexamples, definitions, representations, and results. Prominent features and components of students' concept images were identified for concepts of elementary group theory, including group, subgroup, isomorphism, coset, and quotient group.

Analysis of interviews and written work from five students provided insight into their concept images, revealing ways they understood the concepts. Because many issues were related to students' uses of language and notation, the analysis was essentially semiotic, using the linguistic, notational, and representational distinctions that the students made to infer their conceptual understandings and the distinctions they were and were not making among concepts. Attempting to explain and synthesize the results of the analysis became a process of theory generation, from which two themes emerged: making distinctions and managing abstraction.

The students often made nonstandard linguistic and notational distinctions. For example, some students used the term *coset* to describe not only individual cosets but also the set

of all cosets. This kind of understanding was characterized as being immersed in the process of generating all of the cosets of a subgroup, a characterization that described and explained several instances of the phenomenon of failing to distinguish between a set and its elements.

The students managed their relationships with abstract ideas through metaphor, process and object conceptions, and proficiency with concepts, examples, and representations. For example, some students understood a particular group by relying upon its operation table, which they sometimes took to be the group itself rather than a representation. The operation table supported an object conception even when a student had a fragile understanding of the processes used in forming the group.

Making distinctions and managing abstraction are elaborated as fundamental characteristics of mathematical activity. Mathematics thereby becomes a dialectic between precision and abstraction, between logic and intuition, which has important implications for teaching, teacher education, and research.

CHAPTER I

INTRODUCTION

"Mathematics is the science of order, patterns, structure, and logical relationships."
(Devlin, 2000, p. 74)

In a compelling new book *The Math Gene*, renowned mathematician, expositor, and

National Public Radio commentator Keith Devlin (2000) claims that everyone has innate

ability to do mathematics because "the features of the brain that enable us to do

mathematics are the very same features that enable us to use language" (p. 2). A key

point in his argument is that richer representation rather than richer communication was

the driving force behind the emergence of language.

> In order to properly understand how we acquired language, we should view it as
> a representational structure rather than as a medium of communication. In order
> to communicate some concept, you first need to have a mental representation of
> it. (p. 291)

His argument draws on a broad body of empirical and theoretical work in anthropology,

neuroscience, linguistics, psychology, mathematics education, and also upon his

contention that most people do not know what mathematics is. "Modern mathematics,"

he claims, "is about abstract patterns, abstract structures, and abstract relationships"

(p. 136). And with a sufficiently broad understanding of "pattern," the shorter version

"the science of patterns" says it all (pp. 73-74), suggesting that patterns reveal structure

and relationships. In fact, structure, pattern, and relations are mutually dependent aspects

of mathematical thinking, any of which may be taken as primary. Poincaré, for example,

begins with relations and arrives at structure:

1

> Mathematicians do not study objects, but relations among objects; they are
> indifferent to the replacement of objects by others as long as the relations don't
> change. Matter is not important, only form interests them. (cited in Gallian,
> 1994, p. 102)

And I would argue that structure gives rise to relationships and patterns. Thus, when I speak of a structural view of mathematics, I mean a view that embodies all of these aspects.

Devlin's thesis raises a number of practical questions: What are the implications for the mathematics curriculum? Should all students experience mathematics as abstract patterns, structure, and relationships? If so, how might such ideas be taught? Just when do mathematics students, particularly future mathematicians and secondary teachers, have an opportunity to develop such a perspective about mathematics? And what do students take from such experiences? To what extent is the representational structure of natural language sufficient for reasoning about mathematics? Where does natural language fall short?

With the organization of today's mathematics curriculum, few students ever have an opportunity to develop a structural view of mathematics. Mathematics majors are first exposed to such a view of mathematics in a university course called abstract algebra, typically taken in their junior or senior year. The course usually focuses on elementary group theory and often also includes introductions to ring theory and other abstract structures. It is worth pointing out that examples from group theory form a significant portion of Devlin's description of what mathematics is.

The structural view of mathematics has been an organizing theme in the mathematics research community since the group of mathematicians known collectively as Bourbaki

identified three mother structures: algebraic structure, order structure, topological

structure, although they allow for the possibility of additional fundamental structures (see

Bourbaki, 1950, for an overview). Beginning in 1939, this influential group published a

collection of texts under the title *Éléments de Mathématique*, intended to set mathematics

on a firm footing. In a short expository piece, Bourbaki (1950) simultaneously present a

description of the structural view and an argument for the formal, abstract, axiomatic

method upon which it is based, acknowledging explicitly the difficulty of higher stages of

abstraction and "the great problem of relations between the empirical world and the

mathematical world" (p. 231). And, once again, group theory serves as the canonical

example.

During the 1960s, curriculum developers and some psychologists adopted structure as a

central theme, though not always with the same motivations. Piaget (1970a), for

example, was interested primarily in mental structures, and so structure was a

fundamental characteristic of his psychogenetic theory. He was subsequently taken by

the structures suggested by Bourbaki, such as the analogy between his concept of

reversibility and the algebraic concept of inverse. Ernest (1994) goes so far as to say,

"Piaget was seduced by the Bourbakian account of mathematics as logically constituted

by three mother structures" (p. 2). Bruner (1960/1977), on the other hand, took structure

to be a fundamental characteristic of the disciplines and suggested that structure must be

taught. "The task ... is one of representing the structure of that subject in terms of the

child's way of viewing things" (p. 33). Judging from the movements in the mathematics

curriculum since the "new math," it seems that the structure has largely faded in school

mathematics, but structure has remained at least a dominant background influence in the upper-level undergraduate mathematics curriculum.

Despite this history, little is available in the mathematics education literature about how students learn content, such as group theory, that typifies the structural view. And less is known about the extent to which learning group theory helps students develop a structural view of mathematics. This study aimed to contribute to the empirical and theoretical work in this area of mathematics education by investigating student learning in abstract algebra, or more specifically, group theory. Like Devlin's book, this study was about mathematics, language, and representations, but rather than taking such a global and evolutionary view, it was more exploratory, beginning at the level of individual students in a college mathematics class.

Rationale

The reasons for investigating student learning in group theory are manifold. First, such investigations can contribute to an understanding of advanced mathematical thinking, especially because group theory typifies what modern mathematics is about, as discussed above. Second, students often find the course difficult, and instructors are often dissatisfied with the level of understanding reached by the students. Third, the research in this area is particularly thin. And fourth, because the course is typically required of preservice secondary teachers, there are potential implications for teacher education. These reasons are elaborated below.

Advanced Mathematical Thinking

Tall (1992) suggests that "advanced mathematical thinking ... is characterized by two important components: precise mathematical definitions (including the statement of axioms in axiomatic theories) and logical deductions of theorems based upon them" (p. 495). Over the past decade and a half, the mathematics education community has seen growing interest in the study of advanced mathematical thinking and, simultaneously, in research in the teaching and learning of undergraduate mathematics. Although there was some scholarly work in this area in the 1970s and early 1980s, a community of researchers was formally established with the creation in 1985 of the working group on advanced mathematical thinking within the International Group on the Psychology of Mathematics Education (PME). Since then, accompanying the broader curricular and pedagogical reforms in undergraduate mathematics (Dossey, 1998; Douglas, 1986; National Research Council, 1992; Steen, 1992; Tucker & Leitzel, 1994), scholarly interest in the teaching and learning of undergraduate mathematics has grown and intersected with the broader mathematics community, as evidenced by the increasing numbers of sessions at the Joint Mathematics Meetings devoted to educational issues and particularly by the creation in 1999 of the Association for Research in Undergraduate Mathematics Education (ARUME), which has since become a special interest group of the Mathematical Association of America.

Literature Is Thin

Despite these developments, the research literature in advanced mathematical thinking and undergraduate mathematics education has been and remains sparse, particularly regarding the learning of post-calculus mathematics. This is perhaps a particular

symptom of a general phenomenon that the amount of research literature diminishes sharply as one proceeds from elementary school to secondary school to undergraduate mathematics. One comprehensive survey of the literature in undergraduate mathematics education was conducted in 1995 (Scher & Findell, 1996), at about the time this study was conceived. Based on literature published in journals and known collections (e.g., Kaput & Dubinsky, 1994) between 1985 and 1994, the survey found 312 research articles on the teaching and learning of undergraduate mathematics and categorized them according to mathematical content and research outcome. Of those 312 articles, fewer than 30 could clearly be described as attending to the teaching and learning of content beyond first-year calculus, and only two concerned the learning of abstract algebra. The research about the teaching and learning of undergraduate mathematics has grown since 1994, particularly through the publication of volumes II through IV of *Research in Collegiate Mathematics Education* (Dubinsky, Schoenfeld, & Kaput, 2000; Kaput, Schoenfeld, & Dubinsky, 1996; Schoenfeld, Kaput, & Dubinsky, 1998). And although there is substantial recent work in linear algebra (Dorier, 2000), the literature specific to the teaching and learning of abstract algebra remains thin. A literature search using the same criteria as the previous survey revealed 15 articles on the learning of abstract algebra. Eleven of them had bee published since 1994, of which 9 grew from the work of Dubinsky, Leron, and their collaborators.

Difficulties with Teaching and Learning

Some research has indicated student understanding of the concepts in abstract algebra is less than satisfactory (see, e.g., Dubinsky, Dautermann, Leron, & Zazkis, 1994; Hazzan & Leron, 1996). Leron and Dubinsky (1995) go so far as to declare that the teaching of

abstract algebra is a disaster and to claim that there is wide consensus on this view among both instructors and students. This view may be indicative of a larger problem: The transition to advanced mathematics courses, particularly those beyond calculus, is often problematic.

Harel (1989) proposes several reasons why the learning of linear algebra is difficult for students, which I paraphrase as an initial characterization of the difficulties with abstract algebra. First, the concepts are abstract structures that serve as categories for a broad and diverse range of examples. The objects are defined by their properties, and the properties rather than the examples are primary, making it hard for students to conceive of them. Second, many of the examples themselves are unfamiliar to the students. And third, many students are not yet comfortable with proof and the axiomatic method. Regarding the last point, it is worth mentioning that linear algebra is often studied before abstract algebra. But in some mathematics programs, the approach to linear algebra is fairly concrete, unlike the abstract approach Harel describes. Furthermore, some mathematics programs require that students take a course in "mathematical proof," before they take abstract algebra. Even with such experiences, there is reason to believe that students have not yet transcended the difficulties with proof (see, e.g., Moore, 1994).

Abstract Algebra for Future Teachers

There is widespread agreement on the need for improvements in teacher preparation and professional development in mathematics, as evidenced in the plethora of recent reports that discuss teacher education. The reports recommend ways to improve the system of teacher education (Kilpatrick, Swafford, & Findell, 2001; National Commission on Mathematics and Science Teaching for the 21st Century, 2000; National Research

Council, 2001a), recommend mathematics that should be required of future teachers (Conference Board of the Mathematical Sciences, 2001), and reframe questions about the content and delivery of mathematics teacher education (National Research Council, 2001b), yet there has been little empirical or theoretical work exploring the relevance of particular mathematics courses in the preparation of future teachers. Most certification programs for prospective secondary mathematics teachers require a course in abstract algebra. Thus, by exploring what students do learn in an abstract algebra course, this study provides some empirical and theoretical backing for ways to implement and improve upon the recommendations.

For some time, professional organizations and committees have agreed that the study of abstract algebraic structures is an important part of a secondary preservice teacher's mathematical preparation (see, e.g., Leitzel, 1991; National Council of Teachers of Mathematics [NCTM], 1991; Committee on the Undergraduate Program in Mathematics, 1971). Although these reports provide little in the way of rationale, the dominant point of view is that the equivalent of a major in mathematics should be required of prospective high school teachers (Ferrini-Mundy & Findell, 2001).

The implicit rationale might be that a major in mathematics is necessary in order to understand secondary school mathematics with sufficient depth. And, as elaborated below, powerful ideas from advanced mathematics can explain and unite ideas from school mathematics. A recent report on the mathematical education of teachers (Conference Board of the Mathematical Sciences, 2001) acknowledges, however, that "unfortunately, too many prospective high school teachers fail to understand connections between [abstract algebra and number theory] and the topics of school algebra" (p. 40).

Although the empirical basis for this claim is not stated, there is clearly a perceived need to think about ways to improve the content and effectiveness of the courses that are offered to future teachers. Furthermore, there is a need to think deeply about the rationale for requiring of future teachers a course in abstract algebra, and this study provides some suggestions there.

What Is Abstract Algebra?

> The notion of a "group," viewed only 30 years ago as the epitome of sophistication, is today one of the mathematical concepts most widely used in physics, chemistry, biochemistry, and mathematics itself. (Sosinsky, 1991, cited in Gallian, 1994, p. 68)

School algebra can be seen as a generalization of arithmetic in which the variables are numbers and the expressions and equations are formed with the four arithmetic operations. Abstract algebra is a generalization of school algebra in which the variables can represent various mathematical objects, including numbers, vectors, matrices, functions, transformations, and permutations, and in which the expressions and equations are formed through operations that make sense for the particular objects: addition and multiplication for matrices, composition for functions, and so on. This section provides a short sketch of abstract algebra in order to highlight ideas of structure and to present the terms, concepts, notations, and perspectives that undergird the research questions and subsequent analysis.

Abstract algebra consists of axiomatic theories that provide opportunities to consider many different mathematical systems as being special cases of the same abstract structure. The theories are called *axiomatic* because the structures are defined by axioms.

Group theory is "one of the oldest (and also one of the simplest) of axiomatic theories"

(Bourbaki, 1950, p. 224).

Consider, for example, the following four mathematical systems:

1. The integers $\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ under the operation of addition. This system is denoted Z.

2. The whole numbers less than a given whole number n, $\{0, 1, 2, \ldots, n-1\}$, under the operation of addition, where addition is given by the remainder after dividing the usual sum by n. This system is denoted Z_n.

3. The translations of the plane, where the operation is given by composition, that is, following one translation by another.

4. The set of 2×2 matrices of real numbers with determinant 1, under matrix multiplication.

Each of these examples consists of a set of elements (numbers or translations) together

with an operation that specifies how to combine two of the elements to get an element

that is also in the set. Because the operation combines two elements, it is often called a

binary operation. In order to talk about these examples simultaneously, the operation is

denoted by *, where the interpretations are addition, addition "modulo n," composition,

and matrix multiplication, respectively, in the four examples.

With some work, it is possible to see that each of these systems satisfies the following

axioms:

1. *Associativity*. For any three elements, x, y, and z, in the set (not necessarily distinct), $(x*y)*z = x*(y*z)$.

2. *Identity*. There is an element, e, in the set, such that for any x in the set, $e*x = x = x*e$. (For addition of integers, the identity is 0; for addition modulo n, the identity is 0; for translations of the plane, it is the "identity" translation that leaves every point fixed; for matrices under multiplication, it is the "identity" matrix with 1s on the diagonal and 0s elsewhere.)

3. *Inverse*. For each element x in the set, there is an element y in the set such that $x*y = e = y*x$.

A fourth (or zero[th]) axiom, *closure*, is built into the requirements of a binary operation: that the combination of two elements gives an element that still lies in the set. It should be pointed out that commutativity is not one of the axioms, and it is not hard to see that matrix multiplication is not commutative.

Any set and operation that together satisfy these axioms is said to be a *group*. When the operation is also commutative, the group is said to be *Abelian*. The advantage of the axiomatic approach is that any result (i.e., theorem) that can be proved on the basis of the axioms alone necessarily applies to all four examples and also to any other mathematical system that satisfies the axioms.

The important results in group theory depend upon a collection of related concepts. A *subgroup*, for example, is a subset of a group, which is itself a group under the group's operation. The role of structure again returns to the fore with the concept of *isomorphism*. On a high level, the group axioms define an algebraic structure that applies to a broad collection of mathematical systems. The axioms create the rudimentary structure to which all groups must conform. At a lower level, every specific group is a mathematical system with its own internal structure. An important abstraction can occur when two groups appear in different settings and yet are "essentially the same." The intuitive idea is that two groups are structurally the same, or *isomorphic*, if they differ only in the names of their elements and operation. Demonstrating that two groups are isomorphic requires finding a renaming that preserves the group operation. Such a renaming, which is essentially a function that takes elements from one group to the other, is called an isomorphism.

It should be pointed out that the above mathematical systems and other standard examples may not be familiar to undergraduates in a first course in abstract algebra. Thus, some of the student's energy must be spent trying to build some familiarity with the examples. Taken together, these examples and the concepts of group, subgroup, and isomorphism constitute the fundamental concepts of group theory for the purpose of this study.

I distinguish as "advanced concepts of group theory" those concepts that require the construction of new objects. Given a subgroup *H*, one can create a *left coset* of the subgroup by multiplying an element *a* of the group on the left by each of the elements in the subgroup. The coset is denoted *aH*. When the set of left cosets forms a group by extending the group operation to the cosets, the resulting group is called a *quotient group*, and the subgroup that gave rise to the cosets is said to be *normal*.

Other important mathematical structures are *rings* and *fields*. In ring theory, there are two operations, typically called multiplication and addition. Examples are the arithmetic of integers, of matrices, and of polynomials in one variable with integer coefficients. A field is essentially a ring in which multiplication is commutative and division is also possible, except, of course, division by zero. Examples are the rational numbers, the complex numbers, and the integers modulo *p*, where *p* is prime.

The Big Ideas of Abstract Algebra

A course in abstract algebra is the place where students might extract common features from the many mathematical systems that they have used in previous mathematics courses, such as calculus, linear algebra, and school algebra. Students have opportunities to develop deeper understandings of concepts such as identity, inverse, equivalence, and function. What is shared, for example, by the identity for multiplication of real numbers,

the identity matrix, and the identity function? What is the common idea behind the inverse of a function, the inverse of a matrix, and the multiplicative inverse of a number? In abstract algebra, students can also learn about the importance of precise language in mathematics and about the role of definitions in supporting such precision. Mathematics is also about noticing when things are the same and being able to describe how they are different. In abstract algebra, this naïve notion of "sameness" becomes formalized in the concept of isomorphism.

Thus, it is clear that the concepts in abstract algebra provide guiding themes, principles, and sensibilities that pervade mathematics. It is not so clear, however, what sequence of topics from abstract algebra can be constructed to help students recognize and appreciate such themes. And, in particular, it is not clear whether an abstract algebra course intended for mathematics majors, as it is typically taught, can serve such a role.

When the population of students in an abstract algebra course includes future teachers (which may be almost always), these big ideas, such as inverse and identity, are particularly important because they can help teachers connect advanced mathematics with high school mathematics in ways that can strengthen and deepen their understandings of the mathematics they will teach. Of course, it is also crucial that future teachers are able to employ those new understandings in their teaching, but that concern takes us beyond the scope of this study.

Conceptualizing the Study

In the previous sections, I provide a rationale for an investigation of student learning in abstract algebra and a short description of what abstract algebra is. The remainder of this

chapter describes how I arrived at this particular study and presents the research questions.

My interest in the teaching and learning of abstract algebra stems from my own experiences as a student and as an instructor. I found the subject quite difficult myself, both as an undergraduate and as a graduate student in mathematics. When I first taught abstract algebra to undergraduate mathematics majors at a state college, my hope was to provide more conceptual and concrete support for the students than I had been given. Upon beginning my graduate program in mathematics education, I imagined several possible thesis topics, but foremost among these was learning in abstract algebra. In particular, I was interested in exploring students' conceptual understandings.

Some of the literature on the learning of advanced undergraduate mathematics focuses on students' difficulties writing proofs (e.g., Moore, 1994; Hart, 1994). While this literature confirms that structuring, organizing, and writing proofs presents significant difficulties for many undergraduates, there are also significant obstacles in the concepts themselves (Dubinsky et al., 1994).

As I began to conceptualize this study, I had an opportunity, as part of a graduate course in mathematics education, to interview an undergraduate abstract algebra student on several abstract algebra tasks. That experience and subsequent pilot activity not only served to develop my interviewing skills but also confirmed that students' conceptual understandings in abstract algebra was a researchable area in the sense that the subtleties in students' thinking seemed interesting and worth exploring.

Research Questions

In investigating students' understandings of advanced mathematics, my intent was to begin building a theory: a representation of student's understandings, or, alternatively, an understanding of students' representations. The central theoretical construct for this study was the notion of a concept image (Tall & Vinner, 1981), which denotes the entire cognitive structure associated with a concept, including examples, representations, processes, and the relationships among them. The concept image is distinguished from a concept definition, which is a form of words used to specify a concept, and which I take to be part of the concept image. It is helpful to imagine a concept image as a network in which the links indicate relationships between ideas. The metaphor of a conceptual network accommodates the perspective that new knowledge builds on prior understandings, and so I investigated not only students' understandings in group theory but also how preliminary mathematical understandings were involved in students' learning.

My interest in characterizing students' understandings led ultimately to the following research questions:

- What are the prominent characteristics and components of students' concept images as they are learning the fundamental ideas of group, subgroup, and isomorphism?
- What are the prominent characteristics and components of students' concept images as they are learning the more advanced ideas of homomorphism, coset, and quotient group?
- How do students' understandings of prior mathematics come into play as they are learning elementary group theory?

The context for the study was a nontraditional class in which the instruction was based largely on problem sets that the students completed in collaborative groups of three or four students. In such a setting, and without a comparison group, it was not possible to

determine the causes of many of the events. The goal of this study was not comparison, however, but rather to begin building a theory supported by a thick description of the issues that students grappled with around the mathematical content of elementary group theory while they were in the process of learning.

The following chapter reviews the relevant literature on the teaching and learning of abstract algebra. Chapter 3 sets forth the conceptual and analytical framework that guided this study. Chapter 4 describes the context and methodology. Chapters 5 through 7 address the research questions in turn. And chapter 8 provides conclusions and implications.

CHAPTER II

LITERATURE REVIEW

To synthesize the research on the teaching and learning of abstract algebra, it is useful to

consider first two categories: those articles connected with Dubinsky's framework for

research and curriculum development (Asiala et al., 1996) and those that are not. These

sections are followed by a brief discussion of research on the learning of proof. To

complement the educational research, I include discussion of historical literature

describing the genesis and evolution of algebra and also some of the literature that

provides suggestions for curriculum or instruction. Much of this literature takes a

negative tone, describing difficulties, errors, obstacles, and the ways in which student

understanding falls short of expert understanding. Clearly, the field could benefit from

an approach that begins organically, with students' ways of thinking.

Dubinsky's Framework

The work of Dubinsky and his colleagues is based on a well-articulated framework for

research and curriculum development in undergraduate mathematics education (Asiala et

al., 1996), which grows largely from Dubinsky's (1991) elaboration of Piaget's notion of

reflective abstraction. The core of framework is the theoretical perspective that all

mathematical conceptions can be understood as actions, processes, objects, or schemas

(hence the acronym APOS). The categories may be seen as an extension of the

process/object distinction that is well developed in the literature and that is discussed in

17

detail in chapter 3. It is important to keep in mind that the theoretical perspective provides ways to categorize students' thinking about mathematical concepts, not categories for the concepts themselves. Thus, one student may have an action conception of coset and another a process conception. The categories are roughly developmental, with each new conception requiring new mental constructions.

According to Dubinsky's theory, an *action* conception is different from a *process* conception in that in the former, the student is particularly focused on going through specific procedural steps and is unable to talk clearly about one of the steps until all the previous steps have been carried out. An action conception can become a process conception through a mental construction called *interiorization*. Then, the student can think about the result of the process without actually having done it and, in particular, can imagine reversing the process. A student who has an *object* conception of a mathematical idea can imagine it as a totality and, in particular, can act on it with higher-level actions or processes. Processes can be *encapsulated* into objects, and it is sometimes useful that the student be able to *de-encapsulate* an object to focus on the underlying process. *Schemas* are coordinated collections of actions, processes, objects, and other schemas, which can themselves be encapsulated into objects.

Dubinsky's research and development framework consists of three activities: theoretical analysis, design and implementation of instruction, and observation and evaluation of the implemented instruction. The theoretical analysis describes the actions, processes, objects, and schemas that students might construct in order to develop an understanding of the target concepts. Instructional activities are designed specifically to help students make the constructions identified in the theoretical analysis and typically include

computer activities using the programming language ISETL (Interactive Set Language), whose syntax closely resembles mathematical notation. Evaluation consists largely of interviews and written exams to determine to what extent students made the desired constructions. The framework is cyclical in that observation and evaluation inform revisions in the theoretical perspective, which informs subsequent instructional design, and so on. The research papers primarily report the results of a particular implementation, focusing primarily on characterizing the action, process, and object conceptions of students, reporting the numbers of students in each category, and, sometimes, comparing results with classes that had received traditional instruction.

On the learning of abstract algebra, the evaluation of the first round of curriculum development is reported in a research article (Dubinsky et al., 1994) and the resulting second version of the curriculum has been published (Dubinsky & Leron, 1994). Dubinsky et al. conclude, not surprisingly, that many of the concepts, especially coset and quotient group, seem quite difficult for students, and they offer some explanations. They discuss a number of cognitive obstacles that are common among beginning abstract algebra students. Regarding the group concept, the idea of an abstract binary operation poses a significant obstacle for students, who often think of a group as a set and ignore the operation. Students are often unable to correctly answer questions about cosets in and quotients of noncyclic groups, and they often confuse normality with commutativity. Although some of the students can perform the calculations required for listing the elements in a coset, they have difficulty thinking of cosets as objects that can themselves be manipulated. It may seem obvious that a set is an object, but sets are often described by a process that lists all elements or that would eventually list any element. In this way

a set is a process. A set is not a full-fledged object in the mind of the student until the student can imagine a set as an element of another set. The researchers isolate certain prerequisites for success in abstract algebra, including understanding of functions as both processes and objects.

This research has been criticized by Burn (1996), who characterizes Dubinsky et al. (1994) as a report of a novel teaching procedure using the computer and particular activities. He suggests that the fundamental concepts of group theory may be not group, subgroup, coset, and normality, but rather closure, associativity, identity, inverse, function, and set. Burn further points out that some of the interview excerpts that were regarded as misconceptions may actually reveal insight on the part of the student (e.g., closure is enough to determine whether a subset of a finite group is a subgroup). Furthermore, quotient groups are quite easy to see in some situations (e.g., even and odd integers, rotation and reflection in the transformations of the plane). It should not be surprising, Burn suggests, that the concepts in abstract algebra can be described in the language of sets and functions, but that may be twentieth century analysis imposed on nineteenth century ideas. (I would point out that in order to implement the concepts in ISETL, it is necessary to view them as sets and functions.) Finally, he proposes that automorphisms (specifically permutations and symmetries) may be more profitably viewed as the fundamental concepts of group theory.

Dubinsky et al. (1997) respond by reaffirming that their previous article is not a report of a novel teaching procedure but an attempt to contribute to knowledge of how students understand certain concepts in group theory. Regarding Burns' unsupported claims about

the ease with which students might understand certain instances of quotient group or permutation, they suggest that Burns conduct a study and report on it.

The second iteration of research and curriculum development using the APOS framework to study the learning of abstract algebra is reported in a collection of articles (Asiala, Brown, Kleiman, & Mathews, 1998; Asiala, Dubinsky, Mathews, Morics, & Oktac, 1997; Brown, DeVries, Dubinsky, & Thomas, 1997; see Clark et al., 1997, for an overview). The general conclusion of these articles is that the authors' initial epistemological analyses of the various topics are supported by the data, in the sense that the analyses describe the important processes, objects, and schemas that students need to construct in order to learn the those topics. The authors then typically offer refinements of the epistemological analyses and later offer pedagogical suggestions. Some specific conclusions include the suggestion that the crucial idea in calculating a quotient group may be constructing the binary operation, the importance of being able to choose appropriately between two binary operations defined on a set (e.g., multiplication and addition), and specific misconceptions such as the fact that some students believe Z_n is a subgroup of Z.

Student Thinking

Although the literature on the learning of abstract algebra contains a small number of research articles, the list of misconceptions identified is not short. Selden and Selden (1978) alone list thirteen types of errors, many of which might occur in any undergraduate mathematics course. Some commonly found misconceptions include confusion about the group operation, particularly when the problem involves more than

one group (Hart, 1994; Selden & Selden, 1978), believing that the only solution to the equation $x = x^{-1}$ was the identity element (Hazzan, 1994), using techniques from the real numbers in the abstract setting (Selden & Selden, 1978; Hazzan, 1994, 1999), confusing a theorem and its converse (Selden & Selden, 1978; Hazzan, 1994; Hazzan & Leron, 1996), and difficulty managing the distinction between set and element (Hazzan, 1999; Selden & Selden, 1978). This last distinction is further complicated by the fact that the elements of the quotient group are themselves sets.

Some of the above misconceptions are tied to the use of mathematical notation. Selden and Selden (1978) found, for example, that students often use the same symbol for two different things, and, conversely, they often assume things are distinct because they have different names. Hazzan (1994) suggests, regarding the use of different letters in the axiom for inverses, that it is easier to think of a relation between two different objects than of an object with itself.

Other difficulties seem to be tied to other sorts of representations. As part of a study on visual and analytic thinking, Zazkis and Dubinsky (1996) investigated abstract algebra students' ability to represent the elements of D_4, the group of symmetries of the square and then to find the product of two elements. This task can be approached either "visually," using a geometric representation, or "analytically," using permutation representations. They found that most students used a combination of these approaches, suggesting that the dichotomy between visual and analytic thinkers may be false. They propose an alternative model that assumes visual and analytic thinking to be mutually dependent in mathematical problem solving.

The study also produced some unexpected mathematical results (Zazkis & Dubinsky, 1996). Eight of ten students found as they translated between the geometric and permutation representations that the groups were not isomorphic, causing Zazkis and Dubinsky to conclude that the dihedral groups such as D_4 are not groups until some structure is imposed on them in the sense that the relationship between the group operations in the two representations must be specified appropriately. By careful analysis of the ways to translate from the geometric to the permutation representation, they found that students could focus on the square and where its vertices traveled (an *object* interpretation) or on the four positions and which vertices they contained after the transformation (a *position* interpretation). In computing the product of two transformation symmetries in the geometric representation, students could imagine either that the axes describing the transformations were fixed (a *global* interpretation) or that they traveled with the square (a *local* interpretation). Choosing either the object/global or the position/local pair of interpretations results in the desired isomorphism between the geometric and permutation representations. Most students, however, were drawn to the position/global pair or the object/local pair, which caused the groups to be anti-isomorphic, in the sense that the order of multiplication is reversed. Zazkis and Dubinsky suggest that the embedding of dihedral groups in permutation groups deserves some careful attention in instruction.

Hannah (2000) pursued Zazkis and Dubinsky's ideas through a teaching experiment. Expecting that students would prefer the global interpretation, he encouraged the object interpretation by using additional labels to separate the object from the position. About half the students still preferred the position interpretation, although one of these students

also chose the local interpretation, thus leading to an isomorphism between the geometric and permutation representations. In the second trial of the teaching experiment, after making some additional adjustments in his notation to make the object and frame of reference more salient, all but one student chose consistent interpretations. Hannah also found that permutation notation can lead to the same local/global interpretational issues.

Leron, Hazzan, and Zazkis (1995) discuss the development of the concept of group isomorphism. Some of the difficulties, they suggest, may actually be with quantification. They note that the naïve concept of isomorphism (same group with the names changed) is a good start, but the object isomorphism is defined directionally, with the two groups playing different roles, and requires a sophisticated concept of function. In other words, although there is symmetry in the statement that two groups are isomorphic, actually finding an isomorphism requires choosing one group as the source (the domain of a function) and the other as the target. When trying to construct an isomorphism between two groups, they note that students hope for a canonical (or at least obvious) isomorphism and get stuck when there is a choice.

Hazzan and Leron (1996) argue that the standard formulation of Lagrange's theorem hides its nature and its deep meaning. The standard formulation is:

> Let G be a finite group. If H is a subgroup of G, then $o(H)$ divides $o(G)$.

The notation $o(G)$ signifies the *order* of the group,[1] that is, the number of elements in it. The authors suggest that the contrapositive of the theorem includes explicit quantifiers that make its nature as a nonexistence theorem clearer and reveal its deep meaning:

---

[1] This is Hazzan and Leron's notation. In the class that provided the context for the present study, we used the alternative notation $|G|$ to denote the *order* of a group G.

> **Contrapositive**: If k does not divide $o(G)$, then there does not exist a subgroup of order k.
>
> **Deep meaning**: If there exists a subgroup of order k, then k divides $o(G)$.

This analysis of Lagrange's theorem arose in response to data collected on the question, "Is Z_3 a subgroup of Z_6?" Out of 113 students, 73 gave incorrect answers. Of these, 20 gave some version of, "Yes, by Lagrange's theorem, because 3 divides 6." Hazzan and Leron suggest that students' response may be due to a coping mechanism and may not really reflect thinking about the theorem and the two groups. The authors explore issues of coping more fully elsewhere (Leron & Hazzan, 1997).

In a broader study of learning in abstract algebra, Hazzan (1999) found that students tended to reduce the abstraction level in order to cope with the task at hand. She organized her results according to three ways of looking at the level of abstraction. Regarding abstraction level as the quality of the relationships between the object of thought and the thinking person, she found that students tend to make the unfamiliar familiar by basing their argument, for example, on numbers and number operations. Regarding abstraction level as a reflection of the process/object duality, she found that students tend to personalize formal expressions and logical arguments by using first-person language and that they tend to engage a well-rehearsed procedure rather than rely on theoretical knowledge. Regarding abstraction level as the degree of complexity of the concept of thought, she found that students sometimes reduce abstraction level by replacing a set with one of its elements.

Hirsh (1981) describes an abstract algebra course for preservice secondary school teachers that included a "didactical shadow" seminar in which the mathematical concepts were followed closely by readings and discussions on teaching K-12 mathematics. These

readings were intended to encourage preservice teachers to see abstract algebra as relevant in their future as secondary school teachers. The study found significant improvement in the experimental group in the students' understanding of the real number system and nonsignificant improvement in the control group. On several affective measures, no significant differences between groups were found. The most important result, Hirsch suggests, was that the study demonstrated the feasibility of such a course.

Proof

As stated in chapter 1, mathematical proof is one of the defining characteristics of advanced mathematical thinking, and proof plays a central role in the learning of abstract algebra. Because the role of proof did not play a central role in this study, this section briefly reviews literature that was helpful in framing the study.

One of the leading expositors of the role of proof in mathematics education is Gila Hanna (1991, 1995). She suggests that constructivist theories have led to a mistaken view of the teacher as playing a passive role and of proof as being unimportant. She argues for recognizing and promoting proof in the mathematics curriculum as a key tool for promoting understanding. The research on the role of proof in mathematics education is thin and confused by the typical four-year separation between proof in high school geometry and proof in undergraduate mathematics.

Hart (1986, 1994) describes a research study in which twenty-nine college mathematics majors, taking different abstract algebra courses from beginning undergraduate to beginning graduate, were asked to write six standard abstract algebra proofs, each "doable in 15 minutes or less." On the basis of their performance on three criterion

proofs, students were classified into four levels of conceptual understanding. Analysis of errors made, processes used, correctness of proofs, and student assessment of tasks suggested that the journey from novice to expert in a content domain may be an irregular and unstable developmental process, rather than the dichotomy often assumed in the literature.

In a mathematics course called Introduction to Higher Mathematics, Moore (1990, 1994) found seven major sources of student difficulties in learning to do proofs, including inability to state the definitions, inadequate concept images, inability to use the definition to structure a proof, inability or unwillingness to generate examples, and difficulty with mathematical language and notation. He suggests that the concept image/concept definition dichotomy was not sufficient to explain his results and suggested the term *concept usage* to discuss how students used definitions to generate and use examples, applied definitions within proofs, and used definitions to structure proofs. Although in Moore's work this construct more accurately describes students' use of concept definitions rather than of concepts, thinking about concept usage proved helpful in this study in identifying components and characteristics of students' concept images, as described below.

Taken together, these articles support the idea of investigating not only students' understandings of concepts but also their personal definitions of those concepts. Proof, after all, involves reasoning about concepts, which must be meaningful to the students in order to support such reasoning.

History

Sfard (1995) gives a detailed a description of the historical development of algebra with strong connections to the teaching and learning of both school and abstract algebra, providing compelling support for the claim that historical-critical and psychogenetic studies should converge (Piaget & Garcia, 1989, p. 108). According to Sfard, group theory arose out of the work of Lagrange and Ruffini, who noticed that methods of solving polynomial equations depended on permutations of the roots. Soon permutations and then, with Cauchy, operations on those permutations became objects of attention. Galois defined the notion of a group by declaring interest in the structure imposed on the permutations by the so-called substitutions. Cayley freed the concept from any commitment as to the nature of the elements, focusing instead on the manipulations. With the invention of the concept of group, the seeds had been planted for algebra to become a science of abstract structures.

Kleiner (1986) describes four lines of inquiry that coalesced toward the end of the nineteenth century to form the area we now call abstract algebra. First, the techniques from classical algebra for solving polynomial equations led to the permutation groups. Second, questions in number theory led to the finite Abelian groups. Third, attempts to unify and organize geometry led to transformation groups. Finally, roots in analysis led to investigation of continuous transformation groups. One response to this account is to use historically important problems to provide pedagogical and intellectual motivation in the teaching of abstract algebra (see Kleiner, 1995).

Nicholson's (1993) account of the slow historical development of the concept of quotient group can provide additional sources for cognitive roots to be exploited. She suggests

several obstacles that were overcome by the mathematics community during the development of this concept. First, the community needed an abstract concept of group that was not dependent on any particular representation. Second, the community needed the concept of equivalence (modulo a subgroup). Finally (and most importantly), the community needed to realize that the elements of the quotient group are not like the elements of original group, but are equivalence classes—sets. All of these historical developments provide clues about what might be the issues for students learning the subject. In this study, I paid attention in particular to the ways in which students develop an abstract concept of group and the sense in which they consider sets to be elements of quotient group.

Teaching Suggestions

I close the review of the literature with a discussion of articles that informed the development of the course, that provide additional rationale for investigating learning in abstract algebra, and that collectively support the decision to investigate learning in a nontraditional course.

In "An Abstract Algebra Story," Leron and Dubinsky (1995) condense the principles and research behind their textbook (Dubinsky & Leron, 1994) into a dialogue with an "idealized reader." They begin by asserting that *"The teaching of abstract algebra is a disaster, and this remains true almost independently of the quality of the lectures"* (p. 227). They suggest that the ISETL computer activities provide an experiential basis for the abstractions that follow, asserting that "if the students are asked to construct the

group concept on the computer (by programming it), there is a good chance that a parallel construction will occur in their mind" (p. 230).

In Dubinsky and Leron's approach, before seeing the concept of quotient group, the students have already explored a construct they call GmodH, which is the set of cosets in G of a subgroup H, independent of whether H is normal, an approach consistent with that recommended by Benson and Richey (1994). Leron and Dubinsky acknowledge that the notation GmodH is unorthodox, particularly when H is not normal, but defend their approach by noting that students realize they need to look into the properties of H that make GmodH a group and come to appreciate that the main issue is closure. Furthermore,

> by building on the material that the students bring up, the instructor is able to state most naturally and smoothly the definition of a normal subgroup, the theorem that when H is normal then GmodH forms a group, and the (now very easy) proof of this theorem. Normality is naturally introduced here as the condition which insures that GmodH be a group, and the definition most often discovered by students is $aH = Ha$ for all $a \in G$. Except for the new name, the students can really feel that the instructor merely summarizes what they have found in their investigations. In the session that follows, the instructor makes the final ties with the "standard" approach by explaining that when H is normal, GmodH is commonly denoted G/H, and is called the quotient group of G modulo H and coset product is commonly defined by the formula $(Ha)(Hb) = H(ab)$.
> (p. 238)

Freedman (1983) also rejects the lecture method, quoting Halmos, "A good lecture is usually systematic, complete, precise—and dull; it is a bad teaching instrument" (p. 631) and Moise, "It is simplistic to suppose that people remember what they are told and understand the things that are explained to them clearly" (p. 631). He discusses an approach he used in London in which students in a small seminar were each required to read and lecture on some original papers in abstract algebra. Although this approach may

seem quite radical to instructors in the United States, he claims that the students worked together and were quite successful.

In response to the difficulties students usually have with Lagrange's theorem, Johnson (1983) notes that the traditional proof involves cosets and equivalence relations, both of which are new concepts to most students. As Lagrange's theorem is usually used to prove the more intuitive theorem that the order of an element divides the order of the group, Johnson suggests proving the latter result first, for it follows quite naturally from the decomposition of a permutation into disjoint cycles. Of course, this approach assumes the students are familiar with permutation groups, and such an assumption might be unwarranted.

Holton and Wenzel (1993) describe an abstract algebra course in which Lagrange's theorem is preceded by cooperative learning via examples. Rejecting the traditional approach of "exposition, exhortation and regurgitation" (p. 883), they found that students were able to conjecture the theorem and many of the necessary lemmas. Although it was not a formal research study, the description of the classroom environment was compelling.

Conclusion

This review has shown that although there have been few published research studies on the learning of abstract algebra, there is a theoretical and empirical base on which to build. To complement the work embedded in the APOS framework, this study is more exploratory in nature, taking a broader view of the ways of thinking that students exhibit

while trying to make sense of the concepts in the course. The next chapter describes the conceptual and analytical framework designed to support such an approach.

CHAPTER III

CONCEPTUAL AND ANALYTICAL PERSPECTIVE

This chapter sets forth the conceptual perspective that guided this study and that

contributed to the design of the analytical framework. Fundamentally, learning is seen as

a process of making sense of experience and of building understanding, a viewpoint that

is consistent with various forms of constructivism. The central theoretical construct is the

notion of a concept image (Tall & Vinner, 1981). The concept image is contrasted with

the concept's definition, which leads to a discussion of the role of definitions in

mathematics, in thinking, and in learning. The chapter continues with a discussion of

other important constructs that are useful in describing the growth and character of

concept images, particularly in advanced mathematical thinking, including abstraction

and generalization and the distinction between process and object conceptions of

mathematical ideas. The chapter also includes a discussion of the role of metaphor in

mathematical thinking, with particular attention to thinking in abstract algebra. Next, I

discuss issues of naming and notation, setting the stage for a discussion of semiotics,

which provides much of the analytical and theoretical backing for the study. These

various theoretical constructs are then brought together at the end of the chapter in an

analytical framework that undergirded the analysis of the data.

<u>Learning with Understanding</u>

This study was based upon the following fundamental theoretical assumptions that are consistent with a large body of work in cognitive science, psychology, and mathematics education (see, in particular, Bransford, Brown, & Cocking, 1999; Hiebert & Carpenter, 1992). First, human beings are conceptualizers in that they try to make sense of their percepts by developing concepts. People try to understand their experiences by organizing them, abstracting from them, creating categories, making connections, particularly with prior knowledge, and making distinctions. In high school and college mathematics, for example, students create a category called "function" by abstracting the common features of the many mathematical creatures called "function" in their experience. These abstracted features are not necessarily the properties that are isolated in the formal mathematical definition, as is elaborated below.

Second, knowledge is represented internally in the mind. People create internal representations for objects, processes, properties, and relationships; for images, sounds, smells, sensations, and impressions; and also for categories and networks of these. These mental representations do not match an external world but rather fit one's experience with some degree of viability (von Glasersfeld, 1990). Because mental representations are not observable, discussions of how ideas are represented in someone's head must be based largely on inference. Such inferences can be facilitated by building and testing models of individual understanding, as is elaborated below. A fundamental goal of research in the psychology of learning is to understand mental representations of ideas, by building models, describing their features, and so on, based on observation of learning situations. It is not necessary that the models match the underlying neural processes

(Kosslyn & Hatfield, 1984). Rather, the goal is that models fit the observations with some degree of viability, particularly with regard to explanation and prediction. In order to build such models, it is reasonable to assume that the external entity being represented influences and constrains the internal representation. In mathematics, these external entities are often themselves representations, such as symbols or diagrams, because mathematical ideas are accessible only through their representations (Duvall, 1999; see also Pimm, 1995, p. 119).

Third, internal representations can be connected to one another in useful and hence meaningful ways. Successful learning may be described as learning with understanding, where understanding is characterized by connectivity. While in the process of learning, connections are made internally in the mind of the learner and over time the concepts, processes, properties, examples, and the connections among them grow to form cognitive structures that might be described as networks. In general, the more connections, the more intricate and encompassing are the networks, and the deeper are the understandings. In this study, individual conceptual understandings are described via the term *concept image*, which denotes the entire cognitive structure that a particular individual associates with a particular concept, as elaborated below. In considering the notion of a concept image, it is important to contemplate not only a concept's structure and connections to other concepts but also the boundaries that distinguish the concept from related ideas.

Concept images and other cognitive structures are actively built up over time through experience and through active reflection on that experience. The structures, of course, depend heavily upon prior experience and also upon the nature and extent of the reflection. Thus, in response to an experience, the actual constructions are personal and

idiosyncratic, which implies that learning and knowing, too, are personal. It is for these reasons that phrases such as "construct personal meaning" or "construct knowledge" are helpful in describing the learning process. This is not to say that all conceptual structures are equal. Some conceptual structures are particularly weak or fragile or lack long-term viability in light of future learning goals. Other structures are strong and persistent. Some conceptual structures are unproductive and fade as a result. Other structures are productive and will support and promote future learning. And, of course, when measured against established knowledge, sometimes conceptual structures contain ideas that are incorrect.

The real quandary lies with strong, productive, but faulty structures with incorrect ideas—often called misconceptions. Independent of whether these are called knowledge, such structures are personal conceptions that are held with conviction and are based upon some reasoning, however incomplete or fallacious.

Piaget describes two mechanisms by which a subject makes sense of experience: *assimilation* and *accommodation* (see, e.g., Steffe & Wiegel, 1996). When an experience fits within the existing cognitive structures, the experience has been assimilated. If, on the other hand, the experience evokes cognitive structures that do not fit with the experience, we say the learner has been *disequilibrated*. To re-equilibrate, the learner must reorganize his or her cognitive structures in light of the new experience. It is this reorganization that Piaget calls accommodation.

The point is that new information is not simply received but is actively interpreted and filtered through prior experience. The experience must either make sense within the existing structure, in which case it is assimilated, or it must be "moderately novel" so that

the experience creates a disequilibration, which can lead to an accommodation. The experience must be only moderately novel, for it must be sufficiently interpretable to create some cognitive conflict.

This balance between assimilation and accommodation makes the point that learning is sometimes difficult, and thus faultless communication is fiction. Papert suggests that, "Anything is easy if you can assimilate it to your collection of models. If you can't, anything can be painfully difficult.... What an individual can learn and how he learns it depends on what models he has available" (cited in Steffe, 1990, p. 173).

Given the above positions about the nature of learning, what then are the implications for the teaching of mathematics? First, mathematics itself is a highly structured and organized domain. For mathematical knowledge to be useable (or perhaps even accessible), it must be organized in some way in the mind. It is clearly not possible to transmit whole structures from the mind of the instructor to the mind of the student. Rather, the student must do some constructing in his or her own mind. Second, it is impossible to know in advance what a person will learn from a given activity. Moreover, it is impossible to know with certainty what a person has learned, although an instructor or researcher can develop approximate models by asking questions. Explicit reflection, with the corrective mechanisms of the observations and responses of the teacher and other students, is likely to lead to strong, viable, and productive connections.

Relationship with Constructivism

Many of the above positions are consistent with the assumptions of any of several forms of constructivism.

> What the various forms of constructivism all share is the metaphor of carpentry, architecture, of construction work. This is about the building up of structures from preexisting pieces, possibly specially shaped for the task. The metaphor describes understanding as the building of mental structures, and the term *restructuring*, often used as a synonym for accommodation or conceptual change, contains this metaphor. (Ernest, 1996, p. 335)

A key expositor of constructivism in mathematics education is Ernst von Glasersfeld, who proposes two principles for *radical constructivism*:

> (a) Knowledge is not passively received but actively built up by the cognizing subject;
>
> (b) the function of cognition is adaptive and serves the organization of the experiential world, not the discovery of ontological reality. (von Glasersfeld, 1989, p. 162)

Adopting only the first of these principles is to take a position that is sometimes called "weak constructivism" (Ernest, 1996) or "trivial constructivism" (von Glasersfeld, 1996). As Kilpatrick (1987) and others have noted, the first of these principles is broadly accepted and "almost no mathematics educator alive and writing today claims to believe otherwise" (p. 7). The second principle, on the other hand, is much more controversial. My position is that whether one believes in an objective reality or Platonic ideals or denies both is, in a sense, immaterial because the student's cognitive structures will match neither reality, nor an ideal, nor the teacher's or researcher's cognitive structures but instead will fit each of these with varying degrees of viability. This is a particularly important point regarding the learning of mathematics, since mathematical concepts exist not in the physical world but rather in abstractions from activity in the physical world and in the mind.

In order to understand what constructivism provides, it is important to recognize that the theories arose in part as a response to what was missing from behaviorism, which refused to posit any meaning behind student's actions. Stimulus-response mechanisms were

supposed to explain all behavior. Thus, constructivism was one of many efforts during the twentieth century to insert meaning and understanding into theories of knowledge and learning. But not all behavior is meaningful.

Vinner (1997) describes some behaviors as pseudo-conceptual and pseudo-analytical because they may be taken as an indication that true (i.e., meaningful) conceptual and analytical processes have occurred when in fact the behavior is little more than simple association and imitation based on superficial similarity. For example, a calculus student who immediately responds "$2x$" when hearing "x^2" is not responding meaningfully if the response is merely a verbalized association. In a calculus class, sometimes this simple association will yield the correct answer, and it is impossible to know, without asking further questions, to what extent the student can construct (or resurrect) some meaning for the response. Students are bound to have such associations. Vinner's point is that in mathematics class, students should evaluate their associations consciously and critically, rather than merely verbalizing them in hopes of getting "credit." He argues that such verbalized associations should not be interpreted as indicating misconceptions or anything about a student's cognitive structures, because cognitive structures are not involved.

Part of the reason many students exhibit pseudo-conceptual behavior in mathematics is that they have found such behaviors to be viable in mathematics classes. Many students are successful in mathematics by relying almost exclusively on simple association and imitation, practicing problems that are just like the ones demonstrated in the textbook or by the teacher. Yet the severe filtering effect of high school and college mathematics

suggests that for most students, mere imitation fails at some point, in arithmetic, algebra, geometry, or calculus.

Thus, learning with understanding requires the development of cognitive structures in which the connections are not simple associations but relationships that are rich in meaning.[2] Mathematical learning is rarely effective without such meaning, in the sense that it is unlikely to be durable, flexible, and supportive of future learning. Mathematical learning that is based only on simple associations is not mathematics at all, not to mention that such skills are fragile and essentially useless today.

There are certain meta-cognitive behaviors that may support learning even if they are simple associations. For example, my students learn that in response to their statements I am likely to say, "Okay. Why?" Some of them internalize this behavior and begin to ask the question themselves. Deborah Ball's class learned that she was likely to ask, "Are these all the solutions?" (Ball & Bass, in press; Suzuka, 2001). And many of Polya's (1957) suggestions (Can I think of a similar problem? Can I simplify the problem?) can be seen in a similar light. The list of desirable behaviors also includes many so-called *habits of mind* that describe successful mathematics knowing and learning. Cuoco, Goldenberg, and Mark (1996) provide a compelling list of such habits, suggesting, for example, that students should learn to look for patterns, to watch for things that change, and also to watch for things that do not change.

---

[2] When I asked my 19-month-old daughter, "When will you be two?" she responded, "November." But how much meaning might have been behind her response? It is rather overwhelming how much conceptual knowledge she will need to construct before she will be able to give a detailed account of the meaning adults might take from her response.

Thus, investigations of mathematical understanding must look at behavior, because that is all that is observable, but should also address meaning, which requires probing beneath the simple associations to explore and make inferences about the meaning that students bring to the situation. My goal as a researcher is to understand how meaningful mathematical understanding is built and how meta-cognitive behavior can help.

Concept Image

The assumption that learners build up cognitive structures as they learn mathematics requires some terms to discuss these structures. I borrow a term from Tall and Vinner (1981):

> We shall use the term *concept image* to describe the total cognitive structure that is associated with the concept, which includes all the mental pictures and associated properties and processes. It is built up over the years through experiences of all kinds, changing as the individual meets new stimuli and matures. (p. 152)

In this seminal paper, Tall and Vinner contrast the concept image with the term *concept definition*, which is a verbal description of the concept and about which I say more below. Because only part of a cognitive structure is brought into consciousness during a particular task, the term *evoked concept image* refers to that portion of the concept image that is evoked in response to a given task (Tall & Vinner, 1981).

The ideas of concept image and evoked concept image are consistent with the work of Hart (1994), who found that when students approach mathematical tasks, "strategies are *evoked* [rather than chosen], based on the interaction between the task at hand and the current conceptual schema" (p. 61). Furthermore, he explained his results by suggesting that "processes, metacognition, and misconceptions are actually part of one's conceptual schema" (p. 62).

Characterizing Concept Images

Concept images consist of examples and nonexamples, representations (symbolic, graphical, pictorial, verbal, enactive, iconic, etc.), definitions and alternative characterizations, properties, results, processes and objects, contexts, and impressions from previous experiences. Solving a mathematical problem (or any mathematical activity) involves recalling or reconstructing examples, representations, objects, or processes and establishing connections to other examples, representations, objects, or processes.

Concept images are not monolithic, for the various examples, properties, and representations play different roles. Michener (1978), for example, distinguishes among start-up examples, reference examples, model examples, and counterexamples. Some properties hold for all examples of the concept (e.g., all groups have an identity element). Other properties, on the other hand, are useful for categorizing examples (e.g., some groups are Abelian). For many concepts, there are also lists of key properties for describing examples (e.g., when making computations or deriving results about a group, it is useful to know the group's cardinality, a set of generators, or an alternative representation).

Because individuals are sometimes more able than at other times to make particular connections or to reconstruct particular examples, representations, or processes, concept images not static entities but rather are always in a state of flux as one thinks about a concept, focusing and refocusing one's thought on various aspects of the concept image. Thus, it is useful to consider not only the components of a student's concept image but also the students' concept usage (Moore, 1994), which in turn can provide

characterizations of a concept image. Concept images, for example, can be dominated by particular examples, representations, or ways of thinking. Dubinsky et al. (1994) observed, for example, that some students focused on the process of computing a coset, whereas more successful students were also able to conceive of cosets as objects that themselves could be acted upon by other processes. Thus, concept images can be limiting when they inhibit an individual from making certain constructions or generalizations.

A concept image is built through all previous experiences with the concept. Experiences that are assimilated make sense within the evoked concept images. Experiences that require accommodation, on the other hand, cause structural changes in an individual's concept images such as the construction of a new concept, the creation of new connections to other concepts, or the reorganization of the connections within or among concepts

A key theme that emerged in this study is the complicated relationship between a concept and its name. I make only two points here and provide additional theoretical discussion below. First, a student's concept image might not reasonably be described as a subset of a mathematician's concept image. A student's concept image may instead include misconceptions or may even be of a different character entirely. Second, the notion of concept image presents something like a chicken-and-egg problem: Which comes first, the concept or the name? One might begin with the name of a concept and then gradually build experiences underneath. On the other hand, as individuals gain experience, they build mental structures that are not necessarily part of a named concept but at some point subsume those structures (and experiences) under a single name. In either scenario, at

what point is there a concept image? The resolution of the problem lies in the realization that the notion of concept image is merely an analytical tool. People do not have concept images in the same sense that they have internal organs. Thus, in the analysis I try to maintain a notion of concept image that is flexible enough to accommodate all of these possibilities.

Relationship to Schema

The term *concept image* shares some similarity with the term *schema*, used by Piaget and many researchers in both the constructivist and cognitive science traditions (see, e.g., Bransford et al., 1999; Piaget, 1970a). In the problem-solving literature, particularly in cognitive psychology, schemas are associated with problem types, and each schema has "slots" that are filled by the specific information provided in the problem. (For an overview of this literature see Mayer, 1992.) This view is problematic because it seems to suggest that learning consists of constructing a new schema for each new problem type.

For Dubinsky (1991), "A schema is a more or less coherent collection of objects and processes" (p. 102), which typically might be named as a concept. For example, "The concept of group can be understood as a schema that consists of three schemas: set, binary operation, and axiom (Brown et al., 1997, p. 192). For Skemp (1987), on the other hand, a schema is a suitably connected collection of concepts.

For the purposes of this study, I was primarily concerned with the ways that students think about particular concepts. Thus, a *concept image* was associated with a particular concept, typically given by name. And with the term, I considered both the way it is structured and the ways it connects to other mathematical ideas.

Concept Definition

As described above, Tall and Vinner (1981) introduced the term *concept image* to contrast with a *concept definition*, which is a form of words used to specify a concept. This distinction serves as a reminder of two simple ideas about students' learning of mathematics. First, around any (mathematical) concept, students' thinking is strongly influenced by the examples, nonexamples, representations, and contexts in which they have previously experienced the concept. Second, students do not typically employ (or naturally adopt) the mathematical habit of consulting a formal definition in response to mathematical tasks but rather rely entirely on their concept image. Furthermore, Vinner (1992) found that even when students can recall a concept definition, the concept definition and the concept image might conflict or contradict one another. He calls this phenomenon *compartmentalization*, suggesting that the concept definition and the concept image are not evoked at the same time.

Perhaps because of the phenomenon of compartmentalization, Vinner and Tall often separate the concept definition from the concept image, in describing cognitive structure (see, e.g., Vinner, 1992), and even go so far as to discuss a "concept definition image" to describe a concept image built up around the definition (Tall & Vinner, 1981). For successful mathematicians, however, a formal concept definition constitutes an integral part of the cognitive structure built around that concept. The definition is routinely consulted and is well integrated into the rest of the concept image. Thus, for this study, a concept definition (personal, formal, or otherwise; see below) was considered a subset of a concept image. In the analysis, I explored the definitions that the students provided as a means of making inferences about the nature and connectivity of their concept images.

Theoretically, including the definition as part of the concept image is reasonable even when the definition is compartmentalized, because the term concept image implies nothing about the connectivity of that structure. In fact, an individual's concept image may include several essentially separate components, each evoked for different kinds of problems.

Definitions are not easily remembered verbatim. And in everyday life, a definition's precise wording is often forgotten shortly after it is used, introduced, or consulted. When terms are introduced via a definition, the definition sometimes provides only scaffolding: When the construction is sufficiently complete, the scaffolding is taken away. To overcome this tendency, some instructors, in mathematics as well as other subjects, recommend that students memorize definitions. But it is not at all clear to what extent mathematicians or other experts recall rather than reconstruct definitions that they use in their professional work.

Because definitions are not easily remembered, it seems likely that they are constructed, and this is the point of view that informed this study. According to Tall and Vinner (1981), a student, when asked to define a concept, may respond with a *personal concept definition*, which may not agree with a mathematically acceptable *formal concept definition* but which instead might be described as an ad hoc description of his or her concept image. Thus, some parts of the concept image function as definitions. For example, in Deborah Ball's third grade classroom, Cassandra shows that six is even by pointing to the number line: "Six can't be an odd number because this is (*she points to the number line, starting with zero*) even, odd, even, odd, even, odd, even" (Ball & Bass, 2000, p. 213). For her, the alternating pattern provides the definitions of even and odd.

For other students in the class, grouping by twos serves to provide the definitions. Still other students prefer to separate numbers into two groups. As another example, the literature on the learning of functions is replete with examples of students defining *function* as synonymous with *formula* or *equation* (see, e.g., Vinner, 1992; Ferrini-Mundy & Graham, 1994). Students implicitly use personal concept definitions when asked to determine whether a particular thing is an example or a nonexample of a concept. This is reasonable behavior in contexts—including many mathematical contexts—where precise definitions are not necessary for the task at hand, particularly when one's concept image is sufficiently rich and robust.

Lakoff and Johnson (1980, pp. 117-125) point out that from a cognitive point of view, definition is not a matter of giving a list of necessary and sufficient properties for a concept, although this is sometimes possible. Instead, concepts are defined by prototypes and by types of relations to the prototypes, and there need be no fixed core of properties of the prototypes that are shared by all instances of the concept. Furthermore, some properties of a concept are not part of the thing itself but are functional, purposive, or otherwise involve interaction with an instance of the concept. And finally, concepts are not fixed but can be systematically modified by metaphors and by hedges such as "technically" or "loosely speaking."

In advanced mathematics, on the other hand, the definition of a concept becomes primary; the definition becomes the touchstone whose role is to ensure rigor (i.e., precision and consistency) within, between, and among concept images. Because this perspective on definitions is unusual outside of mathematics and the hard sciences, it represents a significant adjustment for students. The nature and role of definitions in

mathematics did not play an explicit role in the course that is the subject of this study, but because these ideas inform my analysis, the topic deserves more attention here.

Types of Definitions

Linguists distinguish among several types of definitions (see Kemerling, 2001). A *lexical definition* is an attempt to describe the meaning of a word as it is commonly used. These are the kinds of definitions found in dictionaries, which, contrary to some beliefs, portray current usage not timeless truths, in full acknowledgement that languages evolve. A *stipulative definition*, on the other hand, specifies what a term is to denote. Such definitions are commonly found in technical, legal, and scholarly writing. From the viewpoint of some writers, a stipulative definition freely assigns meaning to a new term and thus is intended to be the touchstone for all subsequent uses of the term. Nonetheless, the expositor is somewhat constrained by what the reader might be willing to accept. Thus, one common approach is to use a *precising definition*, which begins with a lexical definition of a term, and then proposes to sharpen it by stipulating more narrow limits on its use.

Theoretical definitions are stipulative definitions made within the context of a broader intellectual framework. It is worth noting that the validity of a theory depends upon the definitions on which the theory is built. Thus, an appropriate interpretation of Newton's laws of motion, for example, depends upon imposing particular definitions of terms such as mass, inertia, and force onto experience. For example, I presume that separating the concept of weight into mass and acceleration due to gravity was a major conceptual advance. When they were introduced, Newton's particular set of definitions provided an extremely elegant description of objects in motion. But one should recognize that the

precise definitions were required for the creation of the theory. It seems backward to teach students the theory only to conclude later that they have misconceptions about some of the terms. Why not instead try to provide them with experiences that will help them see the importance of precise definitions and the usefulness of particular definitions and of the distinctions among them?

Formal definitions in mathematics are, in a way, peculiar precising definitions—peculiar because of the inflexible totality of the implied precision (i.e., no more, no less) and because the formal definition sometimes bears little relationship to the term's informal usage (e.g., "cycle"). This use of definitions may be peculiar to mathematics and the hard sciences. In the social sciences, precise definitions are hard to find. Rather, an idea is given a name (often a common word), and then the researcher spends paragraphs describing what does and does not fit under the name.

In the analysis of the data in this study, I followed Vinner (1976) and restricted my attention to formal and lexical definitions to discuss the two primary roles that definitions play in mathematics and mathematics learning, but it is worth pointing out that in the above discussion I have presented stipulative definitions of several terms including *concept image* and *stipulative definition* itself. None of these, however, carries the precision of formal mathematical definitions. I believe that such precision is not possible because ideas about language and cognition are messy, fuzzy, and dependent upon the phenomena that the definitions are intended to describe. Mathematical ideas, on the other hand, are ideal—abstracted from phenomena and no longer dependent on the "real world," at least in formal mathematical practice.

Definitions in Mathematical Practice

> "When I use a word," Humpty Dumpty said, in a rather scornful tone, "it means just what I choose it to mean—neither more nor less." (Lewis Carroll, *Through the Looking Glass*)

In the words of Polya (1957), the "definition of a term is a statement of its meaning in other terms which are supposed to be well known" (p. 85). But this seemingly innocuous statement hides four crucial aspects of the role of definitions in mathematics: the creation of meaning, the need for undefined terms, the substitution criterion, and the use of mental or physical models. These are discussed, in turn, below.

"The mathematician is not concerned with the current meaning of his technical terms.... The mathematical definition *creates* the mathematical meaning" (Polya, 1957, p. 86). This view of definitions, embodied in the character of Humpty Dumpty above, reached its height in the formalism of Russell, Whitehead, Peacock, Hilbert, and others, but in fact, has its early roots in Kant.[3] Formalists maintain that mathematics involves manipulating meaningless symbols according to the formal rules of the system, and the primary criterion is that the system is consistent. Of course, this point of view requires certain ontological and epistemological commitments or at least changes in perspective. Hamilton, for example, insisted that the symbols must stand for something 'real'—if not material objects, then mental constructs (Kleiner, 1987). Nonetheless, some mathematicians were reluctant to adopt a formalist view. Graves, for example, on Hamilton's invention of the quaternions, responded, "I have not yet any clear view as to the extent to which we are at liberty to create imaginaries, and to endow them with supernatural properties" (quoted in Kleiner, 1987, p. 233). By 1844, however, less than a

---

[3] This was the fundamental idea behind Kant's notion of synthetic a priori statements. "Whereas, therefore, mathematical definitions *make* their concepts, in philosophical definitions concepts are only *explained*" (cited in Beth & Piaget, 1966, p. 13).

year after Hamilton had published his work on quaternions, Graves and other

mathematicians begin creating new mathematical structures almost without restraint.[4]

It is well known that in any mathematical system some terms must be taken as primitive,

that is, left undefined, for the only alternatives are circular definitions or infinite regress,

neither of which is logically tolerable. If one accepts that definitions create the meaning

of terms, where, then, do undefined terms acquire their meaning? Just as the axioms of

natural numbers form implicit definitions of natural numbers (Beth & Piaget, 1966, p.

68), the axioms of any mathematical system give implicit definitions of the undefined

terms of that system. Couturat made this point by distinguishing between *direct*

definition and *definition by postulates*, the latter applying not to a single notion but a

system of notions (cited in Poincaré, 1946, p. 453).

To adhere to the principle that all assumptions must be made explicit in the axioms and

definitions, Pascal was apparently the first to put forward the criterion of substitution:

that the definition permits us "to substitute the definition in place of the defined" (cited in

Beth & Piaget, 1966, p. 38). Thus, the substitution principle is a way of ensuring that

every theorem and every proof could, in principle, be written using only the undefined

terms, the axioms, and the laws of logic. Mariotti and Fischbein (1997) clarify the

implications of this view:

> In the new theory, it is not possible to prove anything which was not already
> possible to prove in the old one. From the formal point of view, a definition does
> not enlarge the power of the theory. A definition is rather a correct definition just
> because it can be eliminated. (p. 222)

---

[4] This approach is not without its failures. There is one apocryphal story, for example, of a mathematician
who proved all sorts of theorems based on a set of axioms that, it turned out, were satisfied only by the
empty set. See Wilensky (1991, note 4) for a similar example.

And although this extreme formalism is rarely carried out, the first substitution is a standard mathematical practice. In other words, when proving theorems about a newly defined mathematical concept, the standard approach is to replace the term by its definition. Polya (1957) calls this process "the elimination of technical terms" by "going back to definitions" (p. 89).

If the axioms and the definitions are to be the source of meaning within a mathematical system, then the implication is that all proofs and formal reasoning should proceed from the axioms, definitions, and previously proven theorems. This is what is meant by mathematical rigor. Because physical and mental models of the system might carry meaning that does not follow from the axioms and definitions, such models cannot be trusted and thus are inadmissible in proofs. The validity of a proof is independent of the meaning of the descriptive terms. To emphasize this point, Hilbert once said, "One must be able to say at all times—instead of points, straight lines, and planes—tables, chairs, and beer mugs" (Reid, 1986, p. 57). The implication is that no matter how the terms are interpreted, a counterexample will never be produced (Lakatos, 1976, p. 100). Taken to an extreme, the formalist approach identifies mathematics with metamathematics and with logic, resulting in a rather bleak picture:

> The subject matter of metamathematics is an abstraction of mathematics in which mathematical theories are replaced by formal systems, proofs by certain sequences of well-formed formulae, definitions by "abbreviatory devices" which are "theoretically dispensable" but "typographically convenient." (Lakatos, 1976, p. 1, drawing on Church, Peano, Russell, Whitehead, and Pascal)

But even Russell (1938) admits:

> It is a curious paradox, puzzling to the symbolic mind, that definitions, theoretically, are nothing but statements of symbolic abbreviations, irrelevant to the reasoning and inserted only for practical convenience, while yet, in the

> development of a subject, they always require a very large amount of thought, and often embody some of the greatest achievements of analysis. (p. 63)

Thus, despite the formalism and the claim to disregard meaning, thinking and meaning remain crucial characteristics of mathematical activity. It is true that the words are symbolic abbreviations, but the concepts for which they stand (and hence the meaning that they are intended to carry) took time to formulate and constitute significant human achievements.

When one acknowledges the importance of both rigor and meaning, perhaps it is not surprising that most mathematicians are Platonists on weekdays and formalists on the weekends (P. J. Davis & Hersh, 1981), seeking to discover timeless mathematical truths and simultaneously adhering to meaningless formalism.

Definitions in the History of Mathematics

The history of mathematics is full of examples where the definitions changed in order to correct for unintended consequences, including such "simple" concepts as function, continuity, and polyhedron (see Lakatos, 1976). Much of the history of mathematics has been spent trying to figure out what the concepts are, trying to "get the definitions right," so that they correspond to the intuitions that the mathematicians had in mind.

> We begin with a vague concept in our minds, then we create various sets of postulates, and gradually we settle down to one particular set. In the rigorous postulational approach the original concept is now replaced by what the postulates define. This makes further evolution of the concept rather difficult and as a result tends to slow down the evolution of mathematics. It is not that the postulation approach is wrong, only that its arbitrariness should be clearly recognized, and we should be prepared to change postulates when the need becomes apparent. (Hamming, 1980, p. 86)

The process of "gradually settling down" on a definition deserves elaboration. Drawing on Lakatos (1976), the process goes something like this:

- Get a mathematical idea that can form the beginning of a concept.

- Create an informal definition that seems to describe the concept.

- Formalize that definition.

- Reason from that definition to determine what it implies: Create some examples; prove some theorems; look for equivalent or closely related characterizations of the concept.

- Modify the formal definition to exclude undesired consequences.

- Alternatively, enlarge or otherwise alter one's understandings and intuitions of the concept to accommodate these newfound possibilities.

There are several points to make about this process. First, sometimes the modifications to the definition amount to little more than eliminating undesirable examples through ad hoc redefinitions, a seemingly nonmathematical practice Lakatos (1976) called *monster barring*.

Second, the process incorporates apparent contradictions on the role of definitions: On the one hand, the definition is taken to create a mathematical object and to give a term its meaning. And on the other hand, the definition is carefully chosen to capture a specific meaning and with an instrumental or expository purpose. Because both of these roles are mathematically indispensable, their relationship is better viewed as dialectical.

Third, once agreement is reached, a definition can be taken as primary—as though it had been handed down on stone tablets. In the deductivist, definition-theorem-proof format of much mathematical presentation and exposition, the dialectical evolution of the concept and its definition are subsequently ignored.

> In deductivist style, all propositions are true and all inferences valid. Mathematics is presented as an ever-increasing set of eternal, immutable truths.... Deductivist style hides the struggle, hides the adventure. The whole story vanishes, the successive tentative formulations of the theorem in the course

of the proof-procedure are doomed to oblivion while the end result is exalted into sacred infallibility.[5] (Lakatos, 1976, p. 142)

Such strict adherence to formalism is to ignore the history of mathematics, rendering the teacher and students blind to important epistemological obstacles and ignorant of required changes in perspective. And not only are students deprived of opportunities to see and benefit from the growth of particular definitions, they are thus also deprived of opportunities to appreciate the evolution of the role of definition in mathematics.

It should not be at all surprising that students have difficulty accepting the role of definitions in modern mathematics when it was not fundamental in until the nineteenth century.

Definitions and Mathematical Intuition

By relaxing the demands of pure formalism, one can adopt a position in which intuition and meaning are central to mathematical activity but where logic and rigor are available as tools for verification. As Hadamard said, "The object of mathematical rigor is to sanction and legitimize the conquests of intuition, and there never was any other object for it" (cited in Ahlfors et al., 1962, p. 192). Despite the rhetoric of formalism and rigor, it seems that metaphorical thinking (Sfard, 1994) and intuition remain central:

> It is significant that when a mathematician reads a theorem which conflicts with his intuitive expectations his first move is to doubt not his intuition but the proof. He trusts his intuition more. If after having checked the proof carefully he becomes convinced that it is correct, he then inquires into what may be wrong with his intuition. (Kline, 1973, p. 160)

Thurston (1994) acknowledges putting "a lot of effort into 'listening' to my intuitions and associations, and building them into metaphors and connections" (p. 165). He discusses

---

[5] Lakatos condemned mathematics and science education as a hotbed of authoritarianism and as the worst enemy of critical thought (Lakatos, 1976, pp. 142-143, note 2).

the relationships among definition, understanding, and intuition by presenting several

characterizations of the concept of derivative:

(1) Infinitesimal: the ratio of the infinitesimal change in the value of a function to the infinitesimal change in a function.

(2) Symbolic: the derivative of x^n is nx^{n-1}, the derivative of $\sin(x)$ is $\cos(x)$, the derivative of $f \circ g$ is $f'{\circ}g * g'$, etc.

(3) Logical: $f'(x) = d$ if and only if for every ϵ there is a δ such that when $0 < |\Delta x| < \delta$,

$$\left| \frac{f(x + \Delta x) - f(x)}{\Delta x} - d \right| < \delta \quad [sic].$$

(4) Geometric: the derivative is the slope of a line tangent to the graph of the function, if the graph has a tangent.

(5) Rate: the instantaneous speed of $f(t)$, when t is time.

(6) Approximation: The derivative of a function is the best linear approximation to the function near a point.

(7) Microscopic: The derivative of a function is the limit of what you get by looking at it under a microscope of higher and higher power.

This is a list of different ways of *thinking about* or *conceiving of* the derivative rather than a list of *logical definitions*. Unless great efforts are made to maintain the tone and flavor of the original human insights, the differences start to evaporate as soon as the mental objects are translated into precise, formal and explicit definitions.

I can remember absorbing each of these concepts as something new and interesting, and spending a good deal of mental time and effort digesting and practicing with each, reconciling it with the others. I also remember coming back to revisit these different concepts later with added meaning and understanding....

The list continues; there is no reason for it ever to stop. A sample entry further down the list may help illustrate this. We may think we know all there is to say about a certain subject, but new insights are around the corner. Furthermore, one person's clear mental image is another person's intimidation:

37. The derivative of a real-valued function f in a domain D is the Lagrangian section of the cotangent bundle $T^*(D)$ that gives the connection form for the unique flat connection on the trivial **R**-bundle $D \times \mathbf{R}$ for which the graph of f is parallel. (pp. 163-164)

Despite the fact that Thurston's 37 characterizations are not definitions, for him they may

function as definitions in reasoning within certain problem settings, though perhaps

without the precision of a formal definition. From his use of words such as *flavor, tone,*

and *insight*, it is clear that these characterizations are full of meaning for Thurston. And it is worth emphasizing that attempting to formalize these ways of thinking runs the risk of changing their character and perhaps their usefulness in reasoning.

Definitions and Learning Mathematical Concepts

Drawing on the history and expository literature discussed above, I take the position that meaning is central to mathematical learning and mathematical thought and that careful reasoning from precise definitions is an important capability to be cultivated in mathematics majors. What, then, is the relationship between definitions and learning?

Skemp (1987) proposes two principles for teaching mathematical concepts:

1. Concepts of a higher order than those which people already have cannot be communicated to them by a definition, but only by arranging for them to encounter a suitable collection of examples.
2. Since in mathematics these examples are almost invariably other concepts, it must first be ensured that these are already formed in the mind of the learner. (p. 18)

Skemp does not indicate how or when he came to these sensible conclusions or what sort of empirical data support them. But from the preceding discussion, it should be clear that mathematicians are distinguished by their ability to violate the first of these principles, and it appears that Halmos, at least, transcends the second principle by constructing his own examples: "A good stock of examples, as large as possible, is indispensable for a thorough understanding of any concept, and when I want to learn something new, I make it my first job to build one" (cited in Gallian, 1994, p. 34).

Thus, learning to violate or transcend these principles is a requirement for entering into the mathematical community. Specifically, a student pursuing a degree in mathematics must learn to build understanding (and perhaps create the examples) by reasoning from a

definition. But here we have a conflict: On the one hand, students must learn to reason from the definitions rather than exclusively from their concept images because pictures, metaphors, and informal understandings are sometimes unreliable. On the other hand, the source of their reasoning may continue to be the models and metaphors that they keep in mind.

> Conflicts between the empirical (lexical) approach and the theoretical approach can represent a real obstacle for the students' understanding. That is the reason why the problem of introducing pupils to the mathematical process of defining constitutes a crucial point in mathematics education, which needs to be faced directly. (Mariotti & Fischbein, 1997, p. 226)

Adopting a formalistic approach to definitions may require epistemological and ontological changes in perspective. Suffice it to say that these changes in approach and perspective are rarely made explicit to the student. It is possible that successful mathematicians learned to reason from the definitions without ever being aware of these changes in perspective. And by the time they are teaching courses to undergraduates, this approach has become so natural that they do not realize that nothing has changed for the student.

From the naïve student's point of view, definitions are lexical: They are used to describe or explain ideas that already exist (Vinner, 1976). But as concepts expand, become more general, and allow inclusion of never-before-imagined examples, the natural meaning gets lost. What does it take to understand the importance of formal reasoning, which includes reasoning from definitions?

Students do not understand the role of mathematical definitions in general and, in particular, do not know how to reason from definitions. Mariotti and Fischbein (1997) found, like Vinner, that students may know the definition and yet fail to correctly identify

whether objects satisfy the definition because the "concepts are often, implicitly or explicitly, distorted by gestalts" (p. 244). These distortions can take the form of additional conditions that remain implicit. They suggest that "in empirical domains, one tends to adapt the definitions to the empirical data—and exceptions are admissible" (p. 245).

Borasi's Work

Some of the most fully elaborated work on learning the nature and role of mathematical definitions comes from Borasi. In *Learning Mathematics Through Inquiry* (Borasi, 1992), she presents a detailed analysis of a "mini-course" with two high school students. Although she had broad mathematical goals, she chose to focus on the notion of definition because it "presents a beautiful example of the more humanistic and contextualized aspects of mathematics" (p. 7).

Before presenting any of the data or analysis, she sets forth five criteria for definitions:

- *Precision in terminology.* All the terms employed in the definition should have been previously defined, unless they are one of the few *undefined terms* assumed as a starting point in the axiomatic system one is working with.

- *Isolation of the concept.* All instances of a concept must meet all the requirements stated in its definition, while a noninstance will not satisfy at least one of them.

- *Essentiality.* Only terms and properties that are strictly necessary to distinguish the concept in question from others should be explicitly mentioned in the definition.

- *Noncontradiction.* All the properties stated in a definition should be able to coexist.

- *Noncircularity.* The definition should not use the term it is trying to define. (Borasi, 1992, pp. 17-18)

She then points out that these criteria stem, in part, from the fact that we want a definition to:

1. Allow us to discriminate between instances and noninstances of the concept with certainty, consistency, and efficiency (by simply checking whether a potential candidate satisfies all the properties stated in the definition).

2. "Capture" and synthesize the mathematical essence of the concept (*all* the properties belonging to the concept should be logically derivable from those included in its definition). (p. 18)

During the mini-course the students wrote, created, and modified definitions, extended definitions to new domains, and constructed definitions in new contexts, such as taxicab geometry. In one of the activities, inspired by Lakatos's (1976) example of the evolution of the definition of polyhedron, Borasi asked the students to construct a definition of polygon, believing that such an experience "could help students appreciate that definitions are really *created by us*, even in mathematics, where everything may seem rigid and predetermined (at least to most students)" (Borasi, 1992, p. 44).

Based upon her analysis, Borasi concludes that the students changed their conceptions not only of mathematical definitions but also of mathematics. Through the experience, she also changed her view of mathematical definition, realizing a deeper understanding of the role of context and purpose in the creation and evaluation of mathematical definitions. Furthermore, she reconsidered the role of her five criteria set forth above, for those criteria are satisfactory only in specific mathematical contexts when it is reasonable to imagine the definition is fixed. When the context changes, however, the criteria must be relaxed, at least for a moment, and the definition may change.

Which Definition?

Which of the various equivalent formulations of a concept is chosen as its definition? The choice is not arbitrary, despite the formalist claim to the contrary. In a formal presentation of a concept, the definition that is chosen is usually the one that is most

elegant or most useful in proofs concerning the concept, which implies that it is formal and often that it is minimal and otherwise concise. In a pedagogical presentation, a definition is chosen with a particular pedagogical purpose. Poincaré (1946) makes it clear that the choices should not be the same:

> What is a good definition? For the philosopher or the scientist it is a definition which applies to the objects defined, and only those; it is the one satisfying the rules of logic. But in teaching it is not that; a good definition is one understood by the scholars [students]. (p. 430)

This view toward formalism was echoed by 75 mathematicians who, responding to the excesses of the new math, warned that "premature formalization may lead to sterility" (Ahlfors et al., 1962, p. 190).

From the mathematics education community, Mariotti and Fischbein warn that "the formal approach does not grasp the very process of defining" (p. 222) and suggest, instead that definitions have a constructive and creative role and actually bring new concepts into existence. They propose that "a definition is to be considered a 'good' definition as far as the new object starts to live by itself and may become the subject of a new theory" (p. 223).

There seems to be very little discussion in the literature about the problem of conflicting definitions, other than occasionally mentioning that parallelograms are sometimes but not always included as trapezoids. What is rarely acknowledged is that there are also conflicting definitions of *natural number* (including vs. omitting 0), *ring* (including vs. omitting 1), and *integral domain* (including vs. omitting commutativity). Thus, although a particular definition may be chosen with a particular expository or pedagogical purpose, there is a certain arbitrariness in which objects are thus defined.

Definitions in Textbooks

How are definitions treated in mathematical texts? Explicit definitions can send implicit messages about the role of definitions in mathematics. Raman (1998) found that texts send conflicting messages about the purpose and use of mathematical definitions. Vinner (1991) suggests that mathematics textbooks and classroom practice are partly based on the following assumptions:

1. Concepts are mainly acquired by means of their definitions.
2. Students will use definitions to solve problems and prove theorems.
3. Definitions should be minimal.
4. It is desirable that definitions will be elegant.
5. Definitions are arbitrary. (pp. 65-66)

What conclusions do students draw from such implicit messages? Rin (1982) found that students do not understand that the definition is to be the official source of information about the concept and that textbooks sometimes compound the problem by burying definitions in the text or the exercises, or by leaving implicit the quantifiers or the appropriate range of the variables.

Would it be better if texts were explicit about the nature and role of definitions in mathematics? Textbooks are rarely explicit about the role of definition, although some texts emphasize that all definitions are "if and only if" statements (e.g., Fraleigh, 1989, p. 3; Bittinger, 1982, p. 40), and a few point out that a definition is an abbreviation (e.g., Bittinger, 1982, p. 40). But these are statements about what a definition *is*, which is singularly unhelpful to students, who believe they already know what definitions are and implicitly operate on this basis (Vinner, 1976). Instead, students need to learn what to *do* with definitions.

Summary

Definitions play opposing roles in mathematical thinking and learning, serving simultaneously to describe and support informal mathematical intuition and meaning and to create meaning through the imposition of formalism. These opposing roles are evident in the history of mathematics, in the evolution of definitions of key concepts, in mathematics textbooks, and in research into students' use of definitions in mathematical learning. In order to accommodate both of these roles into descriptions of students' understandings of concepts in abstract algebra, I took a broad view of definitions, with the aim of capturing both meaning and level of precision. Thus, the analysis included not only students' attempts at formal definitions but also the ad hoc personal definitions they provided when I asked for the meaning of a term or statement.

Abstraction Versus Generalization

Mathematically, a definition creates meaning for a new concept, but psychologically, new concepts are created through processes of abstraction and generalization. Abstraction and generalization are fundamental human activities that become critically important in advanced mathematics. Dreyfus (1991) suggests, for example, that the ability to consciously make abstractions from mathematical situations "may well be the single most important goal of advanced mathematical education" (p. 34). I begin with abstraction, which played a central role during the new math era. Here is one view from that era:

> The process of *abstraction* is defined as the process of drawing from a number of different situations something which is common to them all. Logically speaking it is an inductive process; it consists of a search for an attribute which would describe certain elements felt somehow to belong together....
>
> For example the forming of the concept of the natural number two is an abstraction process, as it consists mainly of experiences of pairs of objects of the greatest possible diversity, all properties of such objects being ignored except

that of being distinct from each other and from other objects. The essential common property of all such pairs of objects is the natural number two. (Dienes, 1961, pp. 281-282)

Piaget distinguishes between empirical abstraction, which starts from perceived objects, and reflective abstraction, which starts from actions and operations (Beth & Piaget, 1966, pp. 188-189).

As an adjective, *abstract* is usually contrasted with *concrete*. Wilensky (1991) points out, however, that concreteness is not a property of an object but a property of a person's relationship to an object. Concreteness, he suggests, measures the degree of our relatedness to the object (the richness of our representations, interactions, connections with the object), how close we are to it, or the quality of our relationship with the object. Thus, any object can become concrete for someone. He notes that this point of view turns the old definition of concrete on its head, so that thinking concretely is not narrow but rather opens up a whole world of ideas and relationships. Frorer, Hazzan, and Manes (1997) agree with Wilensky and suggest two additional themes in abstraction: ignoring details and thinking of things in terms of properties rather than actual components.

As for generalization, it should be mentioned that generalization and abstraction are often confounded in the literature (e.g., Dreyfus, 1991) and are sometimes treated as essentially synonymous (e.g., Beth & Piaget, 1966; Vygotsky, 1934/1986).[6] Tall (1991) suggests, however, that generalization simply involves an extension of familiar processes whereas abstraction requires mental reorganization. Thus, *generalization* is the application of an existing process or structure to a broader class of objects (see also Dienes, 1961). Generalization may be contrasted with *specialization*, where the scope of a process or

---

[6] Piaget speaks mostly about abstraction and Vygotsky mostly about generalization, but it is possible that these choices were made by the translators.

structure is restricted in some way. *Abstraction*, on the other hand, creates a new structure on a higher level, which is not a deduction but a construction process. Abstraction may be contrasted with *exemplification*, which creates a specific instance of the abstract structure or idea. A mathematical metaphor may help make the distinction more clear: Generalization and specialization are about creating supersets and subsets; abstraction is about constructing a new kind of set and exemplification involves choosing an element of that set. In mathematical thinking, of course, abstraction and generalization may be operating simultaneously or consecutively. It is not always possible, however, to separate the two processes, such as in the introduction of the notation of an asterisk '*' to serve generally for an abstract binary operation.

Processes Versus Objects

One of the central theoretical themes in advanced mathematical thinking is the distinction between process and object conceptions of mathematical ideas. Although the terminology is diverse, the primary distinction is that a *process* is an activity carried out through some sort of procedure, whereas an *object* can be conceived of as a single entity. Many mathematical ideas can be conceived both as processes and as objects, so the distinction is psychological. Sfard (1991) distinguishes between operational and structural conceptions. Harel and Kaput (1991) distinguish between a process and a conceptual entity. Dubinsky and his colleagues (Dubinsky, 1991; Breidenbach, Dubinsky, Hawks, & Nichols, 1992) also distinguish between processes and objects and offer additional categories described above. In reviewing this literature, Tall, Thomas, G. Davis, Gray, and Simpson (2000) suggest that it is possible to ascertain whether students have constructed a mental object based on the way they talk and write about the concept.

Object conceptions allow for descriptive discourse and attention to structural features of the mathematical ideas. Process conceptions, on the other hand, are confined to narrative discourse.

There are some differences among the perspectives of these researchers, but they are in agreement that a learner cannot meaningfully act on a process with another process until the former has become an object in his or her mind. This kind of mental construction is called *encapsulation* (or reification, or entification). For some concepts, encapsulation seems to be extremely difficult for most students, and coset may be one such concept, as suggested in the literature review above.

On the other hand, there are "natural," implicit instances of encapsulation. For example, from a process of counting, a young child creates an understanding of "4" as an object that describes what is the same about the wheels on a car, the legs on a dog, and the sides of a rectangle. To emphasize the ambiguity in the symbolism for mathematical ideas, Gray and Tall (1994) coined the term *procept*. Thus, "4 + 5" is a symbol that simultaneously denotes both the process of addition and object that results. In abstract algebra, given a subgroup H and a group element a, the notation aH simultaneously specifies the process for calculating the cosets of H and the result of one of those calculations for the particular value a.

Gray and Tall (1994) further distinguish between a *procedure*, where the focus is on step-by-step details, and a *process*, where the concern is with the result (as dependent on the initial state). A procedure, in other words, refers to a specific algorithm for carrying out a process. The process of addition, for example, can be carried out by many different procedures, including counting all, counting on, or pressing buttons on a calculator.

Similarly, there are many procedures that can be constructed to determine whether a subgroup is normal, but any (or all) of them may be conceptualized as the process for determining normality.

Metaphor

Drawing on the work of Lakoff and Núñez (1997, 2000), I take the position that mathematical concepts are predominantly metaphorical in nature. Despite the central role of precise formal definitions, mathematical thinking is usually guided by metaphors. This recent work in the cognitive science of mathematics is based upon a large body of empirical work in cognitive science that has produced three major findings: "The mind is inherently embodied. Thought is mostly unconscious. Abstract concepts are largely metaphorical" (Lakoff & Johnson, 1999, p. 1). In identifying the metaphors that support particular concepts, most of the evidence comes "from language—from the meanings of words and phrases and from the way humans make sense of their experiences" (p. 115).

Lakoff and Núñez (2000) elaborate the metaphorical nature of mathematics, concentrating first on arithmetic and later on concepts such as the real numbers, limits, and continuity, building up to a case study of the equation $e^{\pi i} + 1 = 0$. In their analysis, some mathematical concepts are based upon grounding metaphors, such as Sets Are Containers,[7] that grow out of bodily experience in the world. Other concepts link to, build upon, or coordinate previously established metaphors, so that "much of the 'abstraction' of higher mathematics is a consequence of the systematic layering of metaphor upon metaphor, often over the course of centuries" (p. 47). A metaphor "A is

---

[7] Throughout this section, I have adopted Lakoff and Núñez's convention of capitalizing the names of the metaphors.

B" is a mapping from a source domain B to a target domain A, where the source domain is typically more familiar. The inferential structure of the source domain gives structure to the target domain, often introducing new elements or relationships in the target domain. The Arithmetic-Is-Object-Collection metaphor, for example, provides grounding for the commutativity of addition; the Arithmetic-Is-Motion-Along-a-Path metaphor provides grounding for the concepts of zero and fractions.

Between their detailed treatments of arithmetic and real analysis, Lakoff and Núñez provide a short discussion of metaphorical nature of abstract algebra. A key construct in their analysis is the Fundamental Metonymy of Algebra, which allows us to reason about numbers or other entities without knowing which particular entities we are talking about. This mathematical notion depends upon its everyday version, the Role-for-Individual metonymy, by which we are able to imagine carrying out actions with whoever (or whatever) fills the required role.

Because algebra in general and abstract algebra in particular are about essence, Lakoff and Núñez (2000) discuss the Folk Theory of Essences, which includes such notions as "everything is a specific kind of thing" (p. 107), "kinds are categories" (p. 108), and "the essence of a thing is an inherent part of that thing" (p. 108). Essence is characterized by three metaphors: Essences Are Substances, Essences Are Forms, and Essences Are Patterns of Change. The Folk Theory of Essences was behind Aristotle's definition of *definition* as a "list of properties that are both necessary and sufficient for something to be the kind of thing that it is, and from which all its natural behavior flows" (p. 109) and also behind Euclid's axiomatic (or postulational) approach.

Building on the metaphor Essences Are Forms, Lakoff and Núñez (2000) put forward a foundational metaphor for abstract algebra: The essence of a mathematical system is an abstract algebraic structure, which is taken to include the elements in the structure, the operations used on those elements, and the essential properties of the operations (p. 111). I accept their guiding principles and much of their analysis, but regarding abstract algebra, their analysis falls short on two counts. First, their notion of a mathematical structure is too restrictive because mathematical structures do not always have operations. Second, some of their metaphors are backwards in the sense that the source domain is less familiar and more abstract than the target domain it is intended to describe.

For example, they claim, in effect, that Z_3 is the abstract group with three elements. While this is a true statement, it is not a helpful metaphor. Furthermore, their description of the abstract group with three elements uses the set $\{I, A, B\}$ with the expected operation table. This group, it is important to note, is not the abstract group itself but a particular representation of it—a representation, moreover, that does not easily support calculation. If abstract concepts are metaphorical, as they claim, then the appropriate metaphor is that the abstract group with three elements is Z_3, thereby providing a familiar instantiation of the abstraction. This metaphor does not provide a complete characterization of the abstract group, however, because it leaves out the necessary abstraction. Where do abstractions come from, and by what process do they come about? Unfortunately, the *process* of abstraction (abstracting) is conspicuously missing from their analysis, although the results of abstractions are covered in their Folk Theory of Essences.

To remedy the analysis provided by Lakoff and Núñez, it is necessary first to broaden the notion of structure. Rickart (1995) suggests that there is "not much hope for stating in a few lines a precise and complete definition of structure" (p. 11). In its place, he suggests, puts forward "an admittedly imprecise approximate definition, which is then elaborated and made increasingly more complete through examples and explanations" (p. 11). Rickart's definition is the following: "A structure is any set of *objects* (also called *elements*) along with certain *relations* among those objects" (p. 17, emphasis in original). The advantage of this definition over that of Lakoff and Núñez is that it can accommodate topological and order structures. Furthermore, it is consistent with the notion of structure in fields outside mathematics, such as linguistics, psychology, biology, and anthropology (Rickart, 1995; see also Piaget, 1970b).

Algebraic structures fit this definition by way of an appropriate interpretation of the relations among the objects. A group is a structure, for example, in that the objects are the elements and the relation is a ternary relation defined in terms of the group operation: The group elements in an ordered triple (g, h, k) are related if $gh = k$. (Rickart, 1995, p. 53) The group axioms can be also be specified as relations.

Analysis of the concepts in group theory, focusing primarily on language, leads to the conclusion that group theory is guided primarily by two metaphors:

- Groups Are Sets
- Groups Are Structures

At first sight, these do not appear to be metaphors at all but would be more accurately characterized as obvious statements of fact. A group, after all, *is* a set. But sets and structures are themselves metaphors, which may be traced back to metaphors that are

grounded in bodily experience. With the above definitions of *structure* as a set of elements with relations among them, it is possible to reduce this to one metaphor: Groups Are Structures. A guiding principle behind the concept of structure, however, is that a structure is independent of the elements themselves, depending only on the relations. For this reason and because the Groups-Are-Sets metaphor is so predominant, it makes sense to consider it separately.

The metaphor Groups Are Sets is quickly grounded through the Sets-Are-Containers metaphor, hence groups are containers. This metaphorical thinking becomes apparent in expressions such as "an element *g in* a group *G*." When a set is *closed* under an operation, as all groups are, the container is metaphorically closed, preventing the elements from escaping. The Groups-Are-Containers metaphor takes a slightly different character in the question "Where does this element live?" suggesting a Containers-Are-Territories metaphor that becomes particularly apparent when the group is the *domain* or *codomain* of a homomorphism.

The metaphor Groups Are Structures becomes apparent in the etymological derivation of the term *isomorphism* as meaning "same *form*." The metaphor of structure also suggests that the form is in some sense incomplete, providing only the framework that is the relations among the elements. The elements themselves are unimportant details. When a particular set *under* a particular operation is said to be a group, it is the operation that provides the structure, by sitting metaphorically above the elements and imposing form on the relations among them.

In constructing the above definition of structure and structuralism that applies across diverse fields, Rickart (1995) observes, "The objective of a structuralist approach to a

subject is to extract the essential information from the background of irrelevant or unimportant information" (p. 19). Thus, structures are also essences. And since "essence" and "essentially" share the root *essens* (Latin present participle of *esse*, "to be"), the merged concepts of structure and essence is revealed in the semantic equivalence between the statements "the groups are essentially the same" and "the groups have the same structure."

By building metaphors on top of metaphors, abstraction on top of abstraction, it is possible to create hierarchical chains of metaphors that ultimately depend upon grounding metaphors. What, then, is the metaphorical relationship between essence and structure in mathematics? On the one hand, structures are essences, but on the other hand, the essence of a mathematical system is its structure. This is not circular reasoning, however, but an example of a conceptual blend, where two concepts combine to form a deeper unified concept while also contributing to a more flexible understanding of each of the concepts individually. The conceptual blend Numbers Are Points on a Line, for example, beginning with Descartes, paved the way for profound connections between geometry and algebra. Thus, structures are essences and vice versa. Saying structures are essences highlights the push toward abstraction that is a guiding principle behind the structuralist approach. Saying mathematical essences are structures, on the other hand, gives body and form to an otherwise ethereal concept.

Consider the definition of a structure as a set with relations, along with the metaphor Structures Are Essences and the three metaphors that characterize essence: Essences Are Substances, Essences Are Forms, and Essences Are Patterns of Change. Taken together these metaphors reaffirm the point made in chapter 1 that mathematics can fit under any

of the three themes in Devlin's (2000) characterization that "modern mathematics is about abstract patterns, abstract structures, and abstract relationships" (p. 136).

This discussion provides but a preliminary analysis of the metaphorical nature of the concept of group. Group theory involves many more concepts, some of which are discussed metaphorically in the analysis that follows. I note here only a few key metaphors that informed this study: Subgroups are subsets, homomorphisms are functions, cosets are sets, and sets are objects.

Naming and Notation

Thus far, I have discussed concepts and their definitions, and certainly mathematical thinking and discourse require concepts and definitions. But students often use language and notation incorrectly. Rin (1982) suggests that students' linguistic misbehaviors are interpretable as reflective of deficient understanding or of deficient expressive powers (p. 10). Mathematical learning requires not only constructing concepts but also learning the standard names and notations for those concepts and the appropriate verbal and mathematical syntax for referring to those concepts in mathematical discourse. In this study, issues of naming and notation were central, as they are key components of the larger issues of language and representation.

One commonly advocated approach for teaching that promotes understanding is to provide opportunities for students to explore concepts before giving the concept a name (e.g., Leron & Dubinsky, 1995). After the students have had sufficient experience and have noticed certain regularities, the relevant concepts can be given names. The naming itself is seen as unproblematic. As the mathematician John H. Conway (1995, April 13)

once said when discussing whether spherical geometry is a non-Euclidean geometry, "This is the sort of question that bugs me! Being about names, it's not a mathematical [*sic*], so 'the answer' really doesn't matter."

But this study and at least one other (Lajoie & Mura, 2000) suggest that attaching the name is sometimes not as simple as supposed. As part of a larger study into the learning of abstract algebra, Lajoie and Mura found that when asked about cyclic groups, students seemed to focus on metaphors of coming back to the start, cycles, and images of circles. Most students did not consider infinite cyclic groups to be cyclic because "you don't come back." Lajoie and Mura propose several sources of confusion: inappropriate use of mathematical definition, semantic contamination from everyday language (a la Pimm, 1987), confusion with cyclic permutations, and nonstandard definitions of powers and generators. They point out, first, that incorrect conceptions can lead to correct answers for many questions about Z_n and, second, that in the history of group theory, Ruffini, Cauchy, and Jordan used similar imprecise "circular" language and excluded infinite cyclic groups. As possible solutions, they suggest drawing students' attention to differences between mathematical and ordinary uses of words and explicit teaching of the role of definition in mathematics.

The question here is, What is the relationship between a concept and its name? What is gained by giving a collection of physical or mental entities a name? How is thinking constrained by the particular name chosen? Given the name of a new concept, what understanding do students associate with that name and how? These uncertainties imply that the notion of concept image must be applied flexibly in the analysis to allow for the possibility of nonstandard connections between concepts and names.

Similar questions may be asked regarding mathematical notation. As mentioned above, students often use mathematical notation improperly in abstract algebra (Selden & Selden, 1978; Hazzan, 1994). Mason and Pimm (1984) suggest that students' difficulty may be caused partly by ambiguity in the notation itself. What, for example, does $2N$ stand for? Is it any even number or all even numbers? Is it specific, particular, generic, or general? Perhaps it is not an even number at all, for it would never appear in a list of even numbers. Is it shorthand for $\{2N: N \text{ a whole number}\}$? In this case, $2N$ is not an even number but an instruction to carry out a calculation. Mason and Pimm suggest that, for students, $2N$ sometimes represents any even number and that as a result they may show that the sum of two even numbers is even by writing $2N + 2N = 4N$. What is missing is awareness of $2N$ as a *particular* even number. *Any* has two interpretations: generic and general, and the latter implies "every." Recognizing that in fact "$2N$" is merely marks on the paper, they point out that the meaning has to do with perception.

Durkin and Shire (1991) suggest that some difficulties with language arise from ambiguities in the language itself, pointing in particular to *polysemy*, the property of some words to have distinct but related meanings. There are many examples, such as *function* or *group*, in which an everyday word takes on a specialized meaning in mathematics. Durkin and Shire suggest that the words *some* and *any* may be similarly confounded, providing additional insight into the observations of Mason and Pimm (1984) above. What is more problematic is when words take on multiple but related meanings within mathematical discourse itself. Zazkis (1998) suggests, for example, that the term *quotient* takes on different meanings depending upon whether one is dealing

with whole numbers or rational numbers. In abstract algebra, it appears that the term

cycle is mathematically polysemous (Lajoie & Mura, 2000).

Taken together, these studies suggest that attaching names and notations to ideas involves

subtle distinctions and ambiguities in language to fit with subtle conceptual distinctions

among mathematical objects. Thus, empirical and theoretical work must allow for and

explain the possibility that students might take words and notations to carry nonstandard

meanings. For this study, the analytical tools were furnished by semiotics.

Semiotics

Figure 1. Ceci n'est pas un groupe

*	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Ceci n'est pas un groupe

To paraphrase René Magritte regarding his painting *Ceci n'est pas une pipe* [This is not a

pipe] (see Foucault, 1983), the table in Figure 1 is not a group. To be precise, it is a

representation—a sign. The sign is certainly not itself the abstract group with three

elements. A central theme in this study is the relationship between mathematical notation

and language, concept definitions, and conceptual understanding. This is essentially the

relationship among signs, objects, and meaning, which is the province of the field of

semiotics, or the study of signs. Whereas semantics is the study of meaning in language,

semiotics is the study of meaning in signs, which includes language as a subset.

Semiotics is generally recognized to have been founded in the work of Swiss linguist

Ferdinand de Saussure (1857-1913) and independently in the work of American

mathematician and philosopher Charles Sanders Peirce (1839-1914). This section

presents the influential distinctions made by Saussure and Peirce on the nature of signs

and of sign systems and their relationship to meaning. These basics of semiotics are then

connected to the work of Vygotsky, with particular attention to learning and the

relationship between personal and social meaning. Additional theoretical background is

then provided on various types of signs and on the semiotics of mathematics. The section

closes with a discussion of the analytical framework that was used to guide and organize

the analysis for this study.

The Sign

Saussure's (1959) fundamental contribution was the distinction between the two

inseparable components of the sign: the signifier (e.g., a sound or symbol) and the

signified (the concept represented). The signifier itself is meaningless, for the same

signifier can represent a different signified in a different context. Saussure also

distinguished between speech (sound patterns) and writing, seeing writing as a separate,

dependent sign system. Such a distinction was not necessary for this study, and Saussure

himself arrived at many of his principles by analyzing words and not sound patterns. I do

distinguish between mathematical language and notation, however, as there are clear

psychological differences for mathematics students. The term *multiplicative identity* and

the symbol 1, for example, do not necessarily have the same meaning.

The fundamental unit of semiotic analysis is the sign, which includes the signifier, the

signified, and the crucial connection between them. The sign, it should be recognized, is

arbitrary, in the sense that the bond between signifier and signified is essentially

circumstantial, cultural, conventional, and historical. It is tempting to conclude that

meaning is contained in the sign. But signs have no meaning except as they relate to and are distinguished from other signs:

> In Saussurean linguistics, words do not refer to things themselves. Rather, they have meaning as points within the entire system that is a language—a system, further, conceived as a network of graded differences. (Harkness, 1983, p. 5)

Thus, Saussure's "conception of meaning was purely *structural* and *relational* rather than *referential*" (Chandler, in press, emphasis in original). The structure of a language system or any system of signs depends upon its network of graded differences. The network is built from semantic distinctions that create concepts, for a concept is not a concept until its boundaries are specified. These semantic distinctions, as well as the supporting phonological, syntactic, and symbolic distinctions, are ontologically arbitrary, as evidenced by the fact that translation between languages is sometimes problematic. Anthropological linguists Sapir and Whorf found, for example, that "Eskimo has many words for snow, whereas Aztec employs a single term for the concepts of snow, cold, and ice" (Encyclopaedia Britannica, 1999). In other words, different languages provide for different concepts. This observation puts a twist on Shakespeare's aphorism "a rose by any other name would smell as sweet." The validity of the statement depends, after all, upon a language that distinguishes roses from objects that smell less sweet, and also distinguishes "smell" and "sweet" from related concepts.

Semiotics is concerned not only with what signs mean but with *how* signs mean what they mean (Sturrock, cited in Chandler, in press), which requires studying the structural relations among signs, as mentioned above, and also the relationship between signs and interpreters. But what is meaning? And where is it? There is a long history of philosophical debate about the meaning of *meaning* (see, e.g., Zemach, 1992). For the

purposes of this study, I point out one particularly influential approach, proposed by

Wittgenstein (1973):

> For a *large* class of cases—though not for all—in which we employ the word
> "meaning" it can be defined thus: the meaning of a word is its use in the
> language.
>
> And the *meaning* of a name is sometimes explained by pointing to its *bearer*.
> (pt. 1, sct. 43)

Wittgenstein's solution contrasts sharply with that of Saussure (1959), for whom "the

linguistic sign unites, not a thing and a name, but a concept and a sound-image"(p. 66).

For Saussure, meaning was a psychological phenomenon in that the signifier was a

mental representation of sensory impressions and the signified was also a mental

construct. Wittgenstein's statement, on the other hand, has a decidedly social or cultural

sense, a point of view that fits with Matthews (2000), who argues that "meanings are in

the public domain; they have to be enculturated" (p. 171). In the analysis for this study, I

considered both psychological meaning and social meaning in mathematical discourse,

focusing, in particular, on the relationships between them.

Charles Sanders Peirce developed a semiotic theory that takes into account both the

psychological and social planes. Asserting that nothing is a sign unless it is interpreted as

a sign, Peirce (1955) proposed, "A sign, or *representamen*, is something which stands to

somebody for something in some respect or capacity" (p. 99). This is essentially a

material version of Saussure's signifier. This sign addresses somebody, creating in the

mind of that person an equivalent or more developed sign, which Peirce calls an

interpretant. These are complemented by the *object* to which the sign refers, creating a

triadic relationship. It is important to point out that there is not necessarily any direct

relationship between the representamen and its object, and in fact the only relationship may be through an interpretant, which requires an interpreter.

The semiotic models of Saussure, Peirce, and Wittgenstein are compared in Figure 2. The vertical alignment is intended to indicate correspondences, though the correspondences, particularly between the psychological and social planes, are not direct. Moreover, the fit among the models is not perfect, precisely because Saussure, Peirce, and Wittgenstein were using different categories as well as different names. In particular, an interpretant for Peirce could be either a mental recreation of the representamen or a more developed sign, perhaps approaching a concept.

Figure 2. Comparison of meanings of *meaning*
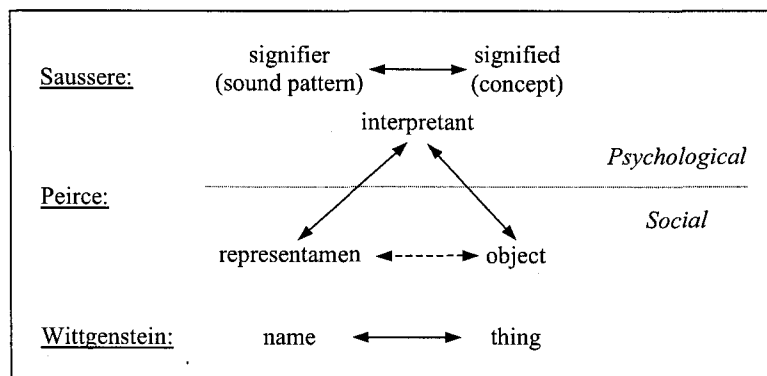
Figure 3 illustrates a semiotic model that blends each of the models described above and expands on them as well. A signifier is a symbol or a word or anything external that is taken to signify something else. A concept is a mental entity, which may be considered the core of a concept image, as described above. A referent is a mathematical object, process, or property, taken to be external in some sense. I make no ontological claims about whether or where the referent exists but say merely that it is useful in the analysis to suppose that it is distinct from the concept and from the signifier. This model of a

sign, then, is what I mean by the meaning of a mathematical word or symbol, with the added stipulation that the meaning must be considered within a system of language and symbols.

Figure 3. A general sign

Figure 3. A general sign
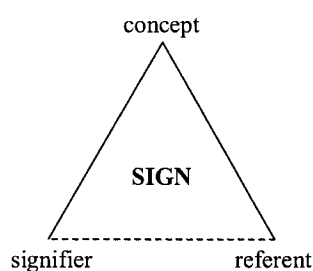
Conceptual Grids

In this study, signs provided access to the students' concept images. The students' use of signifiers provided insight about the meanings and understandings that the signs held in their thinking. The principles of semiotics make clear, however, that signs (and hence meanings) must be interpreted within a system of signs. Thus, the notion of concept image must be sufficiently flexible to pay attention to the ways that the students' concept images related to each other.

Within a system of signs, the meaning of an individual sign is determined, in large part, by its relations to other signs and, in particular, by the distinctions between it and closely related signs. The structure and categories of a system of signs lead those who use the signs to impose a conceptual grid on experience, specifying the way that the experience is cut up and hence shaping the way the experience is perceived. The crucial content in a system is the set of boundaries that are placed around and between the categories, and thus the essence of the system is independent of the particular symbols and names that are attached to the concepts it delineates. The categories, however, are not inherent in the

experience but arise in the way the world is represented in the structure of the system.

Different systems provide different categories that give rise to different concepts. Those

differences manifest themselves in the particular grid that is used to organize an

experience. A semiotician can gain insights about concepts and their meanings by paying

attention to distinctions in the use of the signs—by trying to infer the conceptual grid.

Thus, language and other sign systems play a crucial role in shaping the concepts that are

available. This position is supported by empirical work in linguistics:

> The "real world" is to a large extent unconsciously built up on the language
> habits of the group. No two languages are ever sufficiently similar to be
> considered as representing the same social reality. The worlds in which different
> societies live are distinct worlds, not merely the same world with different labels
> attached. (Sapir, 1929/1949, p. 162)

> We cut nature up, organize it into concepts, and ascribe significances as we do,
> largely because we are parties to an agreement to organize it in this way—an
> agreement that holds throughout our speech community and is codified in the
> patterns of our language. The agreement is, of course, an implicit and unstated
> one, *BUT ITS TERMS ARE ABSOLUTELY OBLIGATORY*; we cannot talk at all except by
> subscribing to the organization and classification of data which the agreement
> decrees. (Whorf, 1956, pp. 213-214)

The above theoretical perspectives led to what is known in linguistics as the Sapir-Whorf

hypothesis, which in its extreme form is called *linguistic determinism*, indicating that

language determines the framework of perception and thought. Although few linguists

accept the hypothesis in this form, its weaker formulation—that language influences

thought—is generally accepted (Chandler, in press).

Semiotics and Learning

As stated above, I consider in the analysis both psychological and social meaning,

focusing particularly on the relationship between them. What is missing from all of the

above treatments is acknowledgment that personal and social meanings may not fit.

Unless an individual makes the same conceptual distinctions as made by the community,

the conceptual grids imposed on an experience will be different. In educational research, the critical question concerns the relationship between personal and social meanings during learning.

Much of Vygotsky's work can be viewed as describing the process of personal acquisition of social meanings, which is essentially the acquisition of speech and language. In developing his research methodology, Vygotsky (1934/1986) was critical of methods of analysis that analyzed psychological processes into components, such as thought and word, to be studied separately, when what is most crucial is to understand how they operate together. In considering a method for analyzing the acquisition of language, he asked:

> What is the unit of verbal thought that is further unanalyzable and yet retains the properties of the whole? We believe that such a unit can be found in the internal aspect of the word, in *word meaning.* (p. 5)

Although it seems that Vygotsky (1978) did not explicitly draw on semiotics in his work, his perspective fits with semiotics. "The sign acts as an instrument of psychological activity" (p. 52), and by *sign* he generally meant a word, which is but one kind of signifier. By comparing the use of signs in thought to the use of tools in physical activity, he maintained that the sign and the tool both mediate activity indirectly, the tool being externally oriented and the sign being internally oriented.

Borrowing from French psychologist Paulhan, Vygotsky (1934/1986) also proposed a distinction between the *meaning* and the *sense* of a word, which are roughly its denotation and connotation, respectively:

> The sense of a word, according to [Paulhan], is the sum of all the psychological events aroused in our consciousness by the word. It is a dynamic, fluid, complex

whole, which has several zones of unequal stability. Meaning is only one of the zones of sense, the most stable precise zone. (pp. 244-245)

Note the striking similarity between this definition of *sense* and Tall and Vinner's (1981) definition of *concept image* as the total cognitive structure associated with a concept. And so, the distinction between meaning and sense may be considered to be roughly the distinction between a concept and a concept image.

Constructing personal meaning requires establishing a conceptual bond between the signifier and the referent. This necessity is well recognized in mathematics education and is seen in the metaphor of attaching or gluing names to ideas (see, e.g., Hewitt, 2001). The fact that the signifier is arbitrary and thus needs to be taught (Hewitt, 1999) fits well with many explicit and implicit theories of mathematics teaching and learning.

What is seldom recognized in mathematics education, however, is that the signified also is arbitrary, in the sense that the conceptual grid is not predetermined. In emphasizing the role of language in creating conceptual grids, Sapir and Whorf seem to have ignored the learning that is required to build the intended distinctions into one's own cognitive structure. I accept the weak version of the Sapir-Whorf hypothesis in the sense that the conceptual grids that are used by the community certainly influence, and to a great extent limit and constrain, those of the learners. But one must also recognize that students' conceptual grids do not always fit with the ones used in the mathematical community. Social and personal meanings will not match but will fit with some degree of viability. Students do not learn social meanings whole and unproblematically but instead make successive approximations, adjusted via accommodation.

Typology of Signs

In addition to his seminal contribution on the nature of the sign, Peirce (1955) also provided a detailed typology of signs. For the purposes of this study, it is sufficient to mention only his distinction between icon, index, and symbol. An *icon* bears a resemblance to its referent, "such as a lead pencil streak as representing a geometrical line" (p. 104). An *index* bears a direct connection to its referent, such as smoke to fire or, in mathematics, as a letter used in text following a figure to refer to a labeled portion of a figure. The label itself, however, is not an index. Finally, the connection between a *symbol* and its referent requires establishment by convention.

Bruner (1966, pp. 10-11) distinguished three ways in which human beings model their experience: enactive, iconic, and symbolic representations, the latter two of which are similar to Peirce's categories. *Enactive representations* embody experience in action and are, in a sense, prior to the other types of representations. Enactive representations, I would suggest, are helpful in describing the gestures that accompany certain metaphorical conceptions of mathematical ideas such as function.

Regarding the signs (or representations) in abstract algebra, it is important to point out that Peirce categorized algebraic equations as icons, in the sense that they are compound signs, composed of symbols and indices, in which the relationship of the signs to one another iconically represents the mathematical expressions and relations they are to represent. But as Peirce pointed out, a sign is not a sign unless someone interprets it as such. Thus, whether a sign is an icon, index, or symbol depends upon the individual using or interpreting the sign. Therefore, when (but not until) an individual has

established a conception of set, element, and set arithmetic in a group, the symbol aH can function as an icon for a coset.

A Semiotics of Mathematics

In a initial semiotic analysis of mathematics, Rotman (1988) identifies three aspects of mathematical discourse: the referential aspect, the formal aspect, and the psychological aspect, which have rough parallels in the mathematical philosophies of Platonism, formalism, and intuitionism, respectively. Each philosophy captures, in part,

> an important facet of what is felt to be intrinsic to mathematical activity.
> Certainly, in some undeniable but obscure way, mathematics seems at the same
> time to be a meaningless game, a subjective construction, and a source of
> objective truth. (p. 6)

Thus, through semiotics, we are back to the metamathematical issues that arose in the discussion of definitions above.

Drawing on Peirce, Rotman distinguishes between the Mathematician (the "self"), who imagines and conducts reflective observations, the Agent (a skeleton diagram and surrogate of the self), who metaphorically constructs objects and carries out processes as demanded by the Mathematician, and the Person (the subject), who operates with the signs of natural language and participates in nonmathematical discourse. The distinctions become clear in Rotman's (1988) observation:

> A mathematical assertion is a *prediction*, a foretelling of the result of performing
> certain actions upon signs. In making an assertion the Mathematician is claiming
> to know what would happen if the sign activities detailed in the assertion *were* to
> be carried out. (p. 13)

The Mathematician cannot directly verify claims that would require infinitely many operations, so she or he sets up a thought experiment in which it is the Agent who performs the necessary actions. The proof of the assertion is presented via the

mathematical *Code*, which consists of "the discursive sum of all legitimately defined

signs and rigorously formulated sign practices that are permitted to figure in

mathematical texts" (p. 15). The proof is guided by an underlying idea, which Peirce

called a *leading principle*. Discussion of neither the leading principle nor knowledge of

the Agent are permitted in the Code. Rather, they are part of the *metaCode*, which

consists of "informal, unrigorous locutions within natural language involved in talking

about, referring to, and discussing the Code that mathematicians sanction" (p. 15). Thus,

it is not the Mathematician alone but the Mathematician in the presence of the Person, the

natural language subject of the metaCode, who can be persuaded by a proof, for

conviction depends upon knowledge of both the leading principle and the actions of the

Agent.

Rotman (1988) then uses this model to provide compelling critiques of the three

mathematical philosophies, largely based upon the aspects of mathematical experience

that they ignore. I will not discuss the substance of these critiques except to mention the

Platonic nature of naming. In present-day mathematical Platonism, the principal function

of language is naming aspects of a pre-existing world—of assigning names to

prelinguistic referents. Rotman argues instead that mathematical language creates reality.

Furthermore,

> what present-day mathematicians think they are doing—using mathematical
> language as a transparent medium for describing a world of pre-semiotic
> reality—is semiotically alienated from what they are, according to the present
> account, doing—namely, creating that reality through the very language which
> claims to "describe" it. (p. 30)

Rotman's point here suggests that the preceding discussion of semiotics, particularly the

Peircian version, suffers from what might be a serious philosophical problem: the

ontological status of the mathematical object that serves as the referent in the model of the sign. In mathematical discourse, the various signifiers exist as marks on paper or perhaps merely as ephemeral vibrations in the air; the concepts exist in the minds of the students or in the collective mind of the mathematical community, as reflected in discourse. But in what sense does the object exist and, more particularly, where does it exist?

This is an age-old philosophical problem that was present in the work of Plato, Russell, Frege, Gödel, Hilbert, and many others. Rather than choosing among the various solutions to this problem, I suggest that for this study (and, I believe, for the semiotic study of mathematical cognition more generally) it was necessary only to suppose that mathematical objects exist in some sense. In particular, this assumption is all that is necessary for semiotics to be a useful analytical tool. From my understanding of philosophies of mathematics, this assumption and the general sign (Figure 3) fit with all the major philosophies of mathematics, with the exception of Hilbert's strict formalism, which maintains that the symbols are themselves the mathematical objects. In particular, this approach can satisfy both Platonistic and anti-Platonistic philosophies (see Balaguer, 1998).

The point is mathematical cognition is primarily a psychological problem, not a philosophical problem, and, as such, theoretical explanations must be psychologically plausible. In other words, psychological considerations must trump philosophical ones. Whether one supposes that mathematical objects exist in an abstract Platonic realm or exist only as fictions, whether abstract objects are created by the community or by an individual's thought, mathematical discourse—including all externalization of

mathematical thought—proceeds *as though* abstract objects exist, and thus the analysis proceeded on this basis.

Unfortunately, this assumption is not sufficient to establish the psychological and philosophical grounding of my version of Peirce's semiotics. There is also the problem of whether the concept is distinct from the object. From a psychological perspective, it is clear that the concept and the object are not identical. A concept of a rock is certainly not identical to a rock that exists as an object in the world. Similarly, it is useful to consider that a concept of the group of integers is distinct from the object that is the set of integers under addition. If the object exists physically, then there is no question that a concept of the object is distinct from the object itself. Thus, once again, independent of where, how, or even whether mathematical objects exist, it was useful in the analysis to suppose that the concept and the object are distinct.

To complete this discussion, I must address the question of whether the signifier is distinct from the referent. In the case of a rock, there is no signifier; the rock is the object and Peirce's triadic structure fails. This is not surprising, however, for Peirce's semiotics is a theory of signs, not of physical objects. One approach, due to Hilbert, is to suppose that the symbols *are* the objects, simultaneously solving the ontological problem above and rendering the current question moot. Hilbert's solution to this problem strikes me as a desperate attempt to construct a coherent, anti-Platonic philosophy of mathematics. The approach is both counterintuitive and anti-psychological, for it ignores the nature of mathematical activity and discourse. Mathematicians feel as though they are working with real objects that exist independent of the symbols and independent of their own

thought, and mathematical discourse suggests such a perspective (see, e.g., P. J. Davis & Hersh, 1981).

Nonetheless, there is sometimes a sense in which students treat symbols as though they are the objects. Nemirovsky and Monk (2000) suggest the construct of *fusion* to describe how some children behave when symbols are used to model something in the world, such as marks on a page to represent people getting on and off a bus, or when a stick becomes a horse during a child's play. This construct is not immediately helpful for noncontextual mathematics, when there is neither a physical object nor a physical activity that the student is attempting to model with the symbols. In the case of a physical object or activity, the student is always able to step back and agree that the stick is not really the horse and the marks on the page are not really people on a bus. In the case of abstract mathematics, the phenomenon is more complex. Sfard (2000) points out that a crucial event in learning about a mathematical concept is when an individual separates a signifier from its referent. At first, the symbol (perhaps an operation table of a group) is the object of thought, much as a particular rock may be an object of thought. The students begins to develop a concept of the symbol by developing some familiarity with it, perhaps relating it to other symbols, transforming the symbol in various ways, and particularly translating it to what is to be another symbolic representation of the same object. Eventually, as the student begins to see the symbol not as a thing-in-itself but as a representation, then the student has a concept of an abstract object and the Peircian triadic sign applies.

Semiotics in Mathematics Education

Because this version of semiotics and the semiotic framework below is not identical to anything currently available in the literature, it is important to point out some similarities

and differences. I have already discussed ways in which this semiotics is similar to the work of Sfard (2000) and others. Several other mathematics education researchers use Lacan's (1977) modifications of Saussure's version of semiotics to describe chains of signifiers that arise in mathematical discourse. For example, Sfard (2000) suggests that "in Lacan's writings, one finds the idea of a sign turning into a signified of another sign" (p. 45). Proceeding in this way, one can create a hierarchy of signs of increasing abstraction. The literature describes how students use chains such as candies → unifix cubes → pictured collections → verbal enumerations (Cobb, Gravemeijer, Yackel, McClain, & Whitenack, 1997) and double-decker bus passengers → beads → nonstandard notations → conventional notations (Gravemeijer, Cobb, Bowers, & Whitenack, 2000). Such chains of signifiers typically proceed from some real-world situation to be modeled, to abbreviated (iconic or essentially indexical) signifiers, to conventional symbols. The trouble with this description is that any chain of signifiers creates a hierarchy that implicitly privileges some signifiers over others. In studying advanced mathematics, where what is important is moving flexibly among the various important signifiers, including names, definitions, and several symbolic representations, one needs a framework that allows more flexibility. The Peircian approach is preferable because it allows consideration of the concept and the object separately, as discussed above, and because it allows a nonhierarchical perspective on the various signs that might come into play during mathematical discourse.

Analytical Framework

The goal of the study was to characterize students' images for concepts in elementary group theory. The preceding sections have described a number of theoretical constructs

that inform the characterizations. In particular, I have discussed the role of definitions in mathematics, the role of metaphor in mathematical intuition, the distinction between processes and objects in mathematical thinking, the distinction between abstraction and generalization in the creation of mathematical concepts, and the role of naming and notation in mathematical thinking and communication. The primary analytical tool, borrowed from semiotics, is the sign, embodied as the distinction between a signifier, a concept, and a referent.

In this section I describe how these theoretical constructs and analytical tools are brought together in an analytical framework. For the purposes of this study, I was interested in the relationship between a concept and three types of signs: symbols, names, and definitions, including informal ones. That symbols and names are signs is obvious; that definitions are signs follows from the substitution criterion described above. A primary activity of thought is replacing one representation with another, and substitution of a definition for the defined accomplishes exactly that.

The theoretical constructs discussed above are partially synthesized in the semiotic conceptual framework shown in Figure 4 for a conceptual object. It is important to note the framework is but a mere skeleton intended to highlight the main relationships for this analysis. The two front-most faces of this pyramid and the vertical cross-section through the vertices labeled concept, name, and referent each constitute a sign in the Peircian sense, in that they are triadic relationships between a concept, a signifier (a name, symbol, or definition), and a referent. Furthermore, the framework suggests consideration of mediating role (in Vygotsky's sense) of the name, symbol, or definition in mathematical activity.

Figure 4. Semiotic conceptual framework
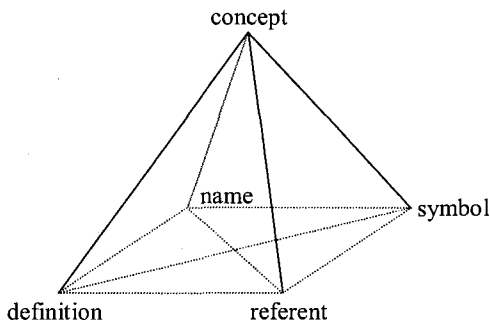
The framework in Figure 4 is not a concept image but rather an organized collection of slices of the concept image that serves as a tool for semiotic analysis. Separating these various signifiers from the referent and from each other provides lenses for looking at students' use of language and notation for the purpose of making inferences about the conceptual structures that the language and notation represents. By paying attention to structural relations among signifiers, one can gain insights on structural relations among concepts.

As for the concepts themselves, they are likely to be metaphorical in nature, and semiotic analysis can serve to reveal some of the operative metaphors. As for the referents, they may be objects, processes, properties, or some combination of these, though in view of the process/object duality of many mathematical concepts, perhaps the nature of the referent is in the relationship between the concept and the referent. Taken as a whole, this framework for analysis can be seen as an elaboration of Gray and Tall's (1994) notion of *procept*, where the semiotic nature of the analysis is made explicit. Furthermore, the analysis takes advantage of the observation that it is possible to ascertain whether students have constructed a mental object based on the way they talk and write about the concept (Tall et al., 2000).

The above framework serves to guide the analysis of individual concepts. But what about collections of concepts and the relationships among them? Here I use the metaphor of a conceptual grid that organizes experience into concepts. The grid is not in an individual's cognitive structure but rather is created because of one's cognitive structure and is then imposed on experience. In other words, a conceptual grid is not something one has but rather something one uses. There is a potential conflict, it should be pointed out, between the notion of concept image, which assumes the concepts to be primary, and the metaphor of a conceptual grid that manifests itself in the way experience is cut up and organized into concepts. Thus, one needs a sufficiently flexible notion of concept image to accommodate not only the possibility that students might have the right concepts but attach the wrong names but also the possibility of having entirely different concepts. In general, this accommodation requires an analysis that gets at the concept without the name (via an activity) and also analysis that tries to determine what is organized under that name. Furthermore, the analysis must provide for the possibility of multiple meanings in the language itself (polysemy or lexical ambiguity) and the analogous phenomenon of compartmentalization in thought, wherein an individual holds two aspects of the same concept under the same name in such a way that they are not evoked at the same time and therefore do not interact.

Summary

On the fundamental assumption that mathematical learning is meaningful learning, the ideas expressed in this chapter combine to create a conceptual and analytical framework intended to support the investigation of the meaning behind students' utterances. The notion of a concept image, as distinct from a concept definition, served to organize the

analysis of learning and using the various concepts in elementary group theory. To characterize the students' concept images, I paid attention to issues of abstraction and generalization, to the sense in which they treated the concepts as processes or objects, and to the metaphors they used explicitly or implicitly. To study the students' use of language and notation, I borrowed constructs from semiotics, focusing in particular on the sign as a relationship among a signifier, a concept, and a referent, such as a mathematical object.

The conceptual perspective evolved over the course of the study. Early versions helped to frame the initial research questions, to inform the design and implementation of the course, and to ground the research methodology and data collection. Later versions served to guide the analysis of the data. These contextual aspects of the study and the evolution of the research questions and methodology are elaborated in the next chapter.

CHAPTER IV

CONTEXT AND METHODOLOGY

The participants in this study were enrolled in a junior-level abstract algebra course at the University of New Hampshire (UNH) during the spring term of 1996. The class was taught by Dr. Steve Benson, a visiting faculty member, and I served as his teaching assistant. The curriculum was designed collaboratively by Dr. Benson and me. The instruction was unusual in that there were no formal lectures, although there were whole-class discussions at least weekly that were led by Dr. Benson or me.

This setting was chosen for the study based, in part, on my theoretical stance and my research questions. They might be paraphrased as, What does students' understanding look like in abstract algebra, and how does it build on their prior experience? Because this was essentially an exploratory study, I wanted to be able to observe some of this knowledge building in a rich, example-driven environment in which the students were encouraged to make their thinking overt and explicit.

The analysis and results are based largely on interviews with five key participants. The methods of analysis were designed to provide characterizations of the students' concept images. This chapter describes the context, curriculum, and instruction in more detail, followed by descriptions of the participants, the data sources, and the methods of analysis.

96

The Context

UNH is a land-grant institution with about 10,500 undergraduate and 2,000 graduate students. The mathematics department consists of 23 full-time faculty, 3 faculty emeriti, 4 adjunct faculty, and 34 full-time graduate students.[8] The department offers 10 undergraduate major programs: a Bachelor of Arts (BA) in Mathematics; a Bachelor of Science (BS) in Mathematics; a BS in Mathematics Education, with Elementary, Middle, and Secondary School options; and a BS in Interdisciplinary Mathematics with options in Computer Science, Economics, Electrical Science, Physics, and Statistics.

The Course

The class that provided the setting for the study was an abstract algebra course intended to be taken by most mathematics majors at UNH and required by the BA in Mathematics, the BS in Mathematics, and the Middle and Secondary School options of the BS in Mathematics Education. The course was offered in only one section in the spring term of 1996.

The class met for 50 minutes, four times per week, for 15 weeks. (See Appendix A for a syllabus.) There were two midterm exams, the first consisting of an in-class and a take-home portion and the second entirely take-home, and a two-hour final exam (see Appendix B). A standard text (Gallian, 1994) was used as a resource for examples, problems, and explanations. The bulk of the class was devoted to collaborative problem sets (classwork) and individual assignments (homework), written by Dr. Benson and me, with Dr. Benson taking the lead role. Problem sets with homework assignments were

---

[8] These data are from 2000 and are reasonably representative of the situation in 1996.

distributed (sometimes separately) approximately every week, though more frequently at

the beginning of the course (see Appendix C). Some of the classwork was completed

using the computer software *Exploring Small Groups* (Geissinger, 1989), which was

available in the department's computer laboratory. The classwork and homework were

periodically collected for comment or grading. Although I provided comments on the

students' work, I had no responsibility for grading.

Mathematical Content

This course focused on group theory, including the concepts of group, subgroup,

isomorphism, homomorphism, coset, and quotient group. This focus is in contrast to

some beginning abstract algebra courses that include introduction to rings and fields. To

provide an experiential basis for the group axioms, these concepts were preceded by

some exploratory work in number theory, particularly modular arithmetic. The course

was highly example driven, focusing especially on the following:

> Z: The group of integers. The elements are the integers, $\{\dots, -2, -1, 0, 1, 2, \dots\}$, and the operation is addition. Sometimes the operation of multiplication was also considered to illustrate the failure of the inverse axiom.

> nZ: The group of multiples of n. The elements are the integers, $\{\dots, -2n, -n, 0, n, 2n, \dots\}$, and the operation is addition.

> Z_n: the group of integers modulo n. The elements are the integers $\{0, 1, \dots, n-1\}$ and the operation is addition modulo n. Sometimes multiplication modulo n was also considered to illustrate the failure of the inverse axiom.

> U_n: the group of units modulo n. The elements are the integers in $\{0, 1, \dots, n-1\}$ that have inverses under the operation multiplication modulo n. An equivalent characterization is the integers in $\{1, \dots, n-1\}$ that are relatively prime (i.e., share no factors) with n. Thus, for example, $U_{10} = \{1, 3, 7, 9\}$.

> D_n: the dihedral group of order $2n$. The elements are the symmetries of a regular n-gon and the operation is given by thinking of the symmetries as transformations and composing them; that is, carrying out one transformation followed by the other. The elements of D_n were represented both geometrically (as transformations) and as permutations of the vertices.

> S_n: The symmetric group of degree n. The elements are the permutations of a set with n elements, and the operation is composition of permutations, thought of as

functions. The elements were usually represented as arrays notation or in cycle notation. For example, a permutation α, where $\alpha(1) = 3$, $\alpha(2) = 2$, $\alpha(3) = 4$, $\alpha(4) = 1$, is represented by the array

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{bmatrix}$$

where the elements in the second row indicate the images of the elements in the first row. In cycle notation, this same permutation would be written (134), indicating that 1 goes to 3, which goes to 4, which goes (back) to 1. The fact that 2 is missing implies that 2 remains unchanged. The identity permutation is denoted (1) in cycle notation, specifying explicitly that 1 goes to 1 and implying that everything else remains unchanged as well.

We also considered the real, rational, and complex numbers; sets of matrices; and various groups and nongroups given by operation tables. During class and on problem sets, these examples were notated as sets, leaving the operation implicit (see chapter 5).

These examples were used to motivate the concepts treated in the course, which included group, subgroup, isomorphism, center, centralizer, order of an element, cyclic subgroups, subgroups generated by elements, homomorphism, coset, Lagrange's theorem, and quotient groups. These concepts are described in more detail in the description of the problem sets below.

Instruction

As was stated above, this class included no lectures. Most of the time for this class, both in class and out, was devoted to working on activities and problem sets designed by Dr. Benson and me. The students worked through most of the activities and problem sets collaboratively, usually in groups of three or four, although some assignments, particularly the take-home exams, were to be completed individually. During class time, Dr. Benson and I worked with the groups and periodically brought the whole class together to discuss common issues, to encourage synthesis of the various results, and to point toward important themes and ideas. Both individually and when working

collaboratively, the students were expected to justify their claims. In this way, student thinking was expected to be overt and explicit. Both Dr. Benson and I held office hours, both regularly and as needed, and students attended both individually and in groups, usually to get help with specific problems on the problem sets.

Aspects of the theoretical perspective described in chapter 3 implicitly and explicitly informed the instruction and the design and implementation of the problem sets. In particular, Dr. Benson and I tried to pay particular attention to what the students were thinking because what they learned might not be what we intended. The class was somewhat like a teaching experiment in that our planning tried to take into account the experiences, including both difficulties and insights, that students were having with previous problem sets. Moreover, because reflection is key to building strong and productive understandings, we encouraged overt reflective activity whenever possible, meaning that the students were expected to explain their thinking, orally or in writing, to us or to other students.

Problem Sets and Homework

The problem sets were designed to provide experience with examples that could be used to motivate the key ideas. Often, concepts were introduced not by a definition, statement, or theorem, but by a problem. Then, as students developed solutions, key features or properties of the problem were drawn out, defined, and given standard names and notations. Often the key terms, definitions, and notations were provided again in subsequent problem sets. In this way, the students might see some of the concepts as growing naturally from the problems they were trying to solve. A sampling of the problem sets is included in Appendix C.

The course initially focused on modular arithmetic, which was used the primary example of a system in which to solve equations. For example, the students were asked to find solutions of $3 + x = 5 \bmod 7$, $3x = 5 \bmod 7$, and $3x = 5 \bmod 6$, and to investigate when such equations had a unique solution, no solutions, or multiple solutions. The students also spent time solving equations of the form $ax = b$, $a + x = b$, or $a*x = b$ in other mathematical systems such as subsets of the real numbers, sets of matrices, and also in finite systems for which the operation was given by an operation table. The group axioms were then presented as a generalized consolidation of what the students suggested needed to be true about a system in order for such equations always to be solvable.

The students were asked to find all possible Cayley tables with 2, 3, and 4 elements in order to motivate the ideas of isomorphic groups, which they initially called *congruent groups*. The isomorphism itself was not explicitly a function, at first, but instead resulted from a renaming process based on looking at the group table.

In order to provide experience for making sense of addition and multiplication of cosets, set addition (and multiplication) were introduced early through examples such as $\{1, 3, 4\} + \{2, 6\} = \{3, 7, 5, 9, 6, 10\}$ and by comparing the sets $3Z, 3Z + 1, \dots, 3Z + 7$. Later, the students were asked to make operation tables for $\{0, 4, 8\}$, $\{1, 5, 9\}$, $\{2, 6, 10\}$, and $\{3, 7, 11\}$ in Z_{12}. And to motivate the usefulness of the normality of a subgroup after introducing the concept of coset, the students performed coset arithmetic at first without concern for whether the subgroup was normal. Additional detail is provided in chapter 6.

Participants

All 29 students enrolled in the class were participants in the study for the purposes of field observation. All were mathematics majors: 24 of them were juniors, and 22 were pursing a Bachelor of Science in one or more of the Mathematics Education options. This high concentration of mathematics education majors is typical in the spring semester offering of this course. They had previously taken an average of seven mathematics classes, typically including a four-course calculus sequence, courses in mathematical proof and statistics, and another course such as geometry or linear algebra. Because the mathematical proof course is a prerequisite for abstract algebra, it is reasonable to assume that all students had taken it previously, although two did not list it on the questionnaire distributed on the first day of class. Of the 29 participants, 25 allowed collection of their written work, 21 were willing to be interviewed, and 19 consented to both. Blank consent forms and Institutional Review Board Approval are provided in Appendix D.

Almost all of the students who completed course evaluations said they found the problem sets and the collaboration helpful in their learning. Some students even said they found the take-home exams particularly helpful.

Key Participants

For key participants, I wanted students who might be considered typical students in the course. I did not want students who were struggling so much that the interviews would not be able to reveal their understanding of the key ideas in the course. On the other hand, I did not want students for whom many of the abstractions and generalizations were quick and obvious. Thus, based on discussions with and observations of the students over the first two weeks of the course, I chose six students who had given permission for

full participation and whom I expected to perform at an average level in the class. As it turned out, their grades were above average, with 3 A, 2 A-, and 1 B in the following distribution: 13 A, 4 A-, 9 B, 2 B-, and 1 D. The study was based primarily on an analysis of the interviews with the five students for whom I was able to collect a full set of interviews: Carla, Diane, Lori, Robert, and Wendy.[9] These students are described briefly below.

The data for these sketches come from three sources: questionnaires distributed to all the students at the beginning of the course, moments in the interviews when the students chose to describe themselves, and conversations with the students after the completion of the class.

Carla. Carla was a junior and was majoring in mathematics education in both the middle and secondary school options. She planned to teach eighth grade. She had taken seven college mathematics courses previously, including calculus, and was taking linear algebra concurrently. Two of these courses were among those taken primarily by preservice elementary teachers.

Carla described herself as follows: "I am a visual learner. So I remember, like, a sequence of letters if I see them" (Interview 1, line 34). She said she was a very successful mathematics student, though she admitted mathematics had not always been easy for her. When a mathematics course was very challenging, she often looked back later and appreciated the struggle. Looking back on this abstract algebra class, she indicated that the class had been stressful and she did not like the fact that that Dr.

---

[9] The key participants have been given pseudonyms that preserve their gender.

Benson and I often answered questions with questions. Nonetheless, she felt that she had learned a lot.

<u>Diane</u>. Diane was a junior majoring in mathematics education under the secondary school option. She had taken six mathematics courses previously, including calculus. She hoped to teach high school mathematics, particularly algebra and calculus.

On her questionnaire, Diane expressed some apprehension: "Because of the approach of this class I'm a little concerned with how I'll do. It's different working through a problem, exploring possibilities and then reaching some conclusions, not just begin told that something is right and here's how to do it." During her second interview, she indicated some frustration with the exploratory approach, which she described as "just playing around with it like this. There has got to be a better way" (line 101). After all, "this is math. There are always rules to follow, and it's always very neat.... But this has already been established somewhere, so I know there's rules" (lines 235-240).

<u>Lori</u>. Lori was a junior pursuing a BA in mathematics and not intending to teach. She was repeating the course. She had been advised to take abstract algebra in the fall because of a perception that the sections offered in the fall were typically geared more toward the mathematics majors and those in the spring were geared more toward the mathematics education majors. Because she had received a poor grade in the fall, she was taking the course again in order to improve the grade on her transcript. She had taken five other mathematics courses previously.

Lori was the weakest student among the key participants. She was the only one who received a B in the class; all of the others received an A or an A-. She indicated that she

appreciated some aspects of the approach of this class: "I don't think that I necessarily understood the concept of closed until we made charts and tables and stuff, and we never made tables last semester" (Interview 1, line 5).

Robert. Robert was a junior and a mathematics education major under the secondary school option. He had taken five mathematics classes previously, including calculus. He intended to teach high school mathematics and was considering a graduate degree in science education.

Robert claimed on his questionnaire that he had had uneven success in his mathematics classes: "I find that I do all right w/ computational math, but I find the theory classes very difficult. Perhaps due to my lack of intuition." When struggling with an unfamiliar problem or concept, he said that he often look at various texts, examples, definitions, to "see if I could make heads or tails out of it, which, typically, I probably couldn't. It's written in mathematics, not English" (Interview 1, line 162).

Wendy. Wendy was a senior and was majoring in mathematics education in the elementary and middle school option. During her program she had decided that she would prefer to teach secondary school, so she was planning to attend graduate school for secondary certification after graduation. She had taken thirteen mathematics courses previously, including calculus. Three of the courses were among those typically taken by preservice elementary teachers.

After the class was over, Wendy indicated that her favorite part about the class was all the writing. She explained that she got a lot out of doing the problems and said she got even more out of explaining the problem and trying to write her solution carefully.

Instructor

Dr. Steve Benson is a mathematician who was a visiting faculty member at UNH and who, at the time, had recently decided to devote his attention to mathematics education. He received his doctorate from the University of Illinois in 1988 and then held a two-year postdoctoral teaching position at St. Olaf College. He taught at Santa Clara University and another year at St. Olaf before coming to UNH in the fall of 1995. His research area was algebraic number theory. While working in mathematics, he published two research papers, two expository papers in journals of the Mathematical Association of America, and one paper on teaching suggestions in abstract algebra (Benson & Richey, 1994).

Dr. Benson had a temporary faculty appointment in the mathematics department at UNH, which he saw as an opportunity to learn about and begin working in mathematics education by interacting with the faculty and graduate students in the Ph.D. program. His teaching was always a priority in his work, as evidenced by a graduate student teaching award at the University of Illinois and by consistently excellent teaching evaluations. In fact, teaching was a primary reason that he pursued and accepted the postdoctoral position at St. Olaf College, which is known for valuing and encouraging quality teaching. He had taught the content of this abstract algebra course three times before, although previously his approach had been more traditional.

Teacher/Researcher

In my interactions with students, I played two roles—teacher and researcher—which brought both opportunities and pitfalls. Ball (2000) suggests that such an approach "offers the researcher a role in creating the phenomenon to be investigated coupled with the capacity to examine it from the inside, to learn that which is less visible" (p. 388).

Simultaneously, the dual roles create challenges with respect to the validity and generalizability of the results, and particularly with respect to causation. Here I pause to discuss how I managed and coordinated these roles.

By assuming the roles of both teacher and researcher in this study, I gained considerable inside knowledge. In constructing the problem sets, I provided some of the ideas and served as a sounding board for some of Dr. Benson's ideas. Like Dr. Benson, I helped and guided groups of students as they worked in class on their problem sets. I led some of the whole-class discussions, and I provided office hours in which students sought extra help. In this way, I provided another pair of ears and eyes to help Dr. Benson learn about the students and their thinking. These duties not only provided detailed knowledge of the context for the interviews but also helped me get to know all the students much better than if I had merely observed from the back of the class and selected a few for interviews.

In class as well as in office hours, the students' thinking was expected to be explicit and was valued, no matter how nascent. Dr. Benson and I rarely told students that they were right or wrong, an approach that served to encourage their own thinking and discussion, although, in retrospect, our implementation of this approach may have been too extreme, as it occasionally led to unproductive discussions (see Chazan & Ball, 1999). It is plausible that the atmosphere we had created in the classroom was partially responsible for the fact that key participants often required little prompting in the interviews.

Many of the pitfalls of being a teacher/researcher arise when the purpose of the research is to study teaching. The problem is gaining sufficient objectivity to ensure the reliability of observations and the validity of conclusions about one's own thoughts and actions.

Such pitfalls were not present in this study, however, because the purpose was to study learning.

Some of the challenges arise in any research that relies on cases. In such research, generalizations must be made carefully and are often only tentative. Any generalization depends upon the extent to which the case resembles other situations. Yet generalizability depends also upon the nature of the claim. When the goal is to establish an existence proof or theory building, as was the case for this study, generalizability is determined outside and after the study and thus is not really an issue.

The most potentially problematic issue for a study such as this is the evaluative role of the teacher. Ball hints at the issue when asking, "What might [the students] not want to say to her? What might it be risky to disclose?" (p. 389). Although it is plausible that this issue is not serious when the students are third graders, as in Ball's research, I am quite convinced that it merits careful consideration when the students are undergraduates. This is why I chose to make it clear to the students, with Dr. Benson's support and assistance, that I was to play no direct evaluative role in the class. I was particularly fortunate that Dr. Benson was comfortable with this arrangement.

The goal of this study was to describe student thinking and, to the extent possible, to build theoretical explanations for the descriptions without necessarily attributing cause as part of the explanations. With such goals, validity and reliability are ensured during analysis through the constant comparative method, as described under methods of analysis. In summary, the conduct of the study and the methods of analysis were designed to take advantage of the opportunities and mitigate the pitfalls of my dual role as teacher/researcher.

Data Sources

The primary source of data was the interviews with five key participants. Some exams and other written work were also collected, including final exam and second midterm papers from all the students, to provide a broader view of the ways students understood the material. Contextual data were provided through a questionnaire distributed on the first day of class (see Appendix A), from field notes I took on 20 occasions, from the problem sets and explanatory handouts, and from audiotaped planning discussions with Dr. Benson.

The Interviews

The interviews took place outside of class, and the students were compensated for their time. All interviews were simultaneously videotaped and audiotaped to aid subsequent transcription and analysis. Each of the five key participants took part in four interviews organized roughly around mathematical content, as described below. To provide some perspective on how the students were working together in the collaborative setting, Lori and Diane were usually interviewed together. (Their third interviews occurred separately because of scheduling difficulties.) Thus, a total of 17 interviews provided the core of the data. The interview schedule is given in Table 1. Two of the interviews took place after the final exam, which was administered on May 10, 1996.

The interviews were intended to address the initial versions of my research questions, which all fit under the guiding question, "In what ways do these students understand the mathematical content of the course?" The interviews were not highly structured but rather were exploratory and contingent. To provide sufficient data on each of the key concepts in the course, the four interviews were organized around mathematical content:

(1) groups and subgroups, (2) isomorphisms, (3) homomorphisms and cosets, and (4) quotient groups. Each interview typically began with a common question and proceeded from there, guided by the student's responses.

Table 1. Interview Schedule

Diane and Lori	Interview 1	04/05/96
Robert	Interview 1	04/05/96
Wendy	Interview 1	04/06/96
Carla	Interview 1	04/12/96
Robert	Interview 2	04/15/96
Diane and Lori	Interview 2	04/17/96
Wendy	Interview 2	04/18/96
Carla	Interview 2	04/24/96
Diane	Interview 3	05/01/96
Carla	Interview 3	05/02/96
Diane and Lori	Interview 4	05/03/96
Carla	Interview 4	05/07/96
Robert	Interview 3	05/07/96
Wendy	Interview 3	05/08/96
Robert	Interview 4	05/09/96
Lori	Interview 3	05/13/96
Wendy	Interview 4	05/13/96

The interviews were opportunities for me to observe the issues that the students were struggling with during their early learning of these new concepts. Thus, most of the interviews were conducted during the several days after which key concepts had been introduced, sometimes immediately following the class. My aim was to try to understand students' utterances as sensible and meaningful from their individual perspectives.

During all my discussions with students (during interviews, office hours, and class), my predominant method was to pose problems, ask questions, and encourage students to explain their thinking, so the students were accustomed to nondirective interaction. During the interviews, however, because I was trying to understand students' understanding, I was typically more probing and less directive than in class or in office

hours, at least until I thought I understood what a student was saying. Then, when I did move on, I typically posed a leading question intended to provide opportunities for the student to correct errors or make new connections among new and old ideas.

The interview tasks and questions were essentially of three varieties: tasks from the literature, open-ended questions such as "What is a homomorphism?" intended to get at the meaning the student had developed, and questions intended to probe the key concepts through standard examples. The key questions and topics in the interviews are given below.

Interview 1: Groups and subgroups. The first interview began with the question, "Is Z_3 a subgroup of Z_6?" During the students' responses, I paid particular attention to the role of the operation. When the students had resolved the opening question, I asked them to find subgroups of Z_6 and then to compare those subgroups with Z_3 and Z_2 to look for the beginnings of the concept of isomorphism.

Interview 2: Isomorphisms. The second interview approached the concept of isomorphism by comparing different groups of order 4, beginning with the four operation tables the students had identified on their take-home exam. Carla's second interview was largely about the concepts of function, domain, and range, prompted by discussions during the class that had preceded the interview. With Robert, we began with a follow-up to the first interview and spent the remainder of the interview representing the elements of D_4.

Interview 3: Homomorphisms and cosets. The third interview began with the question, "What is a homomorphism?" and I asked for examples. Then, I gave the students a

homomorphism from U_8 to Z_4 and asked how they would check whether it was a homomorphism. After checking a few specific examples, I asked them to find the kernel of the homomorphism and the cosets of the kernel. I asked Robert also to find the cosets of the subgroup generated by 3 in Z_{12}. I asked Wendy to try to make a group out of the cosets.

Interview 4: Cosets and quotient groups. The fourth interview was based on comparing the cosets of the subgroup generated by (12) in D_3 with the subgroup generated by (123). The students computed right and left cosets and then tried to construct a group using the cosets. Wendy also computed cosets and the quotient of $4Z$ in Z, and Carla also constructed the cosets and quotient of $\{0, 3, 6, 9\}$ in Z_{12}. Much of each interview was spent sorting out the students' uses of the terms *coset*, *normal*, and *quotient group* to describe the results of their calculations.

Conventions in Transcripts and Figures

All the interviews were transcribed. In the transcripts, I tried to capture all abandoned phrases and restatements, although "ahs" and "ums" were mostly ignored. Because I wanted the analysis to be guided as much as possible by complete thoughts, I chose the paragraph as the smallest unit of coding, although I refer to these paragraphs as "lines" in the transcripts and provide line numbers for all direct quotes. In order to improve the coherence and completeness of paragraphs in the transcripts, I did not interrupt a statement from one speaker to insert inconsequential statements such as "Okay" from another speaker when the statement seemed to have no effect on the train of thought.

Instead, I inserted such statements inside the paragraph of the primary speaker, enclosing

the statements in square brackets to signal the change of speaker.

Numbers were written as numerals in the transcripts except when the use of the numbers

did not seem relevant to the mathematics. In long lists of numbers, semicolons were used

to indicate slight pauses. The notations $\times$, $+$, $-$, and $=$ were used only for the words *times*,

plus, *negative*, and *equals*, respectively; similar expressions such as "added to" or "is

equal to" were written out as words. Notations for standard groups were used throughout

the transcripts, so that "Zee six" was transcribed as "Z_6," for example. Functional

notation was used when the argument of the function seemed clear, so that, for example,

"f of x" was transcribed as $f(x)$. Set notation was used when the context or the written

work suggested, either explicitly or implicitly, that the students were thinking about sets.

Similar conventions were used for permutation notation, transcribing "one two three" as

"(123)," for example. These transcriptional conventions helped me read and analyze the

data more fluently than if I had written out each word in full. Importantly, each of these

conventions is reversible by reading the transcript aloud. The students' written work was

typeset as figures rather than scanned, on the judgment that the essential characteristics of

that work could be more clearly conveyed this way. Thus, it is my conviction that these

conventions improved the clarity of the transcripts and the written work without

influencing the data or analysis by imposing notation inappropriately.

Methods of Analysis

Essentially three types of analysis were employed: detailed analysis of each interview

transcript; global analysis to confirm, refine, and refute the initial hypotheses generated

by the detailed analysis; and conceptual analysis of the mathematical content. Because both the research questions and the methods of analysis evolved as the study progressed, I begin this section with discussion of the fits and false starts that led to the methods. Then I provide a detailed description of the methods and the ways that each type of analysis informed the others. I close with a discussion of the relationship between the methodology of this study and the methodology of constant comparison and grounded theory (Glaser & Strauss, 1967).

Evolution of the Method

In my proposal for this study, the research questions included the following: "In what ways do these students understand the mathematical content of the course? How do these understandings emerge from their experiences?" These broad questions were sufficient to guide the interviews, but, as will become clear, they were initially unhelpful in guiding the analysis because they neither suggested a scheme for coding nor helped me decide what to look for in the transcripts.

Before coding any of the data, I developed a preliminary coding scheme that included categories of mathematical content, such as coset and commutativity; categories from the research literature, such as the proof schemes of Harel and Sowder (1998); categories of student action, such as choosing an example or giving a justification; categories about affect, metacognition, and the nature of mathematics; and categories that described the types of errors that students made, along with categories that described how errors were resolved. The scheme was, from my perspective, exhaustive (and exhausting), including all possible dimensions and aspects of mathematical experience that I could imagine might be present in the interviews. My attempts to use this scheme to code the transcripts

statement by statement proved unsuccessful not only because the scheme was unwieldy but also because the salient portions of a transcript were typically extended exchanges that fell entirely under one or two codes. Simultaneously, other exchanges were straightforward calculations that were not particularly interesting. The fact that both kinds of exchanges received equal emphasis in the coding was clearly not satisfactory. Furthermore, large portions of the scheme did not seem pertinent to the available data.

I temporarily abandoned coding and instead carried out a detailed analysis of each transcript. By comparing the audiotape with the student's written work and, when necessary, with the videotape, each transcript was annotated to clarify the referents of pronouns and what the student was writing. At the same time, the transcripts were segmented into episodes, providing both a chronology and a table of contents for each interview. Additional annotations were inserted to highlight episodes, events, and statements that struck me as interesting or significant, typically because of the use of nonstandard language, an error that seemed nontrivial, a hint of an unusual way of thinking, or a change that suggested learning. Guided by very general questions such as "What was the student doing? What was the student using?" I developed short descriptions of these interesting events. The table of contents and the significant events, together with my description, provided an initial "bottom-up" analysis that also served as a summary of the interview.

During the above processes, the research questions evolved, eventually arriving at questions such as, "What concept images do students demonstrate as they are learning the fundamental ideas of group, subgroup, and isomorphism?" Using the summaries of the interviews, I began the next phase of analysis with an eye toward answering the research

questions. Intending to delineate various components of concept images, I developed a coding scheme that was short, focused on describing concept images, and more relevant to the data. The scheme had categories such as representations, properties, and examples of concepts, as well as definitions, results, and associations about concepts. Once again, however, when I tried to code the transcripts, I had trouble making the scheme fit. It became clear that the most salient features of the interviews were issues of language, notation, and meaning, and the relationship between signs and the concepts that they were intended to represent. These issues still were not sufficiently prominent in the coding scheme.

I again abandoned explicit coding. Reviewing the interview summaries, I instead asked directed questions such as "What can I say about this student and her concept of group?" that led to answers such as the following: "She reasoned from the table; she confused related words; she used idiosyncratic language and syntax; she was confused about the operation in Z_n." The resulting long list of observations about student thinking was then examined for emergent themes. In continuing the episode-by-episode analysis and synthesis, I elaborated the observations with examples of dialogue from the interviews, regularly asking myself, "What is this an example of?" thereby keeping the goal of describing student understanding at the forefront of my thought. As is described in detail below, I also looked for regularities and overarching themes that could be developed into theoretical explanations.

In summary, the method of analysis evolved from line-by-line coding to detailed description of significant episodes and events. Another way to describe the transition is as follows: The unit of analysis was originally the concept, as indexed (not in the Peircian

sense) by the concept name. The students' language use was so unusual and idiosyncratic, however, that it became clear that the unit of analysis needed to be the episode. This change is analogous to Wertsch's (1985) observation that although Vygotsky began with the word as his unit of analysis, many of his colleagues and students (e.g., Davidov, Leont'ev) moved to using the activity as the unit of analysis (see also Wertsch, 1981).

Detailed Description of the Method

The goal of the analysis was, of course, to provide answer to the research questions, which meant describing students' concept images for the key concepts in the course and also describing the ways that preliminary mathematical ideas came into play. In the analysis, I considered both personal and conventional meanings of the concepts and focused on the differences between them, for that is where clues to learning problems lie. Thus, the main fodder for the descriptions of students' concept images was episodes, events, and statements that struck me as significant because of potential differences between personal and conventional meanings. Events were pursued through detailed analysis when my observations about the event seemed sufficiently robust, such as when similar events occurred elsewhere with the same student or with a different student. In this section, I describe some technical and theoretical aspects of the method and also provide additional detail.

Most of the data were managed via *N5*, the fifth major revision of *NUD*IST* qualitative research software (QSR International, 2000). The annotated transcripts and their summaries were imported into *N5* along with excerpts from the midterm and final exams of the five key participants. To provide some context for the interviews and exams, the

discussions with Dr. Benson were summarized, including some verbatim transcription, and imported into *N5*. The field notes were also imported. All electronically available data were coded for mathematical content. In particular, I coded for the following concepts: modular arithmetic; function; binary operation; properties of operations, including the four group axioms and commutativity; group; subgroup; isomorphism; homomorphism; kernel; coset; normality; and quotient group. These mathematical categories formed the primary headings under which I sought to describe students' concept images.

In trying to create descriptions of students' concept images, I found that one of the most puzzling aspects was explaining or even describing students' idiosyncratic and seemingly inconsistent use of language and notation. I was initially at a loss and for a long time found little in the mathematics education literature that helped me understand the students' statements and actions. Eventually, I was led to literature in linguistics, philosophy, and particularly semiotics, from which I borrowed and adapted theoretical constructs that helped explain what I saw and that led to a theory that ultimately connected back to the mathematics education literature.

With these additional theoretical constructs, the analysis of the episodes became essentially semiotic in character. Although semiotics holds no widely shared theoretical assumptions or methodologies, a consistent feature is looking beyond specific signs to discern the relationships between signs and the systems of distinctions operating within them (Chandler, in press). For analyzing student thinking, the approach might be described as looking *at* the students' language rather than *through* it (Sfard, 2000). Lacking direct access to the personal meanings of the students, I relied on semiotic

analysis to help me make inferences about those meanings and to build a theory that fit the data. In the detailed analysis of the significant episodes, I tried to discern meaning in students' utterances and tried to understand their use of mathematical signs, particularly words, notations, and definitions, using the analytical framework described in chapter 3. Not all significant episodes were so analyzed; instead, I focused on those episodes that either spoke to the use of language, notation, and representations; suggested consideration of relevant objects, processes, or metaphors; or raised issues of abstraction and generalization.

To complement the detailed analyses of episodes, I also took a global view, searching for additional uses of words or notation that might confirm, refine, or refute the working hypotheses. For example, I coded and collected the various formal and informal definitions that the students gave of the key concepts in the course. Some of these were in response to very direct prompts such as "What is a homomorphism?" as in the interviews, or "Provide complete definitions for the following terms and phrases" as on the final exam (Appendix B). Other definitions arose without a direct prompt, typically as part of an explanation of something else.

In the global analysis, I used *N5* to search the transcripts and other electronically available data for other instances of the signs (i.e., words and notations) that the students and I were using to discuss the particular concept. On the basis of my familiarity with the data and with the aid of the interview outlines, I also carefully examined portions of the transcripts that were likely to speak to the particular ideas under analysis. The excerpts identified by these searches were considered first for relevance and then for fit with the emerging hypotheses, which were modified to accommodate data that did not fit.

Because the study aimed to characterize students' concept images, the hypotheses were often about a nonstandard conception that a student had during a particular episode. Sometimes the search identified excerpts that suggested that later in the course the student had developed a conception that fit better with standard mathematical usage. I saw such excerpts not as disconfirming evidence of a hypothesis but rather as partial evidence of learning.

The detailed and global analyses produced preliminary descriptions and representations of students' concept images for the key concepts in the course. These were compared with what I took to be standard language usage and descriptions of the concepts in the mathematical community. This comparison was implicit, at first, in the sense that during the analysis, I was particularly interested in language or understanding that did not fit with my own, which I took to be a fair representation of standard mathematical usage. Because both my interview technique and the method of analysis took as a guiding principle the pursuit of that which was interesting, unusual, or unexpected, many discrepancies with standard usage were explored in detail during the interviews themselves, thereby providing substantial supporting data to confirm or disconfirm both the implicit hypotheses that I was generating during the interview and the related hypotheses I was developing during the analysis.

To make explicit the concept images that implicitly guided my analyses of student thinking, I also completed conceptual analyses of the key concepts in the course, as described below. The various analyses were conducted iteratively. By reflecting on the students' statements, I was often better able to conceptualize and articulate what the conventional concepts are and the distinctions between them. Conversely, with a detailed

conceptual analysis I was better able to characterize students' concept images. In this way, comparisons between students' personal meanings and accepted mathematical meaning became increasingly explicit. Because these versions of the accepted mathematical meanings were largely my own creation, some explanation is in order.

In mature discourse, particularly within a professional community, meaning is often "taken as shared" in the sense that individuals converse as though their personal meaning is shared by the community (see, e.g., Cobb, Yackel, & Wood, 1992; Ernest, 1991). Because no one has direct access to the shared meaning of the mathematical community, it was not possible to import conventional concepts directly into my analysis. Furthermore, traditional mathematical exposition would not have been appropriate for this study, because, as Pimm (1995) observes, mathematicians use words as though they are the concepts, as is apparent in mathematical discourse, and symbols as though they are the objects, as is revealed in the metaphor of manipulation. Instead, I created descriptions of the mathematics based on a conceptual analysis that aimed at careful semiotic description of the meanings of the words and representations of the mathematics under study. Guided by my own thinking, frequently consulting resources such as mathematical texts (e.g., Gallian, 1994; Herstein, 1975; Hungerford, 1974), and with careful consideration of accepted formal definitions, I arrived at a particular elaboration of the meaning of a concept, highlighting its semiotic nature and including process, object, and metaphorical characterizations. I take these meanings to be shared by the community, in the sense that the descriptions fit with, though they are not identical to, descriptions I found in texts I consulted.

As the detailed, global, and conceptual analyses proceeded, my preliminary observations about student understanding were combined, reworded, and sometimes dropped, leading to working hypotheses that were categorized eventually under two broad themes: the use of language and notation and the mathematical meanings that students gave to their activity. Although my intent was originally to describe students' concept images for a list of concepts, these themes became increasingly prominent as the analysis continued, eventually overtaking the mathematical content categories in importance. Furthermore, it became clear that these emergent themes provided not components but rather characteristics of concept images. Thus, the research questions were adjusted to reflect this observation, resulting in the versions given in chapter 1.

Characterizing the Method

The working hypotheses evolved over the course of the analysis into a theory that was organized under the two themes. Analysis, synthesis, and theory generation were conducted iteratively and sometimes simultaneously. In other words, by frequently returning to the initial analyses and to the data themselves to judge the faithfulness of the emerging theory and the accompanying explanations, I established the theory in an empirically grounded way.

Very late in the process, I realized that the detailed summaries functioned as codes, the preliminary observations served as initial categories and hypotheses, and the synthesis of the working hypotheses formed the core of an emergent theory. It is now apparent that, disregarding the false starts, the method is consistent with the constant comparative method of Glaser and Strauss (1967; see Cobb & Whitenack, 1996, for a similar discussion). Theoretical constructs were developed as part of the data analysis, and the

constructs are grounded in the sense that they are rooted in the data. The inferences made while analyzing the episodes formed working hypotheses that were constantly compared to the data and modified in light of new data and analysis, and the theory emerged via this process. The methodology also emerged as I abandoned unproductive approaches and instead focused on what the data afforded. Explicit description of the method was made only retrospectively.

What was hardest about the process that eventually led to this method was coming to the realization that there was no need to be apologetic about the fact that my initial research questions were vague and that I could not stick to a coding scheme. From the start, the goal of this study had been to develop new understandings of the ways that students learn abstract algebra. When the study began, the only extant theory had been grown out of Dubinsky's (1991) APOS framework, and I suspected right away that the APOS framework missed and even obscured important issues for the learning and teaching of advanced mathematics. Now it is apparent that my aim all along was theory generation, which is precisely what the constant comparative method is intended to support.

As for the coding, by thinking up the coding scheme in advance, the subsequent attempts at statement-by-statement coding required that I impose (or force) preconceived categories onto the data. The coding did not work precisely because the codes did not fit the data. What I should have done instead was let the codes and categories emerge from the data, and that was the end result, despite the several dead ends that were explored.

To be precise, however, not all of the codes emerged from the data. In particular, coding and categorizing by mathematical content was intended in the early conceptualization of the study and remained important throughout. The fact that these categories were

imposed on the data seems reasonable because they are natural in a sense, because I was

interested in the learning of specific mathematical content, and because these categories

organize important learning goals. Thus, because of these preconceived categories, the

resulting theory is not entirely grounded in the sense of Glaser (1992).

Relationship with Grounded Theory

I have mentioned that the method of this study is consistent with the constant

comparative method of Glaser and Strauss (1967), but that seminal book is more often

cited (with little detail) for the methodology of grounded theory. In this section, I fill in

some of the oft-missing detail and explain the relationship between the constant

comparative method and grounded theory. This discussion is particularly important

because Glaser and Strauss themselves later disagreed about the methodological

requirements of constant comparison and grounded theory (compare Strauss & Corbin,

1990; Glaser, 1992). I follow Glaser's account because it seems to me to be more faithful

to the notions of groundedness and emergence.

The constant comparative method forms the methodological backbone in the

development of grounded theory. Regarding the formulation of a research problem,

Glaser (1992) suggests, "Remember and trust that the research problem is as much

discovered as the process that continues to resolve it" (p. 21). As for reviewing the

literature, Glaser dictates that the researcher *not* review any of the relevant literature in

the field of study (p. 31), because the theoretical constructs in the literature may

contaminate the analysis, steering the researcher toward imposing preconceived

categories on the data. Any theory that grows through the constant comparative method I

will call an *emergent theory*, usage that fits with the work of Glaser and much of the work of Cobb and his colleagues (see, e.g., Cobb & Whitenack, 1996).

How do the constant comparative method and an emergent theory satisfy the traditional research ideals of validity and reliability? Glaser (1992) suggests that the criteria by which to judge the theory are not verification and reproducibility but rather fit, work, relevance, and modifiability. When the goal of a study is theory generation, verification is not necessary if the theory fits, although future studies might undertake verification. Furthermore, it matters not whether another researcher would have produced the same theory but rather whether the theory fits the data, works to explain the variation in the data, is relevant to the context from which the data came, and is modifiable to accommodate the integration of additional concepts.

In the following chapters, I have tried to include enough detail in the analysis to demonstrate that these four key criteria are satisfied. The final analysis and the emergent theory for this study also essentially satisfied Glaser's prohibitions about the formulation of the research problem and the influence of the relevant literature, if the false starts and missteps are disregarded. Certainly, the original statement of the research problem was sufficiently vague, and the statement of the research problems underwent revisions throughout the process in response to what was available in the data. Regarding the review of the literature, although I read much of the literature ahead of time and did try to force some categories on the data, in the final analysis only the process/object distinction was helpful, and that formed but a small part of the resulting theory. Of course, I also used the notion of concept image, but that construct served mostly as a reminder that I

aimed to describe students' understandings broadly, and the construct carried little theoretical baggage that could have been imposed on the data.

Thus, the methodology of this study was consistent with the constant comparative method, and I call the result an emergent theory. A grounded theory, on the other hand, requires additional methodological commitments. For example, in Glaser's version of grounded theory, data analysis and data collection are iterative so that emerging theories can inform and guide subsequent data collection. I do not see this discrepancy as very serious in this study, although I readily admit that the theory could have been developed further and in more detail if I had been able to alternate analysis and data collection.

There are two senses, however, in which the design of this study could not lead to a grounded theory. The first sense concerns the imposition of codes for mathematical content, as described above. That was unavoidable. Because the goal was to understand learning in abstract algebra, it was necessary to keep the mathematical content available in the analysis.

The second discrepancy with the tenets of grounded theory is more fundamental, though it also arises from the attention to mathematical learning. In formulating a research problem, Glaser suggests that the researcher enter the substantive area wondering what the main issue is for the subjects and the processes by which it is handled. Furthermore, it is essential that issue be relevant for the subjects from their perspectives. At least in his field of sociology, it seems that Glaser hopes that the researcher's findings might actually be directly useful to the subjects who participate in the study. Thus, in grounded theory, the subjects' meanings are primary, whereas I was concerned not only with the subjects' meanings but also with the community meanings and the fit between them. As both a

researcher and a teacher, I want not only to understand students' conceptions but also to understand how to guide and direct students toward important mathematical ideas and conventional concepts. This perspective was always in the background in the interviews. In the analysis, this perspective took a different form: How might we improve the teaching of abstract algebra in particular and advanced mathematics in general?

It is hard to imagine a grounded theory that is committed to describing students' conceptual understanding and its relationship with conventional concepts as they exist in the mathematical community. In studying students' conceptual understanding, I would suggest that the criteria of fit and relevance are with respect to teachers and researchers primarily and only secondarily with respect to the students. The fact that a theoretical construct is useful for teachers and researchers does not necessarily imply that it will be directly relevant for students, although it is possible to imagine recasting some constructs in ways that might assist students in reflecting on their own thinking and learning. The point is that in judging the theory, fit and relevance for students is at most a secondary consideration. After all, who would suggest that first graders should begin the year with some lessons on assimilation and accommodation?

Summary

This chapter provides a detailed description of the context for this study, including the curriculum, the instruction, and the participants, and the methodology employed. Briefly, this study consists of a semiotic analysis of interviews with students to support the development of theoretical descriptions of their understanding and learning in elementary group theory. The next three chapters provide the results of that analysis, organized

according to mathematical content and addressing the three main research questions individually.

CHAPTER V

GROUPS AND ISOMORPHISMS

This chapter presents analysis of students' concept images of binary operation, group, subgroup, and isomorphism, which were the mathematical foci of the first and second interviews. The chapter is organized around the mathematics and thus essentially follows the chronology of the interviews. The bottom-up analysis of these interviews revealed two themes that are threaded throughout this chapter: use of language and use of the operation table. Because these themes are well illustrated by Wendy's interviews, a detailed case study of Wendy's concept images forms the bulk of the chapter, with supporting data from other students and other interviews providing corroborating and contrasting evidence. Each section begins with a description of the interview task, which is followed by an analysis of the mathematics. Then portions of the Wendy's interviews are presented and analyzed, followed by related evidence from other students and other interviews. But first, I provide a short description of Wendy's language and reasoning as an introduction to the chapter's main themes.

Wendy's Language and Reasoning

Wendy often misused words. The analysis of the transcripts of Wendy's interviews was complicated by the fact that many of her misstatements were mere slips of the tongue. She would say one thing but meant to say something else. Such an inference is clearly reasonable in two kinds of situations: when Wendy immediately corrected herself and when she restated the idea differently moments later. Because these occurrences were

129

rather frequent, Wendy's language inaccuracies were also interpreted as slips of the tongue when both the context and Wendy's typical usage strongly suggested she intended to say something else. In cases where I do not otherwise call attention to her misstatement, I enclose in brackets what I believe she intended. Not all of Wendy's misstatements were so categorized, however. In particular, her use of the words *inverse, identity, commutativity, associativity*, and *isomorphism* indicated conceptual issues that are explored in this analysis.

Wendy's images of the fundamental concepts in group theory were dominated by the use of operation tables. She often relied on the operation table to provided support for her reasoning and seemed to require that the table be visible in order to begin. The operation table played a metaphorical role in her explanations, appearing to substitute for the group in her reasoning and thinking. Wendy drew conclusions and generalizations from her consideration of the operation tables but also was constrained by her reliance on the tables and found it hard to separate her thinking from them. A related and perhaps consequential phenomenon was that Wendy often considered the group axioms individually, seldom engaging more than one of them at a time in her explanations. Wendy's use of the operation table is explored in detail below.

Groups and Binary Operations

As stated in chapter 4, the first interviews began with a question from the literature: "Is Z_3 a subgroup of Z_6?" The short answer to this question is no because the operations in the two groups are different. More specifically, because Z_3 is the set $\{0, 1, 2\}$ under addition modulo 3, and Z_6 is the set $\{0, 1, 2, 3, 4, 5\}$ under addition modulo 6, the

operations are not the same. For example, $2 + 2$ is 4 in Z_6 but 1 in Z_3. Nonetheless, the subset {0, 2, 4} of Z_6 is simultaneously a subgroup of Z_6 and isomorphic to Z_3, so there is a sense in which the answer is yes. Both of these ideas were explored in the interviews.

The literature suggests two reasons for students' difficulties with this question. First, although students think of a group as a set, they are not always sufficiently aware of the operation (Dubinsky et al., 1994). The second finding in the literature is that some students use a powerful result inappropriately, saying that Z_3 is a subgroup of Z_6 by Lagrange's theorem because 3 divides 6 (Hazzan & Leron, 1996). Because the students had not yet been introduced to Lagrange's theorem at the time of the first interview, my intent was not to explore the students' understanding of Lagrange's theorem but to explore the role of the operation in their conceptions of group and subgroup. Nonetheless, by exploring subgroups of Z_6 in the interviews, I intended to get at some of the divisibility ideas that are behind Lagrange's theorem. Before providing a detailed description and analysis of the interviews, I offer an analysis of the mathematical concepts of binary operation, group, and subgroup. The analysis is semiotic in the sense that I pay particular attention to names, notations, and other representations, particularly those that were used in this class.

Conceptual Analysis

As described in chapter 1, a group is a set and a binary operation that together satisfy four axioms (closure, associativity, identity, and inverse). The operation gives the group its structure. In other words, a group without its operation is merely a formless collection of elements. In some textbooks, this point is sometimes made notationally, but it is more common to use the set to denote the group, thereby leaving the operation implicit.

Fraleigh (1989), for example, at first uses the notation $<G, *>$ to denote the group composed of the set G and the binary operation $*$. Almost immediately he adopts the shorthand notation:

> At some point, all authors give up and become sloppy, denoting the group by the single letter G. We choose to recognize this and be sloppy from the start. We emphasize, however, that when you are speaking of a specific group, G, you must make it clear what the group operation on G is to be, since a set could conceivably have a variety of binary operations, all giving different groups.
> (p. 40)

Using Fraleigh's first notation, $<Z_n, +_n>$ denotes the group consisting of the set $\{0, 1, \ldots, n-1\}$ under the operation addition modulo n. In this class, the instructors and students adopted the shorthand, denoting the group merely by Z_n. Because the most obvious operations to consider are addition modulo n and multiplication modulo n and because the set Z_n is not a group under multiplication modulo n, it is reasonable to say that, for many mathematicians, the phrase "the group Z_n" or "the group of integers modulo n" carries the implication that the intended operation is addition modulo n (see also Gallian, 1994; Hungerford, 1974). Nonetheless, this implication was not always obvious to the students.

The operation on a set may be given in a number of ways, such as by a formula, by a table, or by inheriting an operation from a larger structure in which the set sits. In subsets of the integers, for example, the operations of addition and multiplication may be inherited from the familiar operations on integers. For the sets Z_n, however, the operations addition modulo n and multiplication modulo n are not inherited from Z because, for example, $3 + 5 = 8$ in Z, but $3 + 5 = 2$ in Z_6.

For sets with only a few elements, the table was the predominant representation of binary operations for this class. Even with sets such as Z_3 and Z_6, for both addition modulo n

and multiplication modulo n, the students typically created tables that served to support their reasoning.

Just as a group is a set with structure provided by an operation, a subgroup is not merely a subset of a group but rather a substructure, and the structure is provided by the operation of the group. General insight into the structure can be provided by Lagrange's theorem, which says that in a finite group the order of a subgroup (the number of elements in the subgroup) must be a factor of the order of the group. The converse of the theorem is false in general, as discussed in chapter 3, although it is elegantly true for Z_n: For each divisor d of n, there is a unique subgroup of order d, which consists of the multiples of n/d. In the task at hand, although Z_3 is not a subgroup of Z_6, the multiples of 2 in Z_6 are the subset $\{0, 2, 4\}$, which is a subgroup of Z_6 and which is isomorphic to Z_3, as mentioned above.

Wendy, Groups, and Binary Operations

The beginning of Wendy's interview was marked by uncertainty. She first tried to understand the question:

5   Wendy: Okay. Well on the first question I look at, is Z_3 a subgroup of Z_6? From.... Z_6 is just mod 6, right? Mod 6? So first of all I'd want to.... I am assuming Z_6 is a group if you are going to ask that Z_3 is a subgroup of Z_6.

11   Wendy: I am taking Z_6 to be integers mod 6. And I don't know what's leading me to think that. But, so, but if it is.... Can I just say, "if it is ..."?

12   Brad: Sure.

13   Wendy: A total table. It would consist of 6 items or elements, and for.... It has to be mod 6. Z_6. It has to be integers mod 6, because.... Well, we have to also figure out an operation, also, too. So now you have the elements, you know Z_6. We have to know the operation because that will be [inaudible] whether or not it's going to be a subgroup.

Wendy was unsure of what Z_6 was, what the operation should be, and whether it was a group. Nonetheless, she made some assumptions. Using the wording of my question,

she assumed Z_6 to be a group. She also assumed Z_6 to be integers mod 6, but her statement "Z_6 is just mod 6" (line 5), with its odd syntax, suggests she may have been thinking as much about the process of calculating the remainders as about the set of remainders. Later in the interview, she confirmed this impression, saying, "I'm assuming Z_6 means it's integers mod 6, which means you look at the remainders after dividing by 6" (line 32). The students' understanding of modular arithmetic is considered in detail in chapter 7.

Wendy's phrase "a total table" (line 13) suggests she wanted to create a table, but she quickly realized she would need to figure out what the operation should be. To resolve this issue, she referred back to the question at hand:

17   Wendy: So, therefore you have to find.... That would help you to determine what operation, because maybe if you tried multiplication and if Z_6 wasn't a group under multiplication then you would know that Z_3, you are not talking about whether Z_3 is a subgroup under Z_6 because Z_6 isn't a subgroup [group]. But maybe under addition Z_6 is a group and therefore you can look at the case under addition.

Thus, although Wendy's concept images of group, subgroup, and binary operation were insufficient to provide a quick answer to the interview question, her concept images were sufficient to provide general framing of the question at hand. In particular, she saw that it would be helpful to determine first whether the operation in Z_6 was addition or multiplication. The fact that she didn't say "addition modulo n" suggests that she may not have been distinguishing between addition and addition modulo n, and similarly for multiplication. I did not pursue this distinction in the interview but merely suggested that she try both possibilities. She started with multiplication.

20   Wendy: Okay. Well, Z_6 is not going to be, when I start with my chart, and I do the first row, 0 times any element is going to equal 0, so if you look at that.... Actually, okay. Let me just.... It's not going to have.... You have to.... I'll just finish it. Okay, now it has to hold four properties to be a group. Let's write these down. It has to have an

identity, an inverse, it has to be closed, and it has to be associative, which we're going to leave for last. [Laughs]

During this statement, Wendy set up an operation table and filled out the "0" row and column (see Figure 5). She also wrote down the names of the four group axioms to assist her, it appears, in the process of checking whether Z_6 is a group under multiplication.

Figure 5. Wendy's table for multiplication in Z_6

×	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3				
4	0	4				
5	0	5				

Wendy continued filling out the table, checking the identity and inverse properties as she went along:

22   Wendy: So it has an identity, which is 1, which is the identity ... which is the identity for every element, identity equals 1. But inverse ...

23   Brad: Every element identity? What do you mean?

24   Wendy: An identity means when you multiply the identity by itself, like say if you have the letter m and you multiply it by the identity i, it's going to equal m. [Writes $m(i) = m$.] It is going to give you back the same thing. The identity is ...

25   Brad: Okay. So how does that fit in here?

26   Wendy: If you look at this row, you multiply.... If I.... I am calling 1 the identity. If you multiply 1 by every element, you get the element back, get the original element back. So, like 1 multiplied by this row gives you the same row back.

28   Wendy: So Z_6 does have an identity. Now, inverse. *Inverse* means when you multiply.... If you have a number in Z_6, there has to be a number in which when you, a number so that when you multiply it, you will get the identity.

30   Wendy: So m times the inverse. I don't know how I should represent the inverse. Identity is usually.... I am going to change it so that my identity being represented as e, and then I am going to change the inverse as i. So when you multiply some number m by, it has to have an inverse i, so that when multiplied, it will equal the identity. [Writes inverse $= m(i) = e$].

This excerpt provides the first clear hint that Wendy was thinking about the identity and inverse properties in similar ways. She struggled to articulate each of the concepts and

arrived at definitions (lines 24 and 28) and notations (lines 24 and 30) that were similar.
Furthermore, her syntax suggested that to some extent the names of the concepts had
been swapped in her thinking. For example, when talking about the identity, she used the
phrase "for every element" (line 22), which is more typical when talking about inverses.
Similarly, her phrase "m times the inverse" (line 30) employs syntax more typical of
statements about the identity element. A more correct phrasing would be "m times its
inverse," which makes the dependence on m explicit. A search of the transcript reveals
that Wendy had used similar syntax for the two concepts when listing the group axioms
earlier in the interview: "It has to have an identity, an inverse" (line 20).

Wendy explicitly used the table to verify the identity property (line 26). Similarly, after
explaining her calculations for the row labeled "2," she used the table to explain how the
inverse property failed:

34   Wendy: So if you look at the second row [the "2" row], there is no number when you
       multiply.... If you take m equalling 2, if you take a number equalling 2, when you
       multiply, there is nothing to multiply by 2 to get—in mod 6, cause it has to be an element,
       to be closed, you can only work with the elements within mod 6. And I have tried every
       element, 0, 1 ... 0 through 5, multiplied by 2 to see if I can get the identity, 1, and I can't
       get it. So therefore, Z_6 is not a group under multiplication. So, I don't think we should
       look at it, check to see if Z_3 is a subgroup of Z_6 when Z_6 isn't even a group under
       multiplication.

Wendy was about to begin considering addition but stopped herself to make a comment
about multiplication:

38   Wendy: Actually, up here, in multiplication, I didn't even have to look at the second row
       [the "2" row] because if you look at 0 there is nothing you can multiply by 0 to get the
       identity element back, 1, because 0 times every element is going to equal 0.

It seems that at this point, Wendy had reduced the process of checking the inverse
property to a process of looking for the identity, 1, in a particular row in the table, for not
only was she able to see from the table that the element 1 did not appear in the "2" row,

but she also noticed that 1 did not appear in the "0" row, which provided a more immediate reason for the failure of the inverse property under multiplication. This process either provides a partial explanation for or is partly explained by the close relationship between the identity and inverse properties in Wendy's thinking. The word *inverse* was not present in her justification, however. I asked her to explain:

39   Brad: So what does that say about 0 there?

40   Wendy: 0 cannot be an element in Z_6.

41   Brad: 0. But you are saying it is an element though, because ...

42   Wendy: Oh yeah; 0 is an element in Z_6, but it doesn't have an inverse.

43   Brad: Oh, okay.

44   Wendy: Because you can't.... There's not.... When you multiply 0 by anything, you can't get the identity element. And this doesn't help. So that just doesn't seem.... Like, if you are going to have a group, you couldn't, 0 couldn't be in it. A group under multiplication, it couldn't include 0.

Thus, Wendy's statement that "0 cannot be an element in Z_6" (line 40), was a specific instance of a general principle: A group, under multiplication, cannot include 0. It is not surprising that Wendy wanted to exclude elements that did not satisfy desired properties, because this is essentially the idea behind the construction of the groups of units modulo n. In the introduction to the groups U_n, the class used a more general version of this principle: Include only elements from Z_n that have multiplicative inverses.

Wendy next began considering whether Z_6 is a group under addition. She constructed a new operation table (Figure 6), checking the axioms as she went along.

47   Wendy: Now if you look at addition, I am going to fill out the table the same way, except with addition. I'm going to just look at the remainders when divided by 6. We can see, I can see by filling out the first table [row] that the identity.... Also, I think it is a global property, that since integers, the identity is going to equal zero. That, if you take a subgroup of.... But then we are going to go into another issue, whether Z_6 is a subgroup of, in the integers. But I think if integers has an identity of 0 under addition, that Z_6 will also have the identity 0. It works.

Figure 6. Wendy's table for addition in Z_6

+	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

In class, the word *global* was often used to describe the associative property when checking whether a subset of a group was a subgroup. The term is based on the idea that if an operation is associative on an entire set, then the property must hold for any subset. The term is nonstandard, although the idea closely resembles the meaning of the more conventional phrase "associativity is inherited from the group." This excerpt shows that Wendy had expanded her use of the term to describe a similar idea for the identity property. She was correct, in a sense, in that when verifying the identity property for a subset of a group, it is sufficient to show that the identity element is in the subset, rather than showing that it serves as the identity for all elements in the subset. It is not clear, however, whether she had in mind this precise use of the word. In any case, the "global" idea was not appropriate here because addition in the integers and addition in Z_6 are different operations. Thus, this excerpt suggests imprecision in Wendy's concepts of global and of addition. These issues are explored in more detail below.

Wendy continued verifying the group axioms:

50   Wendy: So next I am going to check the inverse property. And 0 has an inverse so 0 + 1, or.... Excuse me. Since 0 is the identity we have to check that when you add 0 to 0 you get the identity 0. So 0 is the inverse element for itself. And then 1. When you multiply, when you add 1 and 5 it equals 6, but that equals 0 (mod 6) cause 6 is divisible by 6. That's pretty obvious, but.... So 1 has a inverse. 2 has an inverse because 2 + 4 = 6, which equals 0. 3 + 3 has an in-.... equals 0 (mod 6). 4 + 2 = 0 (mod 6). And 5 + 1 = 0 (mod 6). So each element has an inverse. So you know that Z_6 is a group under addition.

Here, Wendy correctly verified the inverse property by using the table to find the inverse of each element. She supported this process by making a check mark alongside each row of the table as she identified the corresponding inverse. She momentarily considered 1 as the identity but corrected this on her own. Apart from her self-corrections, she used appropriate language throughout this verification, which culminated in the statement "each element has an inverse." Wendy's language and calculations, taken together, suggest that she could distinguish identity and inverse properties according to the conventional meanings, although the distinction became less clear again later in the interview.

Although Wendy's verification of the inverse property was essentially correct, she was premature in declaring that Z_6 is a group under addition because she had not yet checked all the properties. Because she immediately went on to check closure, however, it seems likely that she had in mind a preliminary rather than final conclusion. In verifying the closure and associative properties, Wendy explicitly referred to the table to support her reasoning:

51   Wendy: And then it's closed. You can see that there are no elements other than 0 through 5, looking at the chart, because we have all possible combinations on elements in Z_6. So it is closed also.

52   Wendy: And associative. You can see, because the chart has symmetry, that the group will be, is associative. This is how I look at it, anyway, because if you look at 2×5 you are going to get 1 and if you look at $2 + 5$ you get 1. But also you know it is Z_6, is also because it's a global property, because addition is associative, for integers, and you know that this carries over to subgroups and so Z_6 will be associative under addition. Do you want me to explain that further?

Wendy made several errors in her attempt to verify associativity. First she stated that she was comparing $2 + 5$ and 2×5 when, based on her statement about the symmetry in the table, she probably was comparing $2 + 5$ and $5 + 2$. A more significant error was that she

was describing commutativity but calling it associativity. Furthermore, the penultimate sentence implies that she thought that Z_6 is a subgroup of Z. During the interview, I pursued the first two errors.

53   Brad: I want you to explain how you said.... What was it you said, 2 + 5 is the same ...?

54   Wendy: 2 + 5 is the same as 5 + 2.

55   Brad: Oh, okay. So that means it is associative?

56   Wendy: Well that is an example of associative.... No, that's not. That's the commutative property. So we have to check 1 + 2 + 3 is going to equal 1 + 2 + 3. That's the associative property. So in a sense we have to.... But we'd have to go through all of the different combinations including 0 through 5 and all of the different elements, which takes a while. But because we know that the associative property holds under integers, for addition, we know it holds. And that's one of the good things that, good facts about that global property because associativity is so hard, difficult to check. Would you like me to try just to see if this checks?

Thus, Wendy was able to correct both errors on her own. Because it took her a moment to realize that her description was about commutativity, it appears that the commutative and associative properties were closely related, perhaps even overlapping, in Wendy's thinking.

Wendy used the idea that associativity is a global property to complete her verification, but again the idea was not appropriate because addition in the integers and addition in Z_6 are different operations. I did not pursue this issue explicitly in the interview but instead asked Wendy only to verify the property for the example she gave.

At this point, I put aside the case of Wendy to extend the analysis to other students, discussing, in particular, the concept of binary operation, the relationship between associativity and commutativity, and the notion of global properties. In this section, I further develop some of the themes that have emerged thus far, including language use and the use of the operation table.

Binary Operation

<u>Operation confusion</u>. One of the most persistent occurrences throughout the interviews was a phenomenon I initially called *operation confusion*, where students were unsure of the appropriate operation on a set. As might have been expected, operation confusion was more likely to occur when more than one operation was available, such as in Z_n, where there are two natural operations. All the key participants experienced operation confusion during the first interview, and most had at least momentary confusion in the third interview when dealing with a function from U_8 to Z_4. All these groups, it should be pointed out, have elements that look like integers but that do not behave quite like the integers with which the students were familiar.

In the first interviews, none of the key participants was immediately sure about the operations that would be appropriate for answering the question "Is Z_3 a subgroup of Z_6?" Carla, for example, stated at first that the operation must be multiplication "because the addition wasn't a group mod n.... Something about multiples of n" (line 12). She then verified that the group axioms are satisfied under addition modulo n, showed that the group axioms are not satisfied for multiplication modulo n, and realized that she had remembered incorrectly.

Robert, on the other hand, was at first convinced that Z_6 is not a group under multiplication, "because the inverses aren't in Z_6" (line 9), a statement that was essentially correct and might have led him quickly to consider addition modulo 6. Moments later, however, he stated that the inverse of 1 would be "1 over 1, just 1" (line 15), demonstrating that he was thinking of inverses as fractions. Then he used analogous reasoning for addition in Z_6.

20 Robert: So, but that's not a group either because in Z_6 if each inverse is itself, the negative of itself, which isn't in Z_6.

21 Brad: What do you mean?

22 Robert: Like $1 + -1$ will equal 0, so -1 is 1's inverse, but -1 isn't in Z_6.

At this point in the interview, Robert was unsure whether Z_6 is a group at all. By making operation tables for both operations, he was able to resolve these issues, although he first stated that 3 is the inverse of 2 under multiplication because the product is 0, demonstrating some difficulty keeping his additive and multiplicative thinking separate.

It is likely that the students' operation confusion was caused in part by the fact that the class had spent time at the beginning of the course solving both multiplicative and additive equations in Z_n. It may also be, however, that the notational convention of writing the group operation multiplicatively when is not specified promotes multiplicative thinking in additive situations such as this.

During the first interviews, resolving operation confusion consumed considerable time for all of the key participants, but as the semester progressed, the students developed more efficient and accurate methods of determining and keeping track of the operation. For example, they used either the identity or closure properties to deduce that the operation in U_8 is multiplication and not addition:

34 Carla: Let's see U_8 is a group under... [pause] ... I am trying to think if it is a group under addition or multiplication. But it must be multiplication, because if it was addition then 0 would be in there. (Interview 3)

17 Lori: Okay. So, [inaudible]. Is it multiplication? Oh, I was thinking it was addition. 'Cause I'm like $1 + 1$ is 2.

18 Brad: And, why wouldn't that work?

19 Lori: Because 2's not in U_8. I don't know why I was thinking that. (Interview 3)

One might hope that determining a group's operation would become a matter of recall. But the evidence suggests some subtlety in the learning process: The students developed increasingly efficient strategies for determining a group's operation. This hypothesis is analogous to the development of proficiency in other areas of mathematics, most notably in the learning of the basic number combinations: Rather than moving from slow object-based procedures to recall, young children proceed along a trajectory of increasingly efficient procedures until the combinations are based either on recall or on procedures that are indistinguishable from recall (see, e.g., Kilpatrick et al., 2001, chapter 6). Thus, the phenomenon of operation confusion may be viewed as a natural stage in the development of proficiency with group theory and its standard examples.

Diamond and star. Another explanation for operation confusion may be an inevitable consequence of one of the goals for the course: an abstract concept of binary operation. So that all binary operations, including familiar additions and multiplications, might be seen as instances of a single idea, Dr. Benson and I chose sometimes to use a neutral notation for the operation. Clearly the notations +, ×, or · would not provide such neutrality. Thus, we often used ◊ (diamond) or * (star) to denote an unspecified operation. One could argue that * does not provide the intended neutrality because the symbol is often used in computer programming languages to denote multiplication. This is certainly a concern, although it is no more problematic than the common practice in abstract algebra texts of leaving the operation implied, as in ab, a convention that clearly carries overtones of multiplication. Even the more neutral ◊ (diamond), however, was problematic, as is illustrated by Diane and Lori as they tried to determine the operation in Z_3.

20   Diane: Z_3 you would have addition, multiplication, or a diamond. And 0, 1, 2, 3 ... [inaudible]. No, it only goes up to 2.

21   Lori: All right, yeah let's make a table of Z_3 and let's make a table of Z_6.

22   Diane: You need to make 3 tables because we don't know what operation we are talking about.

So for Diane, at least, diamond was not a generic operation that could stand for either multiplication modulo 3 or addition modulo 3; it was another operation entirely. Upon questioning, Diane reiterated her list of three operations, so I asked her to write out operation tables for all three. Lori interrupted:

28   Lori: Like it's integers. What's diamond? Can't you only really have these two? That's what I am thinking.

Diane had begun to construct operation tables but quickly reconsidered:

31   Diane: But I don't know how to do a diamond because I don't know what the operation is.

32   Lori: I don't think that you can do diamond, because we are in Z_3 and it's integers, and what is diamond?

33   Diane: Yeah, that's what I am saying. I don't know what diamond is.

34   Lori: So you can only do like addition and multiplication.

35   Brad: Where does the diamond ...?

36   Diane: Diamond comes in when you don't know what the operation is.

37   Brad: Oh, so you mean when you don't know what you call the operation you just use diamond instead? Do you agree with that?

38   Lori: Yes, definitely. But we kind of know that it's integers. So we know how to add integers. It's not like it's a and b, you know. Then I would probably use diamond because I don't know how to add a and b elements.

Thus, despite their momentary disagreement, Lori and Diane both saw diamond not so much as a label for an abstraction under which a number of familiar operations could sit but rather as a device to use when the operation was unfamiliar or unknown.

Furthermore, it seems that Lori had similar thinking about the notational uses of a and b, in the sense that the letters were not generic labels for group elements but rather

unfamiliar objects that she did not know how to add. Thus, even the most neutral notation did not necessarily lead students to the desired abstraction.

The abstract concept of binary operation continued to be problematic for Lori. During her third interview, while determining whether a particular function f from U_8 to Z_4 was a homomorphism, Lori described an appropriate verification formula, $f(a*b) = f(a)*f(b)$, and seemed to know that the operation * on the left was to take place in U_8, whereas the operation * on the right was to take place in Z_4. Nonetheless, she spent a good deal of time determining what the operations were in each of the groups. She first called both operations addition (line 15), yet after deciding that the operation in U_8 was multiplication, she thought that both operations were multiplication (line 30). Finally, because she could not find multiplicative inverses in Z_4, she decided, "This [U_8] is multiplication, and this [Z_4] is addition" (line 34). I asked her whether it was okay that the operations were different.

38   Lori: If we prove it's a homomorphism, yes. [Okay, but....] Right now I am not sure. [Okay.] So I don't, I guess star right now is just going to have to remain generic until I, if I prove it is a homomorphism, then.... It's neat that they call it star because it could be representing two totally different things. [Oh, Okay.] Do you know what I'm saying? [Okay] Like in U_8 it's multiplication, in Z_4 it's addition. [Okay.] So maybe I should just keep it star.

Thus, Lori continued to prefer to use * when there was some uncertainty about the operation, yet she was becoming comfortable with the idea that * could stand for various known operations. It is not clear why or to what extent these impressions were dependent on whether f was indeed a homomorphism. Would Lori have made better sense of the task if the two operations had been notated differently? In class, the verification formula was typically written as $f(a*b) = f(a) *' f(b)$, thereby making it more apparent that the operations might be different. Lori and Diane were unusual in denoting both operations

as *. Because Lori had taken abstract algebra previously, a reasonable hypothesis about her use of notation is that she stuck to notation she had first learned rather than adopting the class's notation, and, furthermore, her usage had rubbed off on Diane.

Regarding Lori's concept of binary operation, it seems unwise to speculate about the source of her confusion and delayed abstraction. Nonetheless, because most textbooks leave both operations implicit, as in $f(ab) = f(a)f(b)$, it is worth considering the relationship between the notation used in introducing the concept of homomorphism and students' concepts of binary operation. If this was a significant moment in Lori's construction of an abstract concept of binary operation, for example, to what extent did ambiguity of the notation * support or constrain this construction?

Late in the course, some students had developed a reasonably robust concept of binary operation, as evidenced by the fact that they were able to switch effortlessly between additive and multiplicative notation and language. Wendy, for example, compared the expression $a*a$ between U_8 and Z_4 by noting, "Here we're squaring it, but here we're saying $2a$" (Interview 3, line 156). This ability led sometimes to problematic or awkward syntax. Robert, for example, called 6 a power of 3 because "If you operate 3 with itself you get 6" (Interview 3, line 272).

Sometimes, however, the switch between multiplicative and additive notation and language was not so effortless, and multiplicative language seemed to dominate. In one interview, for example, I asked Carla to find the subgroup generated by 3 in Z_{12}.

15   Carla: There would be 3, and 9 would be in it because 3 squared is 9. And 0 would be in it because.... Well, actually, Z_{12} is a group under addition. So it's not.... I can't really think of it as 3 squared.... So 9 is in it, but not because it's 3 squared. 9 is in it because it's 3 cubed when you are adding. So, in other words three 3s. (Interview 4)

It took Carla a moment to establish additive thinking, and still she maintained some multiplicative language, leading to awkward phrases such as "3 cubed when you are adding." Later in the same interview, when explaining the sense in which {0, 3, 6, 9} can be an identity element, she indicated some discomfort with the broad use of multiplicative language:

65 Carla: Okay if you add ... We have often called it multiplying, but I don't like that term because to me it doesn't ... I just don't like the idea of multiplying; it doesn't make sense. So I prefer to think of combining them.

In summary, the notion of an abstract binary operation presented notational, conceptual, and even linguistic issues. Coming to view various operations as instances of the same idea was a slow process. Standard notations such as + or · have associated language and thus associated meaning. New notations, such as diamond, were sometimes seen not to represent new abstract categories but rather new operations.

Regarding the concept of binary operation, the students demonstrated on the one hand that they didn't sufficiently distinguish between various operations called addition. On the other hand, they demonstrated that they imposed nonstandard distinctions between notations for generic operations such as * and notations for familiar operations such as · or +. The students also had trouble maintaining the standard distinctions between associativity and commutativity.

Associativity and Commutativity

It should not be surprising that Wendy sometimes confused the concepts of commutativity and associativity, for the concepts are indeed closely related. And in fact, other students also demonstrated similar confusion. Conceptual analysis, supported by a closer look at the data, provides several possible explanations for the close relationship

and the confusion between the concepts. This section builds an explanation out of description of the definitions, distinguishing examples, and verification processes for the concepts of associativity and commutativity.

<u>Definitions.</u> First, the definitions are quite similar in form. On their final exams, all the key participants except Diane gave largely correct definitions of associative operation and commutative operation, such as the following, provided by Carla:

> assoc. operation - the operation where, for any a, b, c, $a*(b*c) = (a*b)*c$.
> comm. operation - the operation where, for any a and b, $a*b = b*a$.

Not all students were careful about the quantifiers. Wendy, for example, stated on her final exam that "an operation is commutative if for 2 elements a and b, $a*b = b*a$." Students' definitions and statements often noted that commutativity is about two elements and that associativity is about three elements, suggesting that this was a salient distinction between the two concepts. In fact, this is the most obvious difference in the definitions.

<u>Few distinguishing examples.</u> Second, although associativity and commutativity are often discussed in high school mathematics, most elementary examples of noncommutative operations, such as subtraction and division of real numbers, are also nonassociative. Experiences in high school mathematics might lead to concepts of associativity and commutativity that are merged into an "order doesn't matter" property. Wendy said almost exactly this in her fourth interview: "Because it's associative, you can move it all around" (line 213). Furthermore, Diane's final exam included similar claims about commutativity: "This property allows us to switch around the elements in an expression so that it doesn't matter which elements will operate first." Unless students'

linear algebra courses emphasized the fact that matrix multiplication is associative but not commutative, this abstract algebra class may have provided students their first opportunity to separate their conceptions of the two properties. Separating the concepts might require at least a few distinguishing examples, but the group axioms suggest an asymmetry: Associativity is more important. In fact, it seems to be difficult to create an operation that is commutative but not associative, particularly via an operation table (Zaslavsky & Peled, 1996; Benson, in press).

Verification. There are other important differences between the concepts of associativity and commutativity, particularly regarding their verification processes. Commutativity is often easy to infer from a description of an operation, and when the operation is given via a table, commutativity reveals itself as symmetry about the main diagonal. Students often used commutativity to help them reason about groups and subgroups, particularly when filling in an operation table. Thus, commutativity is tied closely to the phenomenon of reasoning from the table.

Associativity, on the other hand, is hard to see in an operation table. When an operation is given via a table, the number of calculations required to verify the property is prohibitively high even for groups with as few as 4 elements. When an operation is given via a description or a formula, there are a number of possible approaches, each with its own subtleties. In class, we took essentially three approaches to the problem of associativity. For operations given via operation tables, we often used *Exploring Small Groups* (Geissinger, 1989) to let the computer perform the tedious calculations. At other times, we verified associativity via symbolic proof. Perhaps the most common approach, however, was to argue that associativity was inherited from a larger structure in which

the desired structure lived. In class, this approach was used uncritically and incorrectly by many students and took on a life of its own under the label "Associativity is global." This phenomenon is discussed in more detail below.

These differences between commutativity and associativity provide a third reason for the confusion and also explain the fact that the confusion was essentially one directional: Students sometimes said *associative* and meant *commutative,* but I found no evidence of the opposite. I do not contend that the opposite confusion never occurs but suggest instead that commutativity is more likely to be present in a student's mind. First, it is easier to think about two elements at a time than it is to think about three. Second, commutativity is such a useful property and such a prominent visual feature of an operation table, students are likely to focus on it rather than associativity, despite the fact that commutativity is not one of the group axioms.

Global Properties

Verifying that a set and an operation satisfy the associativity axiom requires particular attention to the operation. As mentioned above, sometimes the associativity of an operation on a set is inherited from a larger structure in which the set and operation are situated. All key participants applied this "global property" idea inappropriately at some point during the interviews, typically by paying insufficient attention to the operation. Furthermore, many of the key participants uncritically generalized the idea to other group axioms.

Both Lori and Robert, for example, claimed that associativity in Z_6 was inherited from Z:

83   Lori: And it's associative because addition is associative and that's inherited from the larger group Z under addition. So that's why it's a subgroup.

69   Robert: We are talking about Z_6. These are integers, and integers fall in the associative law, so it's associative.

Carla demonstrated similar thinking but with more generality and with idiosyncratic language, calling the set Z_n "mod n." (See chapter 7 for detailed discussion of Carla's use of the phrase "mod n.")

79   Carla: So the next thing to check would be associativity. But mod n is a subset of Z because all of your elements in mod n are integers and Z under addition is associative, so therefore mod n under addition is associative. So therefore mod n under addition is a group.

Earlier in the same interview, Carla had similarly claimed that Z_3 inherits associativity from Z_6:

7   Carla: ... All right, so then Z_6 would be 0, 1, 2, 3, 4, and 5. Okay. So we can see that Z_3 is a subset of Z_6 because 0, 1, 2 are elements within 0, 1, 2, 3, 4, and 5. So because of that we know that the associative property holds because the associative property is global. And if the associative property works on this larger set then we know it is going to work on the smaller set because it is, just has fewer elements to work on.

Carla's description of the idea is essentially correct. In this statement, however, she made no mention of the operation and, in fact, had not yet mentioned operations at all in the interview. This suggests that the notion of global or inherited properties may have been mostly about subsets, with little connection to the operation.

Lori provides additional support for this hypothesis. On her midterm exam she stated, while showing that a subset of a group was a subgroup, "We need not show associativity since it is inherited from the larger group." Similarly, on her final exam, she asserted, "Associativity is a global property, so it is inherited from the group." Thus, Lori was able to use the terms *global* and *inherited* with proper syntax. Elsewhere on her final exam, however, Lori incorrectly claimed that Z_4 is a subgroup of Z and also a subgroup of Z_8. Such statements do not make sense, of course, if one is paying attention to the

operation. This suggests that Lori's other statements about associativity were not properly supported by consideration of the operation, despite their correctness.

Like Wendy, other students broadened the idea of global or inherited properties beyond associativity. The notion that the identity is global was perhaps implicit in many students' claims that 0 is the identity for addition, in the sense that the statement holds for a wide variety of representations of groups, with many distinct operations called addition. Of course, the same can be said of 1 as the identity for multiplication.

Diane and Lori argued explicitly that the identity in Z_3 is inherited from Z_6 (lines 78-80). As in the case with Wendy, it is possible that Diane and Lori intended merely that they did not need to show that 0 behaved as an identity in the subset. This simple explanation seems particularly unlikely, however, in light of Diane's subsequent claim that an element's inverse need not be the same in a subgroup as it is in the group:

94   Diane: The only thing that it says about inherited inverses is that you get the inverse of 1 is 2 here and 2 is an element of this Z_6. It doesn't say that the inverse of 1 has to be 2 in here; it just says that 2 is in this, it doesn't say that it has to be the same.

Diane's concept of inverse seems especially problematic here in the sense that the inverse of an element is unique and thus will not change when restricting to a subgroup. On the other hand, on the assumption that Diane had a broad notion of inherited properties—a notion that did not pay much attention to the operation—then it follows that she would say something about inverses in Z_n being inherited from Z. Then, because the inverse of 5 is −5 in Z but 1 in Z_6, her statement would make sense. This hypothesis is made more plausible on the basis of additional evidence of Diane's broad use of the idea of inherited properties. Particular compelling evidence is provided by an earlier claim of some kind of inheritance by Z_3 from Z_n:

4 4   Diane: Well Z_3 isn't a subgroup of Z_6, it's at least a subgroup of Z_n and we know that Z_n is a group under addition, so it would have inherited property.

Unfortunately, there was insufficient data to make much sense out of what Diane meant by Z_n here. Nonetheless, it is clear that her notion of inherited properties was broader than associativity and was insufficiently tied to the operation.

One potential explanation for students' improper generalization of the idea of global properties is that that term itself is nonstandard and lacks a formal definition. The data suggest, however, that the more standard term *inherited* was also problematic. Furthermore, formalizing either of these terms would have been essentially the same exercise.

Subgroups and Binary Operations

At this point, the discussion returns to the case of Wendy to present detailed analysis of Wendy's concept image of subgroup. Again, the main themes are Wendy's use of the operation table and her use of language. Following the detailed presentation, I broaden the analysis to include other students, discussing first the concepts of identity and inverse and the relationships between them, as these concepts became prominent in students' reasoning about subgroups. Then, following a brief discussion of the students' understanding of the concept of closure, the section closes with a presentation of the findings about their concept of subgroup, focusing particularly on the ways that students answered the main interview question, Is Z_3 a subgroup of Z_6? The central issues are the ways that the students distinguished among various operations called addition and the ways that they used the operation table to support their reasoning.

Wendy and Subgroups

When Wendy returned to the question of whether Z_3 is a subgroup of Z_6, she used the addition table to support her reasoning.

76 Wendy: Now is Z_3 a subgroup of Z_6? Now, we have to check that Z_3 is going to be a group because it has to have all of the elements [axioms] of a group, which means it has to have identity and inverse; it has to be closed. So I am going to start checking Z_3. Z_3 would consist of 0, 1, and 2 under addition. But Z_3, the table is going to be different. See I am going to have to explore right now whether or not.... When you say Z_3 is a subgroup of Z_6, whether it means you are taking Z_3 out of Z_6, or if you are just looking at Z_6 [Z_3] and seeing whether it's a group. See when you say something is a subgroup of something else [pause] I am not quite sure what way to look at it. Like how it exactly, like how Z_3 ties into Z_6, like to be a subgroup of Z_6. What, that.... Like I know how to check whether or not Z_3 itself is a group and whether Z_6 is a group, but to check whether Z_6, Z_3 is a subgroup of Z_6, I don't know exactly what to look at.

Wendy had a sense that the operation table for Z_3 would be different, depending upon whether it was constructed on its own or taken out of the Z_6 table. Consistent with the emerging hypothesis that Wendy's reasoning was highly dependent on *looking* at an operation table, it seems that her statement "I don't know exactly what to look at" meant she didn't know what *table* to look at.

This excerpt suggests a much stronger observation than has been drawn thus far. Rather than saying the operation in Z_3 is different, Wendy said, "But Z_3, the table is going to be different" (line 76), suggesting that the table was not merely supporting her reasoning but rather was substituting for the group in her thinking. The phrase "taking Z_3 out of Z_6" (line 76) suggests again that, for Wendy, Z_6 was not merely a list of elements that appeared on the edges of the table but was in fact the table. This conjecture is further supported in the following explanation in which Wendy referred not to the group Z_6 but again to the table:

78 Wendy: Because if you use the elements of Z_3, which is 0, 1, and 2—are the elements of Z_3. But if you look at them in terms of Z_6, like if you just look at this section of the table

Z_6 (see Figure 7), this isn't going to be a group.

79   Brad: Why?

80   Wendy: Because it is not closed.

81   Brad: Why?

82   Wendy: Because 4 isn't an element of Z_3.

83   Brad: And where did the 4 come from?

84   Wendy: 2 + 2 from Z_6 because it's mod 6 in Z_6, but when you look at Z_3 it is mod 3.

Figure 7. Wendy's table for addition in Z_6, second version

+	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

Thus, through her reliance on the table, Wendy had correctly identified the central issue behind the interview question: whether the addition was to take place based on the operation in Z_3 or in Z_6. Nonetheless, she was not ready to come to a conclusion:

86   Wendy: See, it doesn't make sense. Like, I started over here to do, to look at whether or not Z_3 was a group itself, but that didn't make sense to me.

87   Brad: What didn't make sense?

88   Wendy: To look independently to see whether Z_3 was a group under addition. Actually, I think for the same reasons it is going to be a group under addition, just like Z_6. I think any Z group under addition is going to be a group because 0 is going to be.... Well, I guess it depends what elements are in there, but.... Like Z_3 is going to be a group, it's easy to see after looking at Z_6. But if you just look at it separately, it doesn't really make sense whether, like, to tell whether or not Z_3 is a subgroup of Z_6 to just look at whether Z_3 is group because it has no connection with Z_6.

It is surprising that Wendy was not able to make general statements about "any Z group" but instead stated "it depends what elements are in there." Perhaps this is merely evidence that she needed to see the operation table in front of her. Nonetheless, she was concerned that there should be a clear connection between a subgroup and the group that

it was supposed to come from. Earlier statements indicate that she thought the connection should come via the operation table.

To provide some clarity, I first asked Wendy to compare the two different versions of $2 + 2$ that she had discussed. She explained:

92   Wendy: This is going to equal 1 in Z_3 mod 3 because that equals, $2 + 2 = 4$. In mod 3 that is going to equal, the remainder's 1. But here it's only, it's still going to be 4 because it's mod 6.

Next, I checked briefly why Wendy had not pursued multiplication as the operation in Z_3. (Misused words are set in bold to help call attention to them in the following discussion.)

95   Wendy: Well Z_3 isn't going to be a **subgroup**, isn't going to be a group under multiplication because if you look at the first row it's going to equal the same thing as it was up here. Like, they have very similar relationships, the Z tables. Like, Z_3 under multiplication has a similar relationship to Z_6 multiplication table, as does Z_3 under addition and Z_6 under addition. Like they are going to have the same identity under multiplication and **division** [addition]. So if you look at Z_3 under multiplication I'd know that the first row is going to be—I'm going to fix this—is going to be 0's and from here you know that 0 does not have a, doesn't have an **identity** element, or an inverse, excuse me. So you know already.

99   Wendy: There are no elements in Z_3 when multiplied by 0 will give you the identity 1. That's why you know that, again, for the same reason, Z_3 is not going to be a group under multiplication.

Wendy had trouble saying what she meant here, correcting her language twice (*group* for *subgroup* and *inverse* for *identity*) and also meaning *addition* but saying *division*.

Nonetheless, it seems that she was reasonably confident about the fact that Z_3 is not a group under multiplication (mod 3). But to get some clarity on the extent to which Wendy associated an operation with Z_n, I asked her about Z_{10}.

104   Wendy: Like, I automatically know when you say Z_{10} that, under addition now it's not going to have an inverse element.

105   Brad: Under addition?

106   Wendy: I mean under multiplication it's not going to have an inverse element. Under addition it probably will be a group; it will be a group.

Thus, Wendy still had trouble using the words she meant, saying *addition* when she meant *multiplication*. Furthermore, her syntax "not going to have an inverse element" is more appropriate for talking about the identity.

Then we returned to whether Z_3 is a subgroup of Z_6.

111   Wendy: Like a subset of.... I think we have to look at it as like part of the set of Z_6, which, like subgroup, like as a group in Z_6. So if you look at ... which is why I kind of choose the elements Z_3 out of the Z_6 table.

Wendy was clearly thinking of Z_6 as more than a set and as Z_3 as more than a subset. She was choosing not "elements Z_3 out of the Z_6 table," but rather entries out of the Z_6 table that corresponded to the restriction of the binary operation to the subset Z_3. On the conviction that this was the appropriate method, Wendy decided that Z_3 is not a subgroup of Z_6 because the subset was not closed under the operation.

The fact that she had answered my question was apparently of little concern, however, for she immediately began focusing on the manner in which closure had failed. In particular, she looked at the 3 and 4 that appeared in top left quarter of the Z_6 table (see Figure 7).

123   Wendy: If you just looked at, are, is the subgroup 3.... Well, I don't know really what you'd call that.... But if you kind of just look at the, like, elements 3 and 4.... Actually 4 doesn't have, forget it. 4 doesn't have an inverse.

125   Wendy: I was trying to look. This is closed. Like I was going to say, if you only look at the elements 3 and 4 in Z_6, I was going to say under Z_3 it was going to be closed, but I was just kind of thinking. But I'm not, that doesn't make sense at all.

Wendy may have been considering {3, 4} to determine whether it was a subgroup but saw that 4 doesn't have an inverse in {3, 4}. She may also have seen that 3 is its own inverse. She continued looking for a subgroup.

127   Wendy: Well, I was just kind of looking. Cause 4 isn't an element of Z_6, element of Z_3. So therefore it wouldn't be closed. But the problem are these elements right here: 3, 3 and 4, and 4. So like if you looked at Z_2, it is going to consist of these first two elements. The problem here again is going to be this last multiplication, or addition of 1 and 1,

cause that's going to equal 0.

It seems Wendy abandoned looking at {3, 4} and instead was trying to exclude the

problematic entries in the table by considering Z_2, both as a subset of Z_6 and as a group

on its own. I asked her to explain what she was doing:

129   Wendy: I am just relating to what subgroups would be. What subgroups, what Z mod n subgroups would be a subgroup of Z_6.

132   Brad: And you were trying specifically ...

133   Wendy: Z_2. But the problem is, if you look at Z_2, this 2.... Like 1×1 [1 + 1] in Z_2 equals 0 and that causes ... or that gives you the identity [inverse] element in 1, for 1. But here if you just look at it under Z_6, it doesn't, 1 doesn't have an identity [inverse] element. Just like here [the {0, 1, 2} subset of Z_6], 1 and 2 don't have an identity [inverse] element also, besides it not being closed, there are a lot of reasons why it's not going to be a subgroup.

Once again, Wendy was saying *identity* and meaning *inverse*, and she confirmed this

moments later. But this excerpt provides something of an explanation for her confused

language: She was using the operation table for Z_6 to support this reasoning. In

particular, she was checking the inverse property for various subsets by looking for the

identity inside the appropriate subset of the operation table. Because her process

involved looking for the identity, it is not surprising that Wendy said *identity* rather than

inverse.

This process was in service of a larger question that Wendy was pursuing. She had

generalized the question "Is Z_3 a subgroup of Z_6?" to consider whether Z_n might be a

subgroup for other n. This provided a natural transition to ask Wendy whether she could

find any subgroups of Z_6.

144   Wendy: See Z_6, it's hard to take a subset because you have to make sure you include the identity element in the set that you pick. So let's, just for instance, I'm going to take this. Because if I am looking in the fact that you have to have an identity element. Here, if you look at 1, 2, and 3 they each have and 3, 4.... You can't do that. 'Cause now it's not closed, really. You can't take 3, 4, 5 and 1, 2, 3. It wouldn't work.

Wendy saw that she needed to include the identity, but she was simultaneously considering "blocks" in the operation table, and she saw that this would not work. I suggested that she think more broadly and consider subsets with nonadjacent elements, such as $\{1, 4\}$.

146   Wendy: 1 and 4. Like, what I was saying before, 3 and 4? Like looking just at 3 and 4?

147   Brad: Yeah, or maybe not two that necessarily that are right next to each other. Like what about 3 and 5? Could that work? Or.... Do you know what I am saying?

148   Wendy: It's just easier for me to see [inaudible].

149   Brad: So, what are you doing there? Oh, you're covering up 4.

150   Wendy: 4. It distracts me.

At first, Wendy persisted looking for blocks in the table (line 146), and so I suggested once again that she consider subsets more broadly. This excerpt suggests that she was looking at blocks in the table partly for visual reasons. Because it was hard for her to *see* the operation table for subsets that were not blocks, it was therefore hard to think about subsets that weren't blocks. A related possibility is that she looked for blocks in the table because of an overly limiting interpretation of the Groups-Are-Containers metaphor. If groups are containers, then subgroups must also be containers, but it is difficult to imagine a container that holds every other element from the group table, for example. Wendy tried to overcome this limiting view as the interview continued.

Figure 8. Wendy's table for Z_6, annotated version

+	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	⓪	1	②
4	4	5	0	1	2	3
5	5	0	1	②	3	④

Note: circles added to clarify transcript

150   Wendy: ... Technically you are only looking at the 0, 2 ... 2, 4 [circled in Figure 8].

Right? Because, in other words, you can make that table.... You can't look at the other elements. You can't look at the whole row, 3 and 5. You know what I mean? Because you can only look at the addition of those two. You can't start including 0, 1, 2 added to three, because you have to restrict it to three if you are to restrict it to 3 and 5, in order for it to be closed. So if I am going to look at the addition of just 3 and 5 [pause] 0 [5] doesn't have an id-, a inverse element.

151   Brad: What do you mean?

152   Wendy: Nothing [incomplete thought].... When you add 3 or 5 to 5, you can't get 0. Like, 3, when you add 3 to itself you get 0. So that wouldn't be.

153   Brad: So 3 has a ...

154   Wendy: It is kind of hard. Like, what if you took 1 and 3. Oh, no. Not 1 and 3. You have to make sure you pick 1 and 5. What if I tried three, picking three numbers?

Wendy described how she was restricting her view of the table, describing precisely those entries inside the table (0, 2; 2, 4) that were relevant to whether {3, 5} is a subgroup. Furthermore, she justified this view by noting that "you can make that a table" (line 150). From this view, she noticed that 5 does not have an inverse in {3, 5}, although she said at first that 0 does not have an inverse, perhaps because she had been looking for a 0 in that row. Then, perhaps prompted by the fact that 5 lacks an inverse in {3, 5}, Wendy decided to begin with the set {1, 5} to see whether it was a subgroup of Z_6.

As the interview continued, she focused on the inverse property.

156   Wendy: Let me pick 1 and 5. 1 and 5, and that would give you.... I'll tell you how I am going to do this. 1 and 1 is going to give you 2. 5 and 5 is going to give you 4. And 1 and 5, and 5 and 1, is going to give you 0.

157   Brad: Okay.

158   Wendy: You see that 1 and 5 both have an inverse. So, ooh.

159   Brad: Ooh what?

160   Wendy: 1 and 5 work, so far. It hasn't.... They both have a inverse element. You see what I mean?

161   Brad: Uh huh.

162   Wendy: 1 and 5 are their own inverse, are each other inverse elements. So if you took those two separately, it upholds the inverse property. Identity ...

Although the previous excerpt suggested that Wendy had chosen {1, 5} so that it would satisfy the inverse property, it seems in this excerpt that she was unsure whether this choice would work until she had considered the operation table. Her syntax about the relationship between 1 and 5 was somewhat confused and inconsistent, however, evolving from "have an inverse" to "are each other inverse elements."

At the same time, it is apparent that Wendy was thinking *only* about the inverse property. If she had been thinking about closure, she would have noticed during her calculations that closure was not satisfied. Furthermore, she was ready to consider the identity property only after she had completed her verification of the inverse property.

164   Wendy: It has.... However, it doesn't have an identity element. Like, you have to get 1 and you have to get 5. Like you have to.... If you take something, you kind of have to build from it. Kind of like what we did in abstract class. They gave us, like, one—this confuses me, but—one element of a subset, of a subgroup and they said, "Is this a subgroup?" It wasn't. Well then you kind of have to see what it's missing, and you have to kind of build the subgroup.

165   Brad: Oh, okay, well try that here then. It's a good idea.

166   Wendy: Okay. So, I need.... Well I picked two numbers so that it upheld the inverse property. But now it doesn't have the identity property, which means when added to itself, or when added to another number it gets itself. And that's 0. It has to have 0 in it. So I am just going to move this over. Move this down.

Drawing on a procedure developed in class, Wendy considered adding elements to the set in order to build the subgroup one element at a time. Here she realized that she needed to include the identity element in order to be sure that the identity property was satisfied.

This seems to be a significant moment regarding the identity, for from this point on, Wendy always included the identity early when constructing a subgroup. But at this point, she was ignoring closure and was having trouble reasoning about the set {0, 1, 5} because the table for Z_6 was cluttered with other elements. Thus, she decided to "move

over" the relevant portions of the table into a smaller table containing only the three

elements she was interested in (see Figure 9).

Figure 9. Wendy's table for {0, 1, 5}

+	0	1	5
0	0	1	5
1	1	2	0
5	5	0	4

170   Wendy: Okay. So, here everything has an ... 1 has, everything has.... There is an identity element in here, because 0 is in the table, and therefore 0 added to anything is going to equal the number itself. So zero is everything's identity.

171   Wendy: So now everything has an inverse, even though you added 0, 0 is its own inverse. So, therefore, you didn't have to worry about changing the inverse, like disturbing the inverse property when you added 0.

172   Brad: Oh, okay.

173   Wendy: It's not closed. [Laughs.] Oh, no. It's got 2 and 4 in it. This is just getting really difficult. Like, you're going to have to keep on.... You're going to have to add 2 and 4 now. So the only thing you are missing is 3, and if you ... I am sure if you add 2 or 4 you're going to get.... So if you just do away with 3.... 2 times 1 is going to equal 3. 2 + 1 is going to equal 3. And therefore you're going to need to add 3 in there. So it doesn't work.

There are two points to make here. First, Wendy had been considering the group axioms

one at a time and did not move flexibly among them. From her laughter and frustration

in noticing that the set was not closed, it is clear that she had not considered the closure

axiom earlier in this example. Second, this excerpt reinforces the hypothesis that Wendy

began constructing the set with 1 and 5 because together they satisfied the inverse

property and then added 0 to the set so that there would be an identity element. When

she returned to check the inverse property (line 171), she still was thinking of the process

by which she had constructed the set, but her reasons for choosing 1 and 5 as a pair were

not explicit. This omission may be significant because she seems to have forgotten her

reasons only a few minutes later when, taking advantage of her idea to "build up"

subgroups, I asked what would happen if she had started with different elements or with only one element:

186   Brad: If you just started with 1, would you need 5?

187   Wendy: I don't know. You have to see after you add 2. If you add 2, 2 and 1 is going to equal 3. Then you're going to need 3. 3 and 1 is going to equal 4, and you're going to need 4. And then 4 and 1 is going to equal 5 and you're going to need 5. [inaudible]

188   Brad: Okay. So in other words, if you start with 1, what else do you need?

189   Wendy: 0 for the identity element.

191   Wendy: You need 2 for closure, 3 for closure, [laughs] 4 for closure, and 5 for closure.

Wendy was no longer thinking of the inverse property, or she would have responded (in line 187) more quickly that 5 was necessary. Instead, she was thinking about closure, which is why 2 needs to be in the subgroup. Continuing the processes of adding 1, she decides that 3, then 4, and eventually 5 must also be in the subgroup.

At this point in the interview, the identity axiom was fairly immediate and salient, but the inverse axiom had faded into the background, obscured by the closure axiom. Closure remained dominant as the interview continued. Despite Wendy's frustration, she had built up some ways of thinking that allowed her to proceed more quickly. I asked her to try starting with a different element.

196   Wendy: If you start with 2 you are going to need 0. You are always going to need 0, 'cause, like you said. Okay. So, things are getting kind of messy. I need a new piece of paper. If you start with 2, you're going to need 4.

198   Wendy: And when you're doing 4, you need 0. Well.... Ooh.

199   Brad: Ooh what?

200   Wendy: You need 0, anyway. You need 4 though. [inaudible] So, 2 ... 'cause 2 and 4 is going to equal 0. Uh oh.

201   Brad: Uh oh what?

202   Wendy: It works! You don't.... It's closed. It's got an id-, everything has an identity element ... 0 is the identity element for all, each element. Well, they have to have the same identity element, but.... And it's got an inverse.

Figure 10. Wendy's table for {0, 2, 4}

+	0	2	4
0	0	2	4
2	2	4	0
4	4	0	2

Wendy had found a subgroup, and, for the first time in the interview, she considered several axioms in quick succession. Her explanations were somewhat muddled, however. In particular, her syntax for the identity and inverse was mostly reversed, suggesting once again that these two properties were closely related in her thinking. She did clarify that all elements have the same identity. I asked for clarification about inverse:

203   Brad: What's got an inverse?

204   Wendy: Every element has an inverse. So that's a subgroup.

Because Wendy corrected her syntax regarding the inverse property, it seems that she could distinguish the inverse property from the identity property, even if the distinction was not automatic.

Again, it is remarkable that Wendy considered three of the group axioms almost simultaneously, suggesting growing fluency with the axioms. She had said nothing, however, about the associative property.

205   Brad: Did you check all of the properties?

206   Wendy: No, I did not check associative. [Laughs.] No! [Her tone suggests she'd rather not check the associative property.]

207   Brad: And why? Do you think you need to check it?

208   Wendy: No, because it's a global property. And if it's.... Addition is associative. So no [matter] ... If addition is associative, doesn't, under integers.... Taking any integers, it's still going to be associative. So there's no need to check it.

209   Brad: Okay. So what do you have here?

210   Wendy: A subgroup of Z_6.

Wendy's response to my question indicated some frustration that all her work had not yet produced something that she could be sure was a subgroup. Nonetheless, she overcame this frustration quickly upon remembering that associativity is a global property. Wendy did not correctly use the idea that associativity is global, however, basing her conclusion on the associativity of addition in Z rather than in Z_6. Nonetheless, she was correct that $\{0, 2, 4\}$ is a subgroup.

Next, I asked Wendy whether there are any other subgroups of Z_6.

212   Wendy: I don't know I'd have to play and try. I just found one; I didn't think that we could find one, but I just found one.

213   Brad: Okay. What do you think? Another one?

215   Wendy: Let me try 3, starting with 3. You need 0. You always need ... [Whistles.] Found a group!

216   Brad: You found a group?

217   Wendy: Yeah. Because it's got a identity element. Whoops, I made a little mistake in my calculation, but.... It's got an identity element, 0. It's got an inverse because $0 + 0 = 3$ [0] and $3 + 3 = 0$, so it's got an inverse. It is closed between 0 and 3 and it's associative. So here's another subgroup.

Again her syntax regarding the inverse property is more appropriate for talking about the identity, yet her calculations indicate that she did know that each element must have an inverse. And again she considered the inverse, identity, and closure properties in quick succession.

I asked whether there were any other subgroups.

220   Wendy: No, because 4 you would need 2.

221   Brad: Why?

222   Wendy: Because if you have 4, 4 and 4 is 2. And therefore you need 2.

225   Brad: Well, what if you had 5? What if you started with 5?

226   Wendy: 5 is the same thing as 1.

227   Brad: Why?

228   Wendy: Because 5 and 5 you are going to need 4, and then 4 and 5 are going to need 3.

Okay? And then 3 and 4 you going to need 1, you are going to need 0, and 4 and 4 is going to equal 2. So you need everything.

For both of the examples in this excerpt (starting with 4 and starting with 5), Wendy's arguments were based on the closure property, not the inverse property. The inverse relationship between 5 and 1, which had earlier been quite present in her thinking (see lines 154 and 158), had faded into the background. Instead she focused on the need to satisfy the closure property.

Next, I asked Wendy whether she saw any relationship between Z_3 and the subgroup $\{0, 2, 4\}$ in Z_6. She paused for a moment and then responded:

235   Wendy: You multiply Z_3 by 2, all of the elements by 2, and you get this subgroup. I don't know what you call it. I don't know.

Taking advantage of this relationship, Wendy decided to call the subgroup $2Z_3$. She asserted that Z_3 and $2Z_3$ are not the same but are related by multiplication. Wendy was not satisfied with this description, however, and wanted to find a deeper explanation.

249   Wendy: Yeah, but if you take every element in Z_3 and you multiply it by 2.... I can't really make that connection yet, like, why that exactly works. I know it definitely has the.... Like, I think it definitely affects the fact that 2 and 4 are factors of 6. Not factors.... Oh, no. They're.... Like 3, when added to itself is going to equal 0. When 2 and 4 are added to each other, you're going to ... it's going to keep it closed. Like when you start adding 1 you're switching.... Like, these are two even numbers. The fact that they are two evens, two evens are going to equal an even number. I don't know if it has to do with the evens, but.... I see a definite pattern why these two are going to be subgroups. Because 2 and 4.... 2 and 2 is going to equal 4. 4 and 2 is going to equal 6, and 6 is going to be, 6 is equal to 0. So all these.... Like 2 and 4 when.... I've explained this to you [inaudible] four times. I can't explain.... I don't know. [inaudible]

250   Wendy: Like it makes total sense to me that these two are groups. And I can see why this isn't. So can I re-ask a question or can you re-ask me a question?

Wendy considered factors and evenness to explain why $2Z_3$ would be a subgroup, but neither of these provided a clear explanation. To assist her in searching for the explanation, she sought a new question, suggesting that she saw questioning as a useful means for developing insight and explanation. She continued looking.

254   Wendy: Well, I want to sort of look at $4Z_3$, for no reason at all [inaudible]. If I'm looking at $2Z_3$, why not look at $4Z_3$. And you're going to get.... Actually, we're adding here. This is going to equal 4.

Wendy calculated $4Z_3$ to be the set $\{0, 4, 2\}$ and began constructing the operation table, using that order. Because she was not explicit about how she was doing the arithmetic, I asked her whether she was still doing arithmetic modulo 6:

258   Wendy: Yeah, just as I did here [in the $2Z_3$ table].

260   Wendy: Ooh, wait a minute. But here I said this was a subgroup. This is a subgroup of Z_6, but that's because in Z_3, we're restricting our elements to 0, 1, and 2. Here we are allowing for higher numbers.

262   Wendy: So, technically, this isn't $2Z_3$. Cause it's not in mod 3.

263   Brad: Oh, I see.

264   Wendy: This is definitely not $2Z_3$, $2 \times Z_3$.

265   Brad: But it is the set $\{0, 2, 4\}$, which.... I can see why you want to call it $2Z_3$. I'm not sure.... I mean, maybe that's a good notation [inaudible].

266   Wendy: Well, I guess if you say Z_3, Z_3 has to hold.... Like, if you take all the members of Z_3 and multiply them by 2, who said they still have to hold the stipulations of Z_3? It has to be divisible, like you look at their remainder after dividing by 3. So I guess you could still say $2Z_3$. But I still don't know the connection between $2Z_3$ and Z_3 and why $2Z_3$ is a member, is a group of, a subgroup of Z_6 [inaudible]. I know why Z_3 isn't. But I just don't know why $2 \times Z_3$ would work. This isn't a.... 2×1 equals 2. 2×4 is 8 equals 2.

Wendy was uncomfortable with her notation. She was sometimes adding and at other times multiplying, sometimes modulo 3 and at other times modulo 6. This inconsistency caused confusion that she was not able to resolve. Nonetheless, without prompting from me, Wendy saw a relationship between the tables that she had called $2Z_3$ and $4Z_3$.

272   Wendy: I think this is just a different arrangement of this. Do you see what I mean? This is just a different arrangement of this.

273   Brad: So the thing you're calling $4Z_3$ and the thing you're calling $2Z_3$ ...

274   Wendy: Are the same, just a different arrangement.

The fact that Wendy called two different tables the same suggests that, by the end of the interview, she had begun to separate the table from the group. The question is whether she saw the table (and various rearrangements) as the object of investigation or,

alternatively, saw the two tables as representations of another object. Sfard (2000) points out that "the transition from signifier-as-object-in-itself to signifier-as-a-representation-of-another-object is a quantum leap in a subject's consciousness" (p. 79). Such a distinction between signifier and signified might mark the creation of an abstract object—in this case the group Z_3. This sort of separation between the table and the group paves the way for the concept of isomorphism, which gives rise to the idea of an abstract group that is independent not only of the arrangement of its elements in the table but also of the names of the elements as well. It is not clear, however, to what extent Wendy had made this conceptual leap by the end of the interview. Because isomorphism was the theme of the second interview, this issue is explored in more detail below.

Wendy's reasoning about groups and subgroups was largely external, often requiring that relevant portions of the table be present before her eyes without extraneous information interfering with her perception. When considering whether {3, 5} was a subgroup, she covered up the 4, and when building a subgroup with {1, 5}, she created a new table separate from the Z_6 table. In large measure, the operation table *was* the group for Wendy, although she had begun to separate the group from the table, as evidenced by her suggestion that the table she called $4Z_3$ was a rearrangement of the table she had called $2Z_3$. The operation table both supported and limited Wendy's ability to reason about groups and subgroups. On the one hand, the table helped her see quickly the problem with considering Z_3 to be a subgroup of Z_6. On the other hand, her reliance on the table made it difficult for her to find subgroups.

A symptom of the external, table-based nature of Wendy's reasoning was that she often considered only one group axiom at a time when reasoning about groups and subgroups.

Toward the end of the interview, however, she had developed more fluency and was able to move quickly among the axioms. Moving toward considering the axioms multiply and flexibly might be described as a matter of increasing proficiency and fluency with the group axioms and with the particular examples, which may be a result of internalization of some of the external processes that were based in the table.

The above case demonstrates some of Wendy's difficulties with language and also the ways that her language use shed light on the ways she was thinking about some of the concepts. The case also demonstrates some of the ways that Wendy used the table to support her reasoning. In the sections below, I further illustrate these themes by broadening the analysis to include characterizations of the concept images of other key participants. The theme of language use is particularly prominent in the discussion of the concepts of identity and inverse. The theme of the use of the operation table is central in the discussion of the concepts of closure and subgroup.

Identity and Inverse

Like Wendy, the other key participants demonstrated that their concept images of identity and inverse were closely related. In this section, I first present a synthesis of the definitions and informal meanings that students associated with the concepts, followed by a description of the notational, linguistic, and conceptual expectations that the students seemed to have for each of the concepts. Then I provide some additional examples of confusion between the two concepts and some explanations based in procedures and operation tables.

Definitions and meaning. On the final exam, several students provided definitions of the identity and inverse as part of their definitions of group. Lori, for example, wrote the following:

> There must be the existance [*sic*] of an <u>inverse</u>: $a \diamond a^{-1} = e$ (where e is the identity element).
>
> There must be an existance [*sic*] of an <u>identity</u>: $a \diamond e = a$ (e is the identity element).

The syntactical similarity between the definitions suggests that the two concepts were closely related in Lori's thinking. Furthermore, the quantifiers and other specifications are missing, yet the formulas are correct in the sense that correct definitions could be crafted around these formulas. This characterization fits many of the definitions that students provided.

There were also important differences between students' definitions of identity and inverse. In particular, the definition of identity seems to have been more difficult to formulate than the definition of inverse. Compare, for example, Robert's definitions:

> <u>identity</u> - an element e such that $a \diamond e = a = e \diamond a$.
>
> <u>inverses</u> - for each $a \in G$ there is $a^{-1} \in G$ such that $aa^{-1} = e = a^{-1}a$.

Though Robert's definition of inverse was essentially correct, including the quantifiers, his definition of identity lacked quantifiers entirely. Wendy's definition of identity was also problematic:

> There is an identity element for the group so that every element in G, when multiplied by this identity element, e, will give you back the original element: $\{x \in G \mid xe = x\}$.

Wendy's informal characterization was essentially correct and included the quantifier "every element in G." The formalization at the end, however, is incorrect. A standard mathematical reading of Wendy's symbolism would be, "The set of x in G such that $xe =$

x" or, more informally, "All x in G such that $xe = x$." This is not far from the correct condition, "For all x in G, $xe = x$." Thus, Wendy had specified a set rather than a condition on G, and her unusual symbolism may be interpreted as difficulty with correct symbolic use of quantifiers.

Informal characterizations such as "giving back the original element" were common for the identity element. On her final exam, Lori noted, "The identity of Z_4 is 0 because 0 plus any element in Z_4 gives back that element." Wendy called the identity the "do-nothing element" (Interview 2, line 306). Carla elevated this characterization to a definition: "So you could call 0 the do-nothing element, which is the way we've defined identity" (Interview 1, line 88). Robert combined these characterizations: "R_0, which doesn't do anything to them. R_0 composed with any of them leaves them the same. So there is an identity" (Interview 2, line 178).

Informal characterizations of inverse were more difficult to formulate. Lori, for example, was quite vague: "Because the inverse of something is when you operate two things to equal the identity" (Interview 1, line 87). Recall that Wendy struggled and eventually came to an approximate characterization: "So when you multiply some number m by, it has to have an inverse i, so that when multiplied, it will equal the identity" (line 30). Carla, on the other hand, was more precise in her language, even in her first interview. She stated, "To get the inverse you have to find something that adds with your element that results in the identity, which is 0 in this case" (line 39). In the same interview, she used similar syntax when she described that for something to be an inverse of 2: "It means that 2 times that thing equals the identity" (line 120).

<u>Expectations about the identity</u>. The students' images of the identity property seemed partly tied to what the identity was called. In other words, the symbol that was used for the identity (e.g., 0, 1, or e) seemed to support and facilitate the students' thinking. In fact, the students' use of 0 and 1 (and eventually e) was so flexible that it was not possible for me to distinguish between the symbol and the name. Not surprisingly, there were also strong connections between the name of the operation and the expected name of the identity. When the operation was called *addition*, for example, students expected the identity to be called 0. Then, when determining whether a set was a group or a subgroup, they needed only determine whether 0 was in the set. Similar statements can be made about multiplication and 1. Carla, for example, was explicit about this process:

168   Carla: So then you want to check the identity. We already said that the identity for integers under addition is 0. So we know that the identity for Z_3 is 0. The question is, Is 0 in Z_3? And yes, it is. So therefore we have an identity.

This statement suggests that Carla was not necessarily distinguishing between addition in Z and addition in Z_3 and that 0 being the identity was a global property. Both of these issues are discussed further below. Here I wish to suggest that students also had a sense that the 0 in Z_3 is the same as the 0 in Z. This point brings into question the practice of calling the elements of Z_3 the integers 0, 1, and 2. The alternative is to construct Z_3 as equivalence classes in Z so that the elements of Z_3 are subsets of Z and 0, 1, and 2 are but convenient representative elements. When using representative elements to name equivalence classes, some texts use a bar over the representative element, as in $\bar{2}$, so as to distinguish the equivalence classes from elements themselves (see, e.g., Bhattacharya, Jain, & Nagpaul, 1986). This approach might solve the problem of failing to distinguish between 0 in Z and 0 in Z_3 but might also create a different collection of conceptual and notational issues.

In addition to natural facility with identities called 0 and 1, the students also developed

facility with calling the identity *e*. When groups were given by tables with elements

named by letters, Diane and Lori seemed to prefer that the identity be called *e* and

hesitated when the identity was called something else. Yet they also seemed to know that

the names of the elements do not matter. I asked them whether it mattered that the

identity was called *e*.

422   Diane: That's just convention. I mean we could make it *i* if we want it to be and just say *i* is the identity. That's just conventional.

423   Lori: When we renamed and reordered, we looked at the table after we renamed and reordered. We said, "Okay, what acts like the identity?" And that's what we have to set equal to the identity so that we make sure we get back one of our tables. (Interview 2)

Whatever names were given to elements in a group, most students were able to notice

when elements acted like the identity even when the elements were themselves sets, such

as in a group consisting of two elements, $\{1, 3\}$ and $\{5, 7\}$. In this group, Wendy called

$\{1, 3\}$ the "identity set." Often this sort of reasoning seemed to arise out of the operation

table, but the students also were aware of distinguishing characteristics of the identity.

Wendy, for example, noted on her midterm exam, "The only element that when

multiplied by itself, gets itself is the identity element."

Expectations about inverses. In some representations, particularly when the

representations looked like integers, the students sometimes drew on their experience

with integers and rational numbers and expected inverses to be negative numbers or

fractions, depending upon the operation. Carla, for example, suggested that "the inverse

of 2 mod 3 would be 1/2 mod 3, and 1/2 is not an integer, and it is not in mod 3. And the

only elements of mod 3 are 0, 1, 2" (Interview 1, line 118). Similarly, Wendy stated that

"multiplication is not a group, because there's no inverse ... because they're, under

multiplication they're going to be, like a fraction" (Interview 3, lines 36-38). In his first interview, Robert expected multiplicative inverses to be fractions and also expected additive inverses to be negative: "Like $1 + -1$ will equal 0, so -1 is 1's inverse, but -1 isn't in Z_6" (line 22).

This phenomenon of expecting negatives and fractions for inverses is analogous to expecting that 0 is the identity for any operation called addition and 1 is the identity for any operation called multiplication. In the case of the identity, this phenomenon rarely caused difficulty because 0 and 1 often continue to behave as they do in the integers. In the case of inverses, however, the tendency is potentially more problematic because, for example, -3 and $1/3$ do not have obvious meaning in Z_n. The concept of inverse provides the meaning by which these symbols may be interpreted in Z_n. The students, on the other hand, used their understandings of the rational numbers -3 and $1/3$ as the source of meaning, and those meanings did not fit with their images of Z_n.

Confusing identity and inverse. Wendy's linguistic confusion between inverse and identity continued at least into the second interview. When she was investigating the powers of a specific permutation α, for example, I asked her what α^0 would be. She first called it E and later explained:

308   Wendy: The identity. It's the do-nothing. It doesn't do anything. When you put anything to the power of 0 it doesn't ... Like, any number to the power 0 is going to equal 1. Okay? Because ... And 1 is the multiplicative inverse? You know, like, it doesn't do anything. Multiplicative inverse is the identity. Or not multiplicative inverse. I'm not talking ... I'm not ... I don't know why I just said that. [laughs] But any power, any number to the power of 1 [0] is going to equal 1, which is the multiplicative identity, not inverse. Right? [inaudible] multiplicative identity. So, alpha to the 0 is going, in cycle notation, has to be the cycle identity, which is the do-nothing cycle, which is 1.

Wendy was able to correct her language, though not without a struggle. Other students also sometimes mixed up identity and inverse and corrected themselves. Lori, for

example, said, "Under multiplication there's no identity. I am sorry; there's no inverse for 0" (Interview 1, line 71). Similarly, when discussing $\{0, 3\}$ in Z_6, Robert noted, "3 was its own identity. Was its own inverse, I should say" (Interview 1, line 232). He also sometimes used confused syntax, such as, "3 plus itself is an inverse. 3 is an inverse of itself" (line 189). And sometimes the syntax for the inverse was more appropriate for the identity, such as, when discussing Z_3, he concluded, "So now I find it has an inverse" (Interview 1, line 114). Moments later, however, he was clearer: "Let me say that again, the whole group has an identity 0, and each element in the group, 0, 1 and 2, have an inverse" (line 116).

The confusion between identity and inverse is probably best explained by the close procedural relationships between the concepts. In particular, finding the inverse of an element necessarily involves the identity. As mentioned above, when operations are presented through tables, finding the inverse of an element involves looking for the identity in the appropriate row or column. More generally, checking whether a binary operation satisfies the inverse property is a matter of checking every row and column. Robert was explicit about this procedure: "0 doesn't appear in every row and column, so not every element has an inverse" (Interview 1, line 93). The idea of looking for or creating an identity in order to find an inverse leads to procedures in other mathematical contexts as well. For example, a standard method of finding the inverse of a matrix involves performing row operations on an augmented matrix until part of that matrix looks like the identity.

Another reason for the strong connection between identity and inverse is that it seems to be natural for students to think in terms of inverse pairs, such as $\{1, 5\}$ in Z_6, or in terms

of triples that include the identity, as is described for Wendy above and for Robert below. In this kind of reasoning, elements that are their own inverses are something of a special case that becomes particularly salient when working with operation tables.

Several students noticed that an identity occurs in the diagonal when an element is its own inverse. For example, Carla demonstrated this connection when she described how she knew that a particular table was a group:

394   Carla: So if we rename {1, 3} to e, the set {1, 3} as e and the set {5, 7} as a, we will create a table that looks like e along the diagonal that goes like this and a along the opposite diagonal, which is a group because that's…. Well, for one thing it is one of the tables we came up with when we talked about possible groups for a two element set. And for another thing, we see that each of the elements appears only once in each row or column, which tells us it's group. And we see that it contains the identity, that a is its own inverse, e is its own inverse. So, we're all set. (Interview 3)

As is discussed in the section on Wendy and isomorphism below, Wendy discriminated among groups of order 4 according to the appearance of the identity element along the diagonal. Diane and Lori took longer than Wendy to notice this discriminating feature, but their description makes clear this fundamental connection between identity and inverse.

394   Diane: All the elements squared is e. Each element is its own inverse.

397   Lori: They are all their own inverses.

398   Diane: Well that's just definition. If you take $a \times a = e$, $b \times b = e$, $c \times c = e$, the only way you can get an identity element, if these aren't the identity elements themselves, these have to be inverses of each other, cause that's just the definition [of inverse].

Assuming that e is the identity under multiplication, Diane was saying, "If $a \times a = e$ then a is its own inverse." By calling this a definition, Diane was either trivializing her own reasoning or demonstrating that she did not distinguish this statement from the definition, which might instead be given by "If $a \times b = e$ then b is the inverse of a."

The connection between the identity and inverse became even more apparent when the students looked for the inverse of the identity. As Carla remarked, "The identity's inverse is itself" (Interview 4, line 133). This observation helped her verify that a group containing only the identity was indeed a group.

In summary, although the students often confused the terms *identity* and *inverse* in their language, they often corrected that language on their own. The frequent confusion between the terms is explained by procedural connections between the concepts. In particular, using an operation table to find an inverse is really a matter of looking for the identity element in the table. The confusion is further explained by the fact that the students thought in terms of inverse pairs, for which the product is the identity. Elements that are their own inverses form a special case when thinking about inverse pairs, and the identity element is always its own inverse.

Closure

The students' concept images of closure were similar to those of other group axioms in that the students' reasoning was often tied to operation tables. On the other hand, the students' concept images of closure were different in that there seemed to be fewer linguistic and conceptual confusions. Closure became prominent in service of the concept of subgroup, both in determining whether a subset was a subgroup and in constructing subgroups of a given group. A firm understanding of the concept of closure also relied on distinctions between operations, such as between addition and addition modulo 6. Because all of these issues are treated in detail in the section on the concept of subgroup, here I make only two observations.

First, the students' formal definitions were usually correct. For example, on her final exam Wendy stated that a group must satisfy the property of closure: "For all a and b in G, $a * b$ is also in G." Informal characterizations were quite close to the formal definition. Lori, for example, wrote on her final exam, "So pick any two elements in S, namely a and b and operate them together to see if the answer is in S."

Second, operation tables helped the students see whether or not closure was satisfied. Lori, who had taken the course previously, explained in her first interview that tables helped her understand the concept: "I don't think that I necessarily understood the concept of closed until we made charts and tables and stuff, and we never made tables last semester" (Interview 1, line 5).

More on Subgroups

The students' concept images of subgroup were dominated by the idea that a subgroup is a subset that is a group in its own right. The students often did not explicitly mention the operation and often made no distinctions between various related operations. These themes characterized the students' formal and informal definitions of subgroup, as well as the ways they solved problems involving the concept. In reasoning about subgroups, the students relied on operation tables, on thinking about the processes underlying the operation, and on considering each of the group axioms individually.

After providing the students' formal and informal definitions of subgroup, the bulk of this section presents an analysis and synthesis of students' responses to the question, "Is Z_3 a subgroup of Z_6?" Results of similar questions on the final exam are also presented. The section continues with discussion of two central phenomena that arose during these interviews: the sense that a subgroup should be a block within an operation table and the

belief that addition in Z_3, addition in Z_6, and addition in Z are all essentially the same operation. The section closes with analysis of the ways that the students looked for subgroups of Z_6 and, more generally, how they constructed subgroups generated by elements of groups.

Definitions. On the final exam, the students' definitions mostly characterized a subgroup as a subset that is a group, and the operation was not always mentioned. Robert's definition was typical: "A subgroup is a subset of elements from a larger group G, which form a group under G's operation." Some students did not state explicitly that the subgroup would be a group but rather listed the four group axioms as conditions that the subset must satisfy. Some students mentioned that it was not necessary to check associativity. These formal definitions were consistent with the students' informal definitions found throughout the data. What varied was the kind of attention they gave to the operation.

Is Z_3 a Subgroup of Z_6? When considering whether Z_3 is a subgroup of Z_6, most of the key participants were seduced by the fact that Z_3 is a subset that is a group in its own right. Carla, for example, after verifying that Z_3 is a group, concluded that Z_3 is a subgroup of Z_6 "because Z_3 is a subset of Z_6. That is what makes it a subgroup of Z_6" (line 173). Robert, Lori, and Diane came to similar conclusions. Wendy was alone in her early conviction that it did not make sense to consider Z_3 separately from Z_6. For most students, overcoming this initial conclusion required a coordination of resources and depended upon concluding that addition mod 3 and addition mod 6 are different operations.

As the students continued to ponder the question at hand, the operation tables and the processes behind the operations created suspicions that Z_3 is not a subgroup of Z_6, but the suspicions were usually not sufficient to convince them of that fact. Instead, the notion that a subgroup is a "subset that is a group" was strong enough to create the simultaneous belief that Z_3 ought to be a subgroup. This belief may explain the fact that the students had a strong sense that "addition is addition" despite differences between addition mod 6 and addition mod 3.

Diane and Lori, for example, knew that a subgroup must use the same operation as the group and also saw the operations in Z_3 and Z_6 as different, but that was not sufficient to lead them to a conclusion.

103   Diane: They are both modular arithmetic, they are both modular addition, they are just different mods. So it's kind of weird what you would think of mods, if you are talking.... If you take into consideration the different mods here and still consider it the same operation, then these could be subgroups. This could be a subgroup of this. But you are saying that this and this are different, then you have say they are different operations.

104   Lori: Do we define mod 3 under addition a different operation than mod 6?

Diane and Lori also considered the operation table for the subset {0, 1, 2} in Z_6, which Diane said "would definitely be a subgroup" (line 105). I asked them what the table would look like.

110   Lori: Oh, it's the same as Z_3.

111   Diane: No it's not exactly the same, because you are going to have 0, 1, 2; 1, 2, 3; 2, 3, 4 [in the table].

112   Lori: You can't have 3 and 4. They are not in the set, and then it's not closed.

113   Diane: You're right.

Thus, Diane and Lori had at least two kinds of evidence that the operations are different. Despite this evidence, Lori still wanted Z_3 to be a subgroup of Z_6. She enumerated the group axioms to support her point of view:

118   Lori: I think it is a subgroup of Z_6 under addition because, kind of like.... We were on a roll here with these things, these being, they're both closed, they both have an identity, they both have inverses, they both are associative, so that makes them both groups, and this is in here.

Diane was a bit more skeptical. She was still willing to consider Z_3 to be a subgroup of Z_6 but only if "dividing by 3 and dividing by 6 [is] just a characteristic of mod and you are going to say that it's all right to kind of ignore that" (line 122). In other words, she was unwilling to conclude that Z_3 a subgroup of Z_6 without confirmation that "all mod is fine" (line 124).

Soon thereafter, Lori and Diane went back to considering the tables:

131   Lori: Yeah, that's what I was saying because when I think of something being a subgroup of something else, its table can almost fit right into it since it's the same operation, and I don't see this anywhere down here.

132   Lori: I don't think they are subgroups of each other anymore. I was getting confused with ...

133   Diane: I say no.

Thus, no simple piece of evidence was sufficient, but rather an accumulation of evidence and consideration was necessary for Diane and Lori to conclude that Z_3 is not a subgroup of Z_6.

Robert was similarly hesitant to come to the same conclusion even in the face of evidence. He first used the table for Z_6 to show convincingly that the subset Z_3 was not a subgroup because the inverse and closure properties failed. Nonetheless, he went on to create a separate table for Z_3, and on the basis that Z_3 was a group concluded that it was a subgroup of Z_6. He was unsure whether to call the operations different: "Are we talking addition mod 6 and addition mod 3, or are we just talking addition?" (line 148). Interestingly, Robert was also the only student who was unable to resolve the issue by the end of the first interview. This fact may be partly explained by Robert's sources of

authority for making mathematical claims. At the beginning of the second interview, he announced that he had determined that Z_3 is not a subgroup of Z_6, because the operation is not the same. "For one thing, I read it in the textbook last night. And, for a second thing, Steve told me that today. He mentioned that in class" (Robert 2, line 12). Thus, Robert required evidence, consideration, and external authority to come to the correct conclusion.

For some of the key participants, the conclusions they reached during their first interview were not as enduring as their expressions seemed to indicate. On the final exam, students were asked whether Z_4 is a subgroup of Z_8 and whether Z_4 is a subgroup of Z. Carla, Robert, and Diane all pointed out that the operations were different and therefore Z_4 is a subgroup of neither Z_8 nor Z. Wendy and Lori, on the other hand, both wrote that Z_4 is a subgroup of both Z_8 and Z, arguing, essentially, that Z_4 is a group and also a subset of both Z_8 and Z. Lori's misjudgment was not surprising, for her reasoning had seemed uncertain and ambiguous throughout the discussion of whether Z_3 is a subgroup of Z_6 and throughout her interviews more generally. Wendy's response, on the other hand, is a stark contrast to her thinking in the interview. The most plausible hypothesis for the discrepancy is that, without a table in front of her, it was not readily apparent to Wendy that the operations were different.

Subgroups of Z and Z_n. The prominence of the idea of subset in the students' definitions of subgroup explained not only the sense in which the students considered Z_3 to be a subgroup of Z_6 but also the sense in which they considered them both subgroups of Z. Lori, for example, asserted, "Z_3 is a subgroup of Z. We all agree on that, right? So if they are both subgroups of Z, then maybe they are subgroups of each other" (line 142).

What was harder to explain, however, was Diane's belief that although "Z_3 isn't a subgroup of Z_6, it's at least a subgroup of Z_n, and we know that Z_n is a group under addition" (line 44). From the first half of Diane's claim, it is clear that she did not mean that Z_3 is a subgroup of Z_n for any n. What is puzzling, then, is what kind of object Z_n was for Diane and in what sense Z_3 could be a subgroup. Clearly, Z_n was not an unspecified group that could be Z_3, Z_6, or any of a number of other groups. Instead, Z_n was a different group, distinct from Z_3 and Z_6, but somehow situated so that Z_3 could be a subgroup. Unfortunately, I did not pursue this unusual idea further in the interview. Diane's final exam indicates, however, that, at least at the end of the course, her conception of Z_n was more typical.

> Z_n: Is a group under modular addition. n is a positive integer and tells which mod we are in and which elements are contained in the group $(0, ..., n-1)$. For example Z_4 is a group under addition mod 4 that contains the elements 0, 1, 2, 3.

Portions of the table. Like Wendy, most of the key participants tended to think of subgroups as blocks within the operation table of the larger group. Robert, for example, focused on the top left quarter of the Z_6 table and concluded (momentarily, at least) that Z_3 is not a subgroup of Z_6, because "0 doesn't appear in every row and column, so not every element has an inverse" (Robert 1, line 93). Robert, like Wendy, also initially ignored my suggestion that he pick individual elements from the table and instead continued to focus on blocks such as {0, 1, 2} and {3, 4, 5} (line 176). Later, after he had identified {0, 3} as a subgroup, I asked him whether that constituted a portion of the table. He replied, "Not in the sense that you are just drawing a box around part of the table. This is taking different elements out of the table and putting them into a new table" (line 199). Thus Robert, like Wendy, preferred "blocks" in the table or wholly new tables. Lori's language also suggested that she was thinking of blocks: "When I think of

something as being a subgroup, if you look at its table, you can fit in into a larger table" (Diane/Lori 1, line 129).

By the third interview, Carla's thinking seemed more flexible, although it was still tied, to an extent, to the order of elements in the table. When I asked Carla about the relationship between a table for $\{0, 2\}$ in Z_4 and the table for Z_4, she first called it a portion of the table. She described how she would reorder the table as 0, 2, 1, 3, so that "in the top left corner we'd have an imitation of the table that we just created" (line 125). Then I asked again what she would say about the relationship between $\{0, 2\}$ and Z_4.

127   Carla: I would think it would say it's a subgroup of it, for two reasons. One would be, if you just looked at ... "sub" means a smaller part.... If you do the Z_4 group table, then you have a group table and a corner of it is what you are talking about then that ... it would be a good guess to say that that would be a subgroup. But also you know that 0, 2 that table is a group table, and it is a subset of Z_4, so that means it is a subgroup of Z_4.

When the students used operation tables, they sometimes paid too much attention to the order in which elements were listed in the table and too little attention to the binary operation underlying the tables. This phenomenon may be related to the strong sense that a subgroup is a subset that is a group, coupled with an overly limited Groups-Are-Containers metaphor that made it difficult for the students to think about nonconsecutive subsets. Nonetheless, the tables served a useful purpose in organizing calculations when the students were constructing subgroups or verifying that a set was a group or subgroup.

Addition is addition. This study supports the finding in the literature that students sometimes do not pay sufficient attention to the group operation (Dubinsky et al., 1994). The above analysis shows, however, that even when they do pay attention to the operation, there is still a tendency to say that two operations are the same if they are both

called *addition*, and there is reason to believe that analogous results would hold for operations called *multiplication*.

Dubinsky et al. (1994) suggest that students' progress from thinking of groups and subgroups as sets, to considering them as sets with operations, to considering a subgroup as having the same operation as the group in which it sits. This developmental progression makes good sense and partially describes the data in this study. A key question that emerged in this study is, What is involved in making the second transition? Dubinsky et al. suggest that students need to consider the binary operation to be a function on ordered pairs from the group and then restrict the function to the subset. Then students recognize that the operations need to be the same on the group and the subgroup by coordinating their function concept with their emerging group concept.

This description, quite simply, does not fit the data in this study. First, there was no indication that the students thought of binary operations as functions. Moreover, there is no reason to believe that such a conception was necessary for success, as many students seemed to be successful without it. Regarding the sameness of the operations, all the key participants recognized that the operations needed to be the same, though not always immediately. The issue was that many of the students were willing to call operations the same despite evidence that they were different. All the key participants saw—by looking at operation tables, by considering the processes underlying the operations, or both—that the operations are in Z_3 and Z_6 are different. Nonetheless, they all concluded at various times that the operations are the same because they are both addition.

The issue concerns making distinctions, not only between addition in Z_3 and in Z_6 but also with addition in Z. The above analysis suggests that making such distinctions

requires coordination of evidence and careful reasoning. Furthermore, even after such distinctions have been made, they can become blurred in students' minds moments, days, or weeks later. These findings should not be surprising in view of the fact that addition in Z_3, Z_6, and Z are very much the same. This sense of sameness is not a misconception that must be overcome. On the contrary, this naïve idea, although incorrect, has a grain of truth that can be firmly established only through the concept of quotient group, which is introduced much later.

Constructing subgroups. In constructing subgroups, the students reasoned both from the table and from thinking about the operation. They typically began with a small number of elements and then constructed an operation table to determine whether the elements constituted a subgroup. Most participants stated, either immediately or while reasoning during the process, that any subgroup must contain 0. Robert, for example, chose {0, 2, 4} and {0, 1, 5} as possible subgroups because "0 has to be in there. And we need things that are inverse of each other" (line 230). Diane, on the other hand, initially chose {0, 2, 4} because she "went for the even numbers" (line 224). Diane and Lori were not able to find the subgroup {0, 3}, however, until having considered, as Wendy had, what else would need to be in a subgroup that began with all elements other than 3. Like Wendy, Diane reasoned largely from closure: "If you have 1 you have to have 2, and if you have 1 and 2 you have to have 3" (line 251). Lori also reasoned from the inverse property, noting that if you have 5, "you have to have 1" (line 272). Although we did not use the phrase "subgroup generated by" until much later in the course, Diane, Lori, and Wendy all seemed to pick up this sort of reasoning quite naturally.

In their fourth interview, Diane and Lori used a tabular method of finding the subgroup generated by one or more elements. They developed the method in response to such questions on the second take-home exam. I asked them to explain the method for the subgroup generated by (123) in D_3.

49   Diane: Well I know I need an identity, and I know I have the element (123), so I would go ahead and fill this in as far as it lets me. [Makes an operation table with (1), (123).]

51   Lori: That's going to give her a different element, and she does (132) (132). And then she is going to add it next to it, to keep it ... It's going to get bigger.

61   Diane: And now our table is done because we didn't generate new elements in our [table].

In other words, beginning with the identity and the generators, they constructed a table and filled in its interior. Then they expanded the table, when necessary, adding a row and column corresponding to each new element that appeared in the interior of the table. They continued this process until there were no new elements to append to the table. This process is entirely legitimate for any finite group and with any number of generators. Furthermore, whenever the process stops, the resulting set is necessarily a subgroup.

In summary, the students used operation tables, reasoning about the operations, and the identity, closure, and inverse properties when constructing subgroups. Not surprisingly, associativity was not a consideration. It is legitimate, of course, to assume associativity when the operation considered on the subset is the same as the operation on the set as a whole, but often that was not the case.

Summary. The students' concept images of subgroup may be characterized as subsets that are groups. Their reasoning about subgroups was dominated by the identity and inverse, closure properties, as embodied in operation tables, and without sufficient

attention to distinguishing among various operations called addition. For most students, overcoming this tendency "addition is addition" required multiple sources of evidence, including, for example, operation tables as well as careful reasoning about the operation.

Regarding the main interview question, these data provide additional insight into the finding in the literature that students believe Z_3 is a subgroup of Z_6. There seems to be a strong belief that Z_3 should be a subgroup of Z_6 and, similarly, that they are both subgroups of Z, because they are subsets that are groups of their own right. Furthermore, this belief is strong enough to overpower any suspicion that the operations should be acknowledged as different. As for the inappropriate use of Lagrange's theorem to establish that Z_3 is a subgroup of Z_6 (Hazzan & Leron, 1996), the data and analysis suggest that this phenomenon may not be a matter of confusing a theorem with a converse but rather a matter of grasping a seemingly relevant theorem to support a previously held conviction.

Isomorphisms

Isomorphism was the theme of the second interviews, but Robert's and Carla's second interviews focused on other topics. Nonetheless, the concept of isomorphism arose in interviews with all the key participants, thereby providing sufficiently broad data. Again this section begins with a conceptual analysis followed by detailed analysis of Wendy's second interview, which was particularly rich, and where, once again, the themes are use of language and use of the operation table. The discussion then is broadened to include other students.

It is through the concept of isomorphism that students begin to gain a sense of abstract groups. Thus, this discussion provides a way then to discuss students' concepts of groups and the nature and role of abstraction in such conceptions.

Conceptual Analysis

As discussed above, Wendy noticed during her first interview that sometimes a group's operation table can be made to look like another group's table by renaming the elements. This can happen, of course, only when the groups have the same structure, because, once again, it is the operation that gives a group its structure, and the names of the elements are structurally unimportant. When two groups have the same structure, they may be considered "essentially the same," and the groups are said to be *isomorphic*. Furthermore, both groups may be seen as instantiations of the same abstract group.

It is a hard problem, in general, to determine whether two groups are isomorphic, although it is often possible to see quickly that they are not, such as when they do not have the same number of elements. When two groups are represented by operation tables and if one believes that the two groups might be isomorphic, the naïve approach is to attempt to rename the elements of one group and perhaps reorder the elements in the operation table until the table is identical to the table for the other group. The formalization of this naïve idea is somewhat involved. The renaming and reordering are accomplished via a one-to-one function from one group onto the other. Then the task of comparing the two structures involves comparing two kinds of calculations:
(1) performing the operation on the elements in the first group and sending the result through the function, and (2) sending the elements through the function individually and combining their images under the operation in the second group. If the results of the two

calculations are the same for all pairs of elements in the first group, the function is called an *isomorphism*. Formally, an *isomorphism* is a one-to-one function f from a group G onto a group G', with operations $*$ and $*'$, respectively, such that for all a and b in G, $f(a*b) = f(a)*'f(b)$. Two groups are said to be *isomorphic* if such a function exists. Depending on the manner in which the groups are presented, it may be a very difficult task to find a function that works.

The formal definition of isomorphism, although necessary, obscures the intuitive notion that the two groups G and G' are essentially the same and thus are examples of the same abstract group. The formalization has other negative consequences as well. The idea that two groups are isomorphic is symmetric, in that if G is essentially the same as G', then clearly G' is essentially the same as G. In contrast, the formal definition is asymmetric, in that one of the groups must be chosen as the domain of the function.

The core idea, once again, is that groups that arise in different contexts might actually be different representations of the same group. This point of view provides an opportunity for profound insight into the nature of groups. For example, although there are countless representations of groups with three elements, all of them are isomorphic and thus all represent the same abstract group. It is in this sense that it is legitimate to talk about Z_3 as representing the abstract group with three elements, or, more simply, to talk about *the* group of order three.

Dr. Benson and I had as a goal for this course that students would begin to develop an understanding of such abstract groups, so that they might begin to "see" an abstract group "through" a representation. Once again, the only access to abstract objects is through

representations or, more to the point, through multiple representations. An abstract object emerges in a student's thinking when he or she begins to see a symbol as a representation rather than a thing-in-itself, but that is unlikely to happen "unless there are other symbols that can be regarded as signifying the same entity" (Sfard, 2000, p. 79). For small finite groups, an obvious mode of representation is an operation table. Thus, to pave the way for these ideas in class, the students had been asked on the take-home portion of the first midterm exam (Appendix B) to "fill out all possible operation tables which make the set $\{e, a, b, c\}$ a group," where e was assumed to be the identity. The students had found four such tables.[10] Wendy's tables are provided in Figure 11. It turns out that first three tables in Figure 11 are isomorphic and thus represent the same abstract group. The fourth table, on the other hand, is not isomorphic to any of the first three and thus represents a different abstract group. In this way, there are exactly two abstract groups of order four, just as there is only one group of order three.

Figure 11. Wendy's tables for $\{e, a, b, c\}$

①

*	e	a	b	c
e	e	a	b	c
a	a	c	e	b
b	b	e	c	a
c	c	b	a	e

②

*	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	a	e
c	c	b	e	a

③

*	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

④

*	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

To provide additional grounding for the concept of isomorphism and the ideas behind it, the class spent several days renaming and reordering tables to show that they were the same as other tables. For groups with four elements, they were asked to determine whether they got one of the four operation tables they had identified on the midterm exam (see Figure 11) and, if so, which one. Our hope was that, when asked to show that

---

[10] To be precise, there are exactly four such tables only if one assumes that the elements are to be presented in a particular order. Because all students used the order e, a, b, c, this imprecision did not present a problem.

Z_4, for example, was isomorphic to one of these four tables, the students would choose different ones, and indeed they did. Then when one student showed that Z_4 was isomorphic to Table 1 in Figure 11 and another student showed that Z_4 was isomorphic to Table 2, they might conclude that Table 1 must be isomorphic to Table 2. Dr. Benson and I thought that conclusions such as this would be obvious to students, based on their intuitions about renaming and reordering, and indeed students did make such conclusions.[11]

During this work in class, a student had suggested the word *congruent* to describe the relationship between two groups that could be made the same via renaming and reordering elements in the operation table. This term suggests the intuitive idea that establishing a correspondence between elements of two groups and comparing the operations is, in a way, analogous to establishing a correspondence between vertices of two geometric figures and then comparing the figures. Dr. Benson introduced the word *isomorphic* to give a standard name to the naïve concept of congruence that was emerging in the class, and he indicated that the term *congruent* would be acceptable as well. At the same time, he gave formal definitions of *isomorphism* and *homomorphism* (see chapter 6).

Wendy and Isomorphisms

Wendy's second interview took place just after the session in which Dr. Benson had introduced the term *isomorphic*. I had planned to discuss the concept of isomorphism in the context of the several groups of order 4 that the class had been investigating. I was getting ready to ask a question when Wendy put forward her own question:

[11] It is possible to formalize this kind of reasoning by proving that *isomorphic* is an equivalence relation, but this would have required a formal version of *isomorphism*, which had not been introduced yet.

10   Wendy: What exactly does the word *isomorphic* mean? *Iso-* meaning like, one? Does it mean one?

11   Brad: Well, the etymology of the word, *iso-* means "same."

12   Wendy: Same.

13   Brad: And *-morphic* means "having to do with form."

14   Wendy: Okay. Same form.

15   Brad: Same form?

16   Wendy: Okay. [inaudible] I don't like using a word if I don't know what it means.

17   Brad: Right. But on the other hand, understanding what the word means, "same form" and understanding how it relates to this stuff here, that's ...

18   Wendy: It really relates.

22   Wendy: Because you can reorder.... The way I explained it to the class today.... I am kind of getting a hold on this. [Laughs] (Wendy 2)

Dr. Benson's introduction of the term *isomorphic* apparently had not been enough for Wendy. Her initial confusion followed by her response "It really relates" seems to indicate that my description of the etymology of the word was both necessary and sufficient for her to attach the word *isomorphic* to the idea that she had been developing. As is shown below, however, understanding isomorphic as "same form" was not really sufficient.

Wendy then described how she was thinking about the reordering (rearranging) process.

24   Wendy: Well, the way that I do it.... Because people have.... One person in class kept on putting up, asking a question that didn't make much sense to me, but I explained my way of saying, well, say we have Table 1 and Table 2 [see Figure 11], which we kept on saying were the same.

26   Wendy: Okay? And we wanted to show that 2, Table 2 is like Table 1. But what she kept on saying, was, "Well, can you rearrange it anyway you want?" And in a sense you can, there are different ways to rearrange it to get it to look like Table 1. But you can't rearrange it anyway you want to make it look like Table 1, because the way I know how to rearrange it is that you have to look at the diagonal.

28   Wendy: There might be another way, but this is the way that always works for me. By looking at the diagonal you see that there.... And the reason why you can tell that it's different from this group altogether is because if you look at the squared ...

29   Brad: From your Table 4.

30   Wendy: From my Table 4, is that if you look at the squared elements, all the squared

elements, e, a, b, and c squared equal the same thing. But in these three tables [Tables 1 through 3], two elements equal the identity e, and two elements don't. That's why these three can be arranged to look like each other. (Wendy 2)

So, for Wendy, the differences in the diagonals indicated a fundamental difference between the groups that the tables represented. Furthermore, she saw that the appearance of the diagonal constrained the ways that one table could be rearranged to look like another.

It is worth noting that Wendy's approach is correct and insightful: The squares—in particular the number of them and how often they each appear—provide essential information (in the sense of "essence") about the structure of a group. In the case of groups of order 4, merely counting the number of squares is an accurate and efficient discriminant between the two groups of order 4. The approach is somewhat general in that if the number and multiplicity of the squares is different between two groups, then the groups are not isomorphic, although the inverse is not always true.

At this point, Wendy had a sense that the groups presented by Tables 1 through 3 were isomorphic. But actually showing they are isomorphic requires finding an isomorphism, which for Wendy was a reordering and renaming that would make the tables identical, or as Wendy said, "in the same form."

34   Wendy: But to rearrange them, if you want to see whether they are the same or not, you want to get it in the same form, hence being isomorphic. Same form. So if you want get [Table] 2 to look like 1, if you look at…. e^2 is always equal to e, so you really, you can leave e where it is. But in the first table you have $a^2 = c$, $b^2 = c$, and $c^2 = e$. That means you want to make a^2 and b^2, or these two elements in the middle of the table, their squares to be the same element where the last, the last element in the table you want to equal e^2.

35   Wendy: And if you look at the second table. These two elements, their squares equal e^2, but you don't…. You want, you want this square to equal the nonidentity square. A nonidentity square. You know what I mean?

36   Brad: So you want a^2 to be other than the identity?

37   Wendy: Other than the identity. And you want a^2 to be in this position because that's

where it is in the first table.

39   Wendy: So by this case you know that a and c have to switch. But as Steve [Benson] pointed, as we figured it out in class today, that you don't necessarily.... What I did, is I switched a and c, so the table would have, the new table to rearrange and rename to make it look like Table 1 would be e, c, b, a.

Wendy's language in the statement "you want to get it in the same form" (line 34) suggests that, for her, the form depended on the arrangement of the elements in the table. Rearranging a table would put it into a different form, and the goal was to get two tables into the same form, "hence being isomorphic." Wendy's concept of isomorphism was dominated, at this point at least, by her notion of same form. Thus, her concept of isomorphism depended on the particular arrangement of elements in the table.

Wendy used the table not only to determine which groups could be rearranged to be the same but also to determine how to accomplish a reordering that would work. By focusing on the fact that the two middle elements in Table 1 (a and b) had nonidentity squares and the fact that $a^2 = e$ in Table 2, she knew that something other than a in Table 2 must map to a in Table 1. One way to accomplish that was by switching a and c, but Wendy saw there were other possibilities.

41   Wendy: But if you look at this table, e, b, c, a works, and that is because you can take out a, the a row on this table and slide it up. And that way the diagonal would be e, a, a and if you place a down at the bottom, if you kind of take it out, slide these up and put it back on the bottom you'd get e, b, c, a and that works too. You can do it that way. So there are two ways to do it.

43   Wendy: e, b, c, a works, and so does e, c, b, a.

Wendy's explanations were partly based on manipulations of the table, such as switching rows and removing a row and sliding others up. Her written work from class demonstrates the e, b, c, a reordering and renaming process (see Figure 12). She reordered Table 2 and then used a "renaming function" in order to end up with names that were easier to compare with Table 1.

Figure 12. Showing Table 2 is like Table 1

①
*	e	a	b	c
e	e	a	b	c
a	a	c	e	b
b	b	e	c	a
c	c	b	a	e

To show (2) is like (1)

reorder   rename

②
*	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	a	e
c	c	b	e	a

$\Rightarrow$

*	e	b	c	a
e	e	b	c	a
b	b	a	e	c
c	c	e	a	b
a	a	c	b	e

$\Rightarrow$

*	E	A	B	C
E	E	A	B	C
A	A	C	E	B
B	B	E	C	A
C	C	B	A	E

$f(e) = E$ $ab = c$

$f(b) = A$

$f(c) = B$ $f(a)f(b) = f(c) = f(ab)$

$f(a) = C$ $f =$ renaming function

$\searrow$ homomorphism
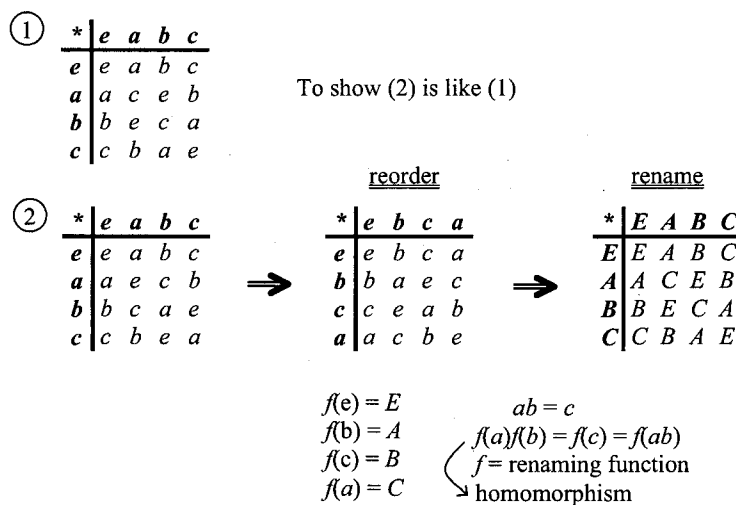
The work in class had led Wendy to several insights about renaming and reordering tables. She also had begun to ask more general questions and to sense possibilities for broader conclusions.

47   Wendy: We can do it more than one way. And that was interesting for me. But what he [Dr. Benson] was saying today that I really haven't thought much about was that all you have to show, to show that they are isomorphic, is one way, and that makes sense. You don't have to show all of the different ways that it's isomorphic.

48   Wendy: But, I wonder how many different [inaudible] ways. Like, say you have Table 3, which has *a*, *b*, or *e*, *b*, *e*, *b* in the diagonal, and you want it to look like the first one. I wonder how many different ways to reorder it before you rename there are. You know, like, there were two ways to reorder 2 to look like 1 before you rename it. I wonder how many there are for 3. I don't know.

Thus, Wendy was concerned not only about whether a table could be rearranged and renamed to be the same as another table but also about the number of ways that it could be done, despite the fact that Dr. Benson had indicated that this was not necessary in showing that two groups are isomorphic. Wendy's question demonstrates noteworthy mathematical instinct, for finding and counting the different ways to reorder and rename a group is the key idea behind the set of automorphisms of a group—a topic that we did not explore in class.

As the interview continued, I asked Wendy to write out operation tables for two familiar groups to see whether they were the same. In particular, I wanted her to compare the rotations in D_4 and a group the class came to call the *military group*, which consisted of the commands "stand as you are," "left," "right," and "about face," abbreviated S, L, R, A, where the operation was following the first command by the second. I was interested in how she would approach this task because, from a certain perspective, these groups are obviously isomorphic, in that the former consists of rotational symmetries of a square and the latter consists of the same rotations of a person as viewed from above.

Before I told her the two groups I had in mind, she was concerned that she was going to have to think of the examples. And even when I named two groups, she preferred to copy the table from her notes rather than reconstruct it, suggesting that her reasoning was largely external and based in the table. As she copied the military group from her notes, she considered reordering it as part of the copying:

59   Wendy: Do you want them in any particular order? I presume not because we're going to rearrange it anyway. But it also makes it interesting, while I am just copying this [inaudible], is that like, you.... It depends on how you set this up. You know what I mean? Like what if I give like.... I don't know. [inaudible] You can write the table down this way or in this way. Rearrange ...

60   Brad: You mean as S, A, L, R or as S, L, A, R.

61   Wendy: Yeah, and that is going to make a difference to how you are going to have to rearrange it to make it look like something else.

62   Brad: Okay. But is it ...? Whether you write it down as S, A, L, R or S, L, A, R are they, are these different operations here? Are these different systems?

63   Wendy: No, it's the same operation. But say for some reason my.... We are doing the rotations in D_4. Maybe if I wrote it this way, I wouldn't have to rearrange it. It might turn out to be the same. Or if I wrote this one, it might turn out to be the same. Do you know what I mean?

Wendy anticipated that in her copying she had an opportunity to choose an order that could make it unnecessary to reorder the table again. And even though she preferred to

look at the tables, she saw them as representing something that was independent of the order that elements were listed.

Wendy began filling out the table for the rotations in D_4 and quickly noticed a connection:

64   Wendy: Okay, the rotations of D_4 are rotate 0, rotate 90, rotate 180, and rotate 270. Rotate 90 and a 90 is rotate 180. Rotate 270, rotate 0. [Begins writing out Figure 13, Table 1.] This is like an addition table.

65   Brad: Oh, really?

66   Wendy: I think so [inaudible] because you just.... Do you know what I mean when I say it is like an addition table? Like if this was 0, 1, 2, and 3; 0, 1, 2, and 3, this is.... R_0 is the identity. And then it kind of rotates 1, it like moves up one.

67   Brad: Why don't you write it down? What the table would be if you call them 0, 1, 2, 3 just like that?

68   Wendy: Under addition?

69   Brad: Well, however you are thinking about it.

70   Wendy: I am thinking about it like an addition table. 0, 1, 2, 3. [Begins writing out Figure 13, Table 2.] Uh oh. I am thinking about it mod 4, a Z_4 addition table.

71   Wendy: You know how you always kind of ...? When you are filling out a table if you pick up a pattern, if it clicks with something else? Like this is how it clicked with me. And that's how I know how to fill out my table.

Figure 13. Wendy's connection: Rotations in D_4 as Z_4

①

	R_0	R_{90}	R_{180}	R_{270}
R_0	R_0	R_{90}	R_{180}	R_{270}
R_{90}	R_{90}	R_{180}	R_{270}	R_0
R_{180}	R_{180}	R_{270}	R_0	R_{90}
R_{270}	R_{270}	R_0	R_{90}	R_{180}

②

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Wendy noticed the connection while working in the table, not by reasoning about the rotations or addition modulo n or abstractly about the groups. This point becomes clearer as the excerpt progresses. She noticed how the elements rotated and moved up one position in the interior of the table (line 66). Furthermore, her surprise in line 70 indicates that it was not until she was carrying out this rotation process to construct the

addition table for {0, 1, 2, 3} that she realized that the operation was not addition, but addition mod 4.

To get some clarity on her thinking, I asked her why she switched to addition mod 4.

75   Wendy: It wouldn't be closed if I turn it, if I didn't switch it to mod 4.

76   Brad: Okay, but ...why is the connection here?

77   Wendy: Because with addition, modulo addition, when you have, the first row is always going to be the identity. Row or column is always going to be the identity. And then when you fill out the next row, it always increases by one, the following row, it almost like cycles. It's almost like a turn [inaudible]. Almost like when we did the permutations. Let's see if I can remember the forms. I mean the notation kind of screws me up. It's like that notation. Permutation notation. Right?

78   Brad: Okay. How, what?

79   Wendy: So, 1 goes to 2, 2 goes to 3.... Well, I have.... It should be 0, 1, 2, 3, but if we are going to look at this case. But, same difference. 0 goes to 1, 1 goes to 2, 2 goes to 3, 3 goes to 4, 4 goes to 1. Same thing as if you wrote R_0, R_{90}, R_{180}, and R_{270}. Same thing. R always goes to.... R goes to.... If you look at the rows and columns, Row 0 goes to Row 90, Row 90 goes to 180, Rotate 170 [270] would go back to zero, and it does this in all of the columns. Like, rotate 90 goes to rotate 180. It kind of moves up.

Wendy's explanation supports the point that she was reasoning from the table, for her description was about how the elements move around the rows and columns of the table and included no discussion of the meanings of the operation in D_4 or the structural aspects of the group. Wendy's statement "the first row is always going to be the identity" (line 77) suggests, however, that Wendy may have thought of the whole row as actually *being* the identity, rather than indicating how the identity acts on the elements of the group. Wendy's comparison with permutation notation is problematic, based perhaps only on notational similarities between the rows in the operation table and one of the two methods the class used for representing permutations. On the other hand, the idea of elements of the group acting on the group as a whole is yet another seed of an important mathematical idea. Furthermore, developing the connection more fully requires strong

connections with permutations, once again demonstrating Wendy's good mathematical instinct.

Moments later, when Wendy compared these tables to the table for S, A, L, R, she first wondered whether she had made a mistake in copying the table but soon suggested that switching things around might make it work.

89   Wendy: Because if you noticed in these two tables [Figure 13, Tables 1 and 2], you can write the same permutation notation because if you look at the diagonal again, it has the same form. These are in the same isomorphic form. They're iso-.... I don't know how to use the word right yet, but.... These have the same form right now. The way that these two are set up, D_4 and, I don't know, military—I don't know what you want to call it—this military form. It's not cut up in the same form, yet, that's the table.

Thus, for Wendy the form was the table, a perspective that may have made it difficult for her to separate the notion of isomorphism from the particular table (and its arrangement) used to represent a group. She had a sense that there was something that stayed the same under rearrangement, but she did not yet have the language for it.

As the interview continued, Wendy reaffirmed that her key to whether two tables were in the same form was first to look at the diagonals, and her goal was to try to get the diagonals to look the same. I asked her to do that with the military group, to try to make its diagonal look like the diagonal of the rotations in D_4.

101   Wendy: Well, I see that rotation 0 is the identity in this table, and it goes.... In this diagonal, I am going to try to set this diagonal up to look like this diagonal and see if everything else will fall into place. In here it's identity, rotation 180, identity, and rotation 180. So am going to see, since S is the identity in this table, I am going to see if I can get the same thing: the first and third squared elements to equal the identity. Like here it's the second. So I am going to try to see if I switch A and L, if I can make the tables look the same.

As Wendy tried to carry out this plan, she stumbled for a moment because she tried to switch the row and the columns at the same time but then completed the rearrangement (Figure 11, Table 3).

Figure 14. Wendy's tables for other groups of order 4

①

	R_0	R_{90}	R_{180}	R_{270}
R_0	R_0	R_{90}	R_{180}	R_{270}
R_{90}	R_{90}	R_{180}	R_{270}	R_0
R_{180}	R_{180}	R_{270}	R_0	R_{90}
R_{270}	R_{270}	R_0	R_{90}	R_{180}

⑤

*	E	A	B	C
E	E	A	B	C
A	A	B	C	E
B	B	C	E	A
C	C	E	A	B

②

*	S	A	L	R
S	S	A	L	R
A	A	S	R	L
L	L	R	A	S
R	R	L	S	A

③

*	S	L	A	R
S	S	L	A	R
L	L	A	R	S
A	A	R	S	L
R	R	S	L	A

④

*	E	A	B	C
E	E	A	B	C
A	A	B	C	E
B	B	C	E	A
C	C	E	A	B

She next renamed Table 3 to E, A, B, C, yielding Table 4. Throughout this process, she paid careful attention to the diagonals, frequently checking the correspondence along the diagonals and often filling out the diagonal of the table first.

Her process began by making each calculation and then each translation, but as she continued, the process became increasingly abbreviated. She also introduced a function f to describe the renaming from Table 3 to Table 4:

116   Wendy: I am saying that there is a function almost that puts this table to this table. And I am going to call that function.... And the function brings S.... Let's call it the renaming function, and it puts S to E, L to A, A to B and R to C. And what I realized as I was filling up this table, that when, if you have S times L for instance it is going to equal L, and instead of figuring out.... All you have to do is look at the function of L, which is A, to figure out, to rename it, to this ...

I asked her to explain how this function helped her abbreviate the procedure:

124   Wendy: So the function that we're calling the renaming function is up here. So R brings, R is renamed to C. So almost $f(R)$ is going to equal C. So that's why I was just saying R is going to be equal to C.

It took Wendy a long time to say this, and she had trouble articulating the way that the function supported the sense of equality between the groups. In particular, saying "almost $f(R) = C$" is redundant, whereas saying "$R = C$" omits the role of the function

entirely. A more accurate description would be "almost $R = C$." The function establishes

the "almost" precisely: $f(R) = C$.

She completed her calculations and described the result:

126   Wendy: So that's the new form [Table 4] after I reordered it and renamed it. So reorder and then rename.

127   Wendy: Here [Table 1] I am just going to rename. And this I am going to use as my function. So rotate 180 is equal to B. Good, it's working. [Both laugh] Rotate 270 is equal to C, rotate 0 is equal to E. Rotate 270 is C, rotate 0 is E, rotate 90 is A, rotate 180 is B, rotate 0 is E, and rotate 90 is A. And these two [inaudible] the same. By renaming, by just renaming D_4 and by reordering and renaming the military one you can get them to hold the same form.

128   Brad: That's nice.

129   Wendy: Isn't that nice?

Thus, Wendy described Table 4 as the "new form" of Table 2 and immediately went on

to rename Table 1 to create Table 5. Upon completing this process, she summarized the

relationship between D_4 and the military group and also expressed some satisfaction in

what she had shown.

At this point, Wendy still thought of the form as the particular arrangement of elements in

the table. The group did not have a form that was independent of the table, but rather the

groups could sometimes be made to "hold the same form" by the processes of renaming

and reordering. In other words, the form was not a property of the group but something

you could do with the group.

Wendy explained further:

134   Wendy: Because if you are looking at the table, the table is like a very specific form.... Like.... You know how we said that these are kind of alike, before we really started asking this isomorphism stuff. There are different arrangements you can have of the same table. Like these two [Tables 2 and 3] we are calling the same table, but they are just different arrangements of the same table. So, technically these would be the.... These are different arrangements of the same.... Like these have different forms, but they are really the same operation. So this takes form into account, this permutation.

So the *operation* and the *table* are independent of the arrangement and they take *form* in particular arrangements of the table. The permutation "takes form into account" by specifying the necessary rearrangement to get tables into the same form.

I asked Wendy whether the groups represented by the tables in Figure 13 and Figure 14 were all the same.

138 Wendy: Yeah, they are the same because you can rearrange them to be the same. They are definitely the same. But when you have it, if you have the original military one, the original, and this is D_4, they are different.... These [Tables 1 and 2] are in different forms of the same, what would you call it, the same group. Like it has the same.... Or actually, I don't know how you'd say it. They have the same form. They are iso-.... I don't really know.... I have to make this.... I think this is going to be the.... Once I make this statement right here I'll totally understand what I am doing.

Wendy saw that many of the groups of order 4 were essentially the same. As she tried to describe this, however, her use of language evolved from "different forms of the same," to "same group," to "same form." She almost said "isomorphic" but held back and instead revealed personal insights about the relationship between her language and her understanding. Perhaps she had a sense that the word *isomorphic* should be about the group and not about the particular order in which the elements are listed. In any case, this seems to have been a significant moment, for her attention turned from the processes of renaming and reordering to resolving her language difficulties.

139 Wendy: What I am trying to understand right now is whether or not these have the same form, or if these have the same form [Tables 1 and 2 versus Tables 1 and 3]. Can you say that these two have the same form, or can you say that these two have the same form? 'Cause I understand that these two [Tables 1 and 2] are the same tables once you rearrange it and rename it, but without rearranging it and renaming it, do they have the same form? They don't. Well, they do because you can rearrange it to have the same form. But I just don't know if you can say that when they're not the same yet.

140 Brad: When it's not obvious.

141 Wendy: When it's not obvious. Do you know what I mean? So what can you say about these two tables and then what about these two tables? [Tables 1 and 2 versus Tables 1 and 3.] These two [Tables 1 and 3] have the same, they're in the same form, right there and then. So can you still say these two [Tables 1 and 2] have the same form? I guess

you can because they do have the same form it is just not in it yet. It's just disguised. Wendy saw that Tables 1 and 3 had the same form, demonstrating that that the form (and hence the abstract group) was independent of the names of the elements. She was conflicted, however, as to whether the arrangement of the elements determined the form. She saw renaming and reordering as different processes with different consequences. Renaming left the form intact, whereas reordering seemed to change the form. Wendy was uncomfortable with this point of view, however, perhaps because she had a sense that the order in which elements were listed in a table ought to be unimportant.

142   Brad: Here is a question to ask about it. Does...? You might ask whether the tables have the same form, is one way of asking it. But another is to ask sort of more abstractly, do the groups have the same form?

145   Wendy: See, the tables don't have the same form. That's what I was trying to get at, but the groups do. Like they're, they definitely, these elements under their operations have the same form, because it's just the way you made up your operation table that you disguised it and made it look like it didn't have the same form. But if you take it and erase it, like we did, and like reordered it and renamed it, it really does have the same form. You can see it.

The distinction between the form of the group and the form of the table was helpful to Wendy. At this point it seems as though she wanted the phrase *same form* to be tied to the group but not to the particular arrangement in the table. But then she became concerned about her work on the midterm exam.

147   Wendy: Like we were saying with any 4-order table in our take-home exam, we had, we had that there are 4 different.... I don't.... Do you remember how, what exactly the question is that he asked us? Because I am curious to know whether we actually answered the question right on the exam or not.

148   Brad: The question was something like, "Assuming you have four elements, *e, a, b, c* where *e* is the identity, write down all the group tables that you can."

149   Wendy: All the.... So it was right that we wrote all the different arrangements, because there are four different arrangements of this group table. Like these are all group tables you can fill out.

151   Wendy: But we figured out that this one is different. These, all three of these are the same, can be rearranged to hold the same form, if you rearranged them and renamed them.

Wendy was reassured that her work on the midterm exam had been correct, but she had a sense that there was a deeper question that would also incorporate her sense that three of the four tables were the same.

153 Wendy: But, if you really look at the form and see which ones are actually different.... But I want to know what question.... And I think maybe he asked this. Which, what would you ask to get the answer, to get the two different forms?

154 Brad: You mean.... On the, the way the midterm was worded the answer was four different tables. Here they are.

155 Wendy: Right, different arrangements. What would you ask...? How many congruent or...? What question would you ask to only get...? I want to try to figure out what question you would try to ask to only get two.

Again, the interview took a new direction. Wendy knew an answer for which she could not formulate a question. In referring to the "two different forms," she was talking about abstract entities that are independent of both the names of the elements and the order in which they might be listed.

She continued trying to formulate the question:

157 Wendy: These are all little links now. I am trying to figure out exactly what's going on. So what is this actually called? Like, what are we actually doing? Like, maybe list all.... List the different forms.... List the tables.... [inaudible] Maybe ...

159 Wendy: Show the different tables of different forms? Or [inaudible] form of order 4.... Of a group of order 4.

I asked Wendy what if the question on the exam had been to write down all the groups of order 4 and began recalling some of the many groups of order 4 we had discussed in class. Wendy interrupted:

163 Wendy: I think there is an ... like endless amount of tables you could write down. [inaudible]

164 Brad: Endless, if ...

165 Wendy: If you consider all of the different operations, the different.... I am sure there are tons of different 4-like element groups.

168 Brad: The idea of having to write down this many is kind of annoying, maybe.

169 Wendy: Yeah, because there are all the same, most of them are the same thing.

170   Brad: They are the same in what sense?

171   Wendy: It breaks down to two forms. Like, if you look at any, every single one of those 4 ordered tables, there are only two forms that they can have. This form or pick any one of those three forms. Because you can.... There are only four different arrangements of those four.... This is what we figured out in the exam. If you take any 4-ordered group, it has to hold one of these arrangements. Okay? But, so, we realized when we started working with isomorphic groups that these are just different arrangements.... Like you can.... Say you wrote down this form, you can arrange any of these to look like this one. So really these three are the same, have the same form. And so if this is one form and this is one form because there is no way you can make it look like this because it has different elements in the diagonal. But you.... It has to.... Any 4 element, 4 ordered group will either hold this form or this form.

It is readily apparent that Wendy had developed some conviction about the idea that there are two groups of order 4. Furthermore, her use of the word *form* was becoming less tied to the to the table and more associated with the abstract groups of order 4.

I asked how she might reword the question from the exam:

173   Wendy: List all of the arrangements of a 4-ordered group ... which have different form? Which have a different form? Would that narrow it down?

She considered using the words *isomorphic* and *congruent* but eventually stuck to the word *form*.

189   Wendy: Or you could even say, How many forms are there? [inaudible] And what are they? And then it could be any combination of these 2.

190   Brad: You mean Table number 4 and ...

191   Wendy: Any one of those [Tables 1 through 3].

Thus, by the end of this episode Wendy had a firm conviction that there are two groups of order 4. A glance back at the beginning of the interview reveals, however, that she already had a sense of this when the interview started. What had she learned during the interview? She had changed her use of the word *form* so that it was no longer tied to the particular arrangement of elements in the table, but it is hard to point to any other learning.

During the interview, the word *isomorphic* had not been very helpful to Wendy, although it seems likely that she would still say that isomorphic means "same form." She preferred to use the word *form*, perhaps because neither the terms *isomorphic* nor *isomorphism* provide appropriate syntax for what she saw as the essential idea. When one says that two groups are isomorphic, what is it, then, that the two groups have in common? The answer is something like "their essence" or "their form" or "their structure." The term *isomorphic* is not necessarily helpful.

During this interview, Wendy was simultaneously developing concepts of the two abstract groups of order 4 and developing language to talk about them. This process involved separating her concept of the groups from the names of the elements and also from the order in which elements were listed in the table. She spent most of the interview generalizing her use of the word *form* to accommodate this abstraction. It appears that these processes can take a good deal of time and mental effort.

Mathematical habits of mind. The most prominent feature of Wendy's second interview is that she had noticed a profound mathematical idea: There are essentially two groups of order 4. During this interview, she had a sense that this idea was separate from the names and arrangement of the elements in the tables, but her reasoning was so tied to the tables that she had trouble making the separation. Furthermore, she was going a step beyond this observation and in doing so adopted an inherently mathematical point of view. She wanted to know what question to ask in order to get the answer, "There are two groups of order 4." In other words, she not only saw the mathematical elegance of this statement, she also wanted to be able to talk about it.

In retrospect, perhaps I should have been more helpful in her struggle, for it is highly unlikely that Wendy would have been able to phrase the question in anything like the conventional mathematical way: "Up to isomorphism, how many groups of order 4 are there?" (see Fraleigh, 1989, p. 112; Hungerford, 1974, pp. 76, 82). The subtlety and difficulty in the idea are perhaps underscored by the fact that the phrase "up to isomorphism" is not an obvious metaphor.

This episode is noteworthy for another reason: It illustrates some inherently mathematical habits of mind. Throughout Wendy's interviews, her calculations were sometimes slow and seemingly unaided by mathematical insight. At the same time, she often showed good mathematical instincts and asked deep mathematical questions that sometimes led to important insights. For example, she decided that 0 could not be an element of a group if the operation was multiplication.[12] She focused on the squares of elements in a group as an indication of something essential about the group. She demonstrated interest not only in how to rearrange an operation table but also in counting the number of ways that it could be done. After concluding that Z_3 is not a subgroup of Z_6, she looked at other possibilities, including Z_2. She chose a useful name, $2Z_3$, for a subgroup of Z_6 and then decided to investigate whether $4Z_3$ was a subgroup. She sought to understand the meaning of specialized terms, such as *isomorphism*, and was conscious of her language difficulties. In a later interview, she noted, "If you operate any two-cycle groups that don't equal the identity, it is going to equal a three-cycle" (Wendy 4, line 481), demonstrating seed of a good idea here: In S_3, the product of any two (distinct) two-

---

[12] Wendy did not distinguish among various kinds of multiplication.

cycles is a three-cycle. Furthermore, the observation generalizes to S_n, although it is necessary to add the hypothesis that the two-cycles overlap.

Habits of mind such as these ought to be cultivated. They might be missed, however, in a traditional class that does not encourage students to articulate their nascent ideas. From the ways that students are typically assessed and from the ways that mathematics texts are written, the implicit message is that in mathematics what is valued is the result of thinking, not the process of thinking. The case of Wendy suggests the potential of changing the message.

Other Students and Isomorphisms

As described above, the concept of isomorphism was introduced informally as a process of renaming and reordering operation tables, and the formal version was introduced later. The interviews and the students' exams together provide evidence that the connection between the formal and informal conceptions was not made very well. The students in general had a good intuitive sense of when two groups were isomorphic, though they were often drawn to other language, such as "similar," "corresponding," "the same as," "equal," and, particularly, "congruent." Their concept images were dominated by patterns and relationships they saw in operation tables and in renaming processes, which they were sometimes also able to imagine without operation tables. The formal concept definition, in contrast, was rarely evoked. In a nutshell, the students demonstrated shallow understanding of "isomorphism" as a function with particular properties but rich understanding of "isomorphic," including the ability to see two different operation tables as being the same abstract group. To illustrate this result, I describe below the definitions

the students provided on their final exams and the ways in which they used the operation tables to support their informal understandings.

<u>Definitions</u>. On their final exams, the students' definitions of *isomorphism* were not very close to the formal definitions that had been provided in class. In particular, although Carla and Robert noted that an isomorphism is a function, none of the key participants explicitly included a condition such as $f(a*b) = f(a)*'f(b)$. Instead, most students gave approximate informal definitions of *isomorphic*, using phrases such as "congruent" and "renaming and reordering." Lori's definition was the least formal:

> f) <u>isomorphism</u> Two groups are isomorphic to each other, if the groups are similar or congruent. This means that there are two totally different groups that can be renamed and possibly reordered to be represented exactly the same. The two groups have the same number of elements, they have an identity element that acts similarily and they have elements that have similar inverses. In other words, the two groups are completely congruent after renaming and reordering.

Robert's definition, in contrast, was essentially correct, although different from what had been presented in class:

> f) An <u>isomorphism</u> is a special kind of homomorphism. It is a 1-1 and onto function. Things that are the same after renaming and reordering are said to be isomorphic to each other. <u>Note</u>: a <u>homo</u>morphism does not need to be 1-1.

In class, the concept of homomorphism had been introduced as a generalization of the formal version of isomorphism. Robert had reversed the relationship, making possible a very simple definition of isomorphism. Furthermore, his definition combined formal and informal descriptions clearly and correctly. In fairness, all key participants except Lori provided answers elsewhere that suggest they knew that an isomorphism is a one-to-one and onto function that is a homomorphism. Only Robert demonstrated such clarity when asked to provide a definition of isomorphism.

<u>Renaming and reordering</u>. The act of rearranging the table was procedurally difficult to carry out and fraught with possibilities for error. For example, it was tempting to try to reorder the columns and rows at the same time without coordinating the two. Renaming and reordering were particularly difficult when the two sets of names overlapped. To overcome this difficulty, some students preferred to rename both groups to some neutral representation, although this approach brought the new difficulty of determining which element "acted like the identity."

Renaming sometimes presented conceptual difficulties as well. I asked Robert, for example, whether he could rename the $\{0, 2, 4\}$ table, just as he had the $\{0, 3\}$ table.

279   Robert: Yeah, we could maybe call the.... But the thing is, if I had renamed these 1's in
        this one, then we wouldn't have had a group under mod 6, like addition under mod 6.
        But if I was going to do something similar to that, I would just call this 0, 1, 2.

Thus, renaming can present cognitive obstacles when the operation has a meaning because the meaning of the operation must change to accommodate the new names. This obstacle is related, of course, to Wendy's concern about attaching the name $2Z_3$ to the subgroup $\{0, 2, 4\}$ in Z_6. In mathematical discourse, one talks about isomorphisms (and homomorphisms) as preserving the group operation, and abstractly that is accurate. There is a sense, however, in which renaming modifies the operation, or at least the way one must think about it.

<u>Seeing form in the table</u>. Tables were very present in the students' concept images of isomorphism, particularly for groups of order 2, 3, and 4. Without prompting from me, the students often noticed, usually based on patterns in the table, that a group given by one table was isomorphic to another group. Carla noticed, for example, that both a group

whose elements were the sets {1, 3} and {5, 7} and the subgroup {0, 2} in Z_4 could be

renamed to what she might have called the {e, a} group (see Figure 15).

97　　Carla: So if we rename {1, 3} to e, the set {1, 3} as e, and the set {5, 7} as a, we will
　　　　create a table that looks like e along the diagonal that goes like this, and a along the
　　　　opposite diagonal, which is a group because that's.... Well, for one thing it is one of the
　　　　tables we came up with when we talked about possible groups for a two element set. And
　　　　for another thing, we see that each of the elements appears only once in each row or
　　　　column, which tells us it's a group. And we see that it contains the identity, that a is its
　　　　own inverse, e is its own inverse.

116　　Carla: So again if you rename 0 to e and 2 to a, you end up with e's on the diagonal and
　　　　a's on the opposite diagonal, just like the table of the left coset.[13]

Figure 15. Carla's groups of order 2

×	{1, 3}	{5, 7}
{1, 3}	{1, 3}	{5, 7}
{5, 7}	{5, 7}	{1, 3}

	0	2
0	0	2
0	2	0

	e	a
e	e	a
a	a	e

From the way Carla discussed the diagonals, it appears that she was noticing visual

patterns in the tables. When she noticed these patterns, she did not use the word

isomorphic, yet when I asked her what was the relationship between the first two tables in

Figure 15, she responded immediately and firmly, "They are isomorphic.... Yeah.

Congruent" (lines 118, 120).

For Wendy, the diagonal of an operation table was a distinguishing feature in groups of

order 4. The diagonal was particularly salient for other students, as well, particularly

when it contained only the identity element. Carla was momentarily convinced, for

example, that a group of order 3 should have the identity along the main diagonal, which

is impossible for a group of order 3, although it is necessary for groups of order 2 and

works for one of the two possible groups of order 4.[14]

---

[13] Carla's use of the phrase "the left coset" is explored in detail in chapter 6.

[14] These statements assume that elements are listed along the rows in the same order as along the columns.

Diane noticed similar relationships in her fourth interview, pointing out that two groups were isomorphic to the group she called e, a. At the end of the interview, as she gathered her papers, she looked at two operation tables and said, "Oh, wait, wait, wait. These are isomorphic!" (line 551). When I asked her to think also about Z_2, she asserted, "It has only one possible table for all groups with 2 elements.... So anything of order 2 will be isomorphic to each other" (lines 561, 563). Thus, Diane had begun to develop a concept of the abstract group of order 2.

Robert naturally renamed group elements to representations that were more familiar, and he did this even in his first interview, before there had been any explicit attention to the idea of renaming. For example, while he was considering whether $\{0, 3\}$ is a subgroup of Z_6, he noted in passing that "You couldn't call it Z_2" (line 189). Because his comment suggested he saw a connection, I asked him what Z_2 looks like.

193 Robert: Just replace all these 3's by 2's. Oh no, what am I saying? Replace all of these 3's by 1's. So that's what it would look like.

When I pursued this connection again later, he disagreed that the subgroup $\{0, 3\}$ was "like" Z_2. Instead, it had reminded him of Z_2 only because it had two elements (lines 220-225).

Robert also eventually renamed the table for the subgroup $\{0, 2, 4\}$ in Z_6 as $\{0, 1, 2\}$ and noticed that the table was then the same as Z_3. The process led him to a more general conclusion:

318 Robert: All right, cool. Well, I am thinking now that if you have a 3-element set, no matter what we call the elements, you get the same type of table.

It seems likely that Robert was seeing the form in the table. He frequently noticed symmetries, "cycling," and other patterns in the operation tables, as did all the students.

Robert was alone, however, in calling such features "geometric," and saw even more surprising connections, such as analogizing tables to matrices and renaming to performing row operations. Independent of Robert's unusual associations, it is clear that patterns in the tables were involved in his emerging concept of abstract groups.

I interviewed Diane and Lori on the concept of isomorphism the day before I interviewed Wendy and before the word *isomorphic* had been introduced. Nonetheless, their work in class had provided sufficiently rich experience for me to investigate how they were thinking about the ideas. They were convinced from their work on the midterm exam that their four tables exhausted the possibilities for tables with four elements. Furthermore, from their work renaming and reordering other tables with four elements, their conviction had become deeper:

5   Lori: Right. These are the only four, no matter what ... [inaudible]. No matter what group of order 4 you make a table of, it's going to be congruent to one of these four after renaming them and reordering, I think.

Based on this knowledge, they tended to fill in tables based on the patterns and sometimes provided justification for their actions. At the start of the interview, however, they were skeptical as to whether any of their tables were "similar" or "congruent" to each other. Nonetheless, they proceeded to rename and reorder tables, eventually showing, correctly, that three were isomorphic. I asked them what they thought about the three tables that they had shown to be isomorphic.

411   Diane: I think they are the same table. They could be the same table. Like they came from the same abstract table.

413   Diane: Well, I mean, just because you rename and reorder something doesn't change what it means, what it defines. Like you can call number 1 *a*, you can call it whatever you want to, but it still stays number 1.

So for Diane, neither the names of the elements nor the order mattered for what the table "means, what it defines." Through or in each of the three tables, she saw the same abstract table, where what matters is "the value that [the elements] hold" (line 431).

The abstract group. The fact that several students used the $\{e, a\}$ group as a canonical representation of a group with two elements conflicts with my suggestion that that the metaphor "Z_3 is the abstract group with three elements" (Lakoff & Núñez, 2000, chapter 3) is backwards. For these students it may be that they were noticing isomorphisms not with the abstract group with two elements but with the $\{e, a\}$ group. In order to notice an isomorphism, however, they must have had some notion of the abstract group, yet it is possible that what they were seeing was patterns in the table rather than the abstract group.

On the one hand, using $\{e, a\}$ as the canonical representation of the two-element group makes perfect sense because then the letters can be anything. One could argue that the letters are the names of the elements, but some students saw the letters as variables that can take on any "values." On the other hand, the representation $\{e, a\}$ does not provide any support for thinking about the underlying binary operation. In fact, it is hard to imagine an underlying mechanism that would give meaning to an operation on elements that themselves seem to carry no meaning. In contrast, Z_2, as $\{0, 1\}$, brings plenty of meaning for the operation. In fact, Z_2 brings so much baggage from operations on integers that it is hard to think of $\{0, 1\}$ as representing something else.

Another explanation is that it may be easier to see (and remember?) $\{e, a\}$ as an object. Diane demonstrated this possibility by calling the group "e, a, a, e," listing all the entries in the interior of the table (Interview 4, line 559). The table clearly supported object

conceptions, for the students thought of it as something that could be acted on as a whole, compared as a whole with something else, and manipulated as a whole (via renaming and reordering) in ways that do not change its fundamental nature.

Comparing this result to the findings about the concept binary operation reveals a contradiction. On the one hand, Dr. Benson and I introduced diamond to represent a generic (particular but unspecified) operation, intending that it would stand for any operation. We found that the students saw it as a new operation, distinct from whatever addition and multiplication might make sense on the set. Yet, on the other hand, we also introduced a new group on the letters $\{e, a\}$, notationally distinct from familiar groups, in hopes that students might notice that it is "essentially the same" as some of those familiar groups. We found that students treated it, in a sense, as a generic group, capable of representing any of a number of specific groups.

So who is right? What is the difference between a generic object in a category and a new unfamiliar object in that category? Are the two cases different? In the case of the abstract group with 2 elements, is there a way to represent it generically? Is it better to consider Z_2 or $\{e, a\}$ as the canonical representation of the abstract group with 2 elements? The data and analysis above suggest that it might be most profitable to imagine the abstract group with 2 elements as something that lies in a coordination between Z_2 and $\{e, a\}$. That way, both process and object conceptions are supported, and, more importantly, neither representation is seen *as* the group.

Summary

The students developed rich, nuanced, and largely informal concept images of isomorphism, based on processes of renaming and reordering operation tables. They

used operation tables to see when two groups were isomorphic and to construct

isomorphisms. These processes supported the emergence of concepts of abstract groups,

independent of the names of the elements and the order that elements were listed in the

operation tables. Through these experiences, students came to see that there is one group

of order 2, one group of order 3, and two groups of order 4. These results suggest that the

operation table provides a viable experiential root for the concept of isomorphism.

Students had considerable difficulty, however, articulating formal versions of the concept

of isomorphism. No doubt, this result stems partly from the manner in which the

concepts were introduced (first informally, then formally), but it seems unlikely that it

would be productive to introduce the formal definition without having developed some

sense of what was being formalized. Thus, the pedagogical problem is what might be

done to help students connect these informal understandings with the formal version.

Among the concepts investigated in this study, the concept of isomorphism is perhaps the

most striking example of the general problems of connecting formal and informal

conceptions, learning to use quantifiers, and learning to reason from definitions. I return

to these issues in the chapter 8.

Groups and Abstraction

Most of the analysis above discusses the students' understandings of the groups Z_n and

groups given via operation tables. As was mentioned in chapter 4, the students in this

class also had experiences with U_n, D_n, S_n, as well as other standard and nonstandard

examples. There were less data on the students' understanding of these classes of groups

than on Z_n. Nonetheless, there were sufficient data to support a few observations.

Regarding the groups U_n, the observations are similar to what was detailed above about the groups Z_n. The students had a tendency to view the elements of U_n as though they *are* integers, whereas a more sophisticated point of view is that they are equivalence classes of integers, or that they are merely arbitrary names for an abstract group. The students often were not immediately sure of the operation in U_n and carried properties of the elements as integers over into U_n.

Students often had trouble writing down the elements of the dihedral groups D_3 and D_4 and were even unsure of how many elements there should be. Robert, for example, described 12 elements of D_4, not realizing that he had listed 4 elements twice. This was particularly true when the groups were represented as geometric transformations, using letters, such as R_{90} for a 90-degree rotation and H for a reflection about a horizontal axis. The students also had difficulty when the groups were represented as permutations of the vertices, although, in this case, the difficulty seemed to have more to do with the permutation notation.

The students took a long time to become comfortable with the cycle notation for permutations in S_n and openly expressed frustration early in their learning. Wendy, for example, complained, "This notation drives me nuts…. It scares me. Do you see how intimidated I am right now?" (Interview 2, lines 218-220). The students had trouble maintaining the distinctions between the meanings of the array and the cycle notations, and the notational confusions were sometimes compounded by the similarities between the array notation and the rows of an operation table. As the course progressed, however, all the key participants used the cycle notation fluently and with few errors in finding subgroups, cosets, and quotient groups in D_3 during their fourth interview.

The data and analysis suggest that with growing fluency, students' expectations are more often fulfilled, in the sense that they know ahead of time what the operation table ought to look like, whether a subset is likely to be a subgroup, or whether two groups are isomorphic. This idea makes good theoretical sense and seems to fit the data in this study, although it would be hard to verify empirically because expectations often remain implicit. This observation suggests that the literature on the learning of abstract algebra would benefit from a notion called something like *group sense*, analogous to number sense (Greeno, 1991; Markovits & Sowder, 1994) or symbol sense (Arcavi, 1994), to describe particular kinds of fluency and proficiency that students might develop as they gain familiarity with the examples, notations, language, and results of group theory and the objects and properties that they are supposed to represent.

The difficulty that the students experienced with the standard examples of groups presents another pedagogical dilemma. On the one hand, abstract algebra is about abstraction, intended to rise above specific examples to see generalizations that apply to whole classes of mathematical systems. On the other hand, the students spent much of their time in this class making sense of specific examples, such as Z_6 and D_3, and, more generally, classes of examples (Z_n, U_n, D_n, S_n) that were not available to them previously. The number systems of school mathematics (natural numbers, integers, rationals, reals, complex numbers) do not seem diverse enough for students to develop rich, robust, and sufficiently abstract concepts of group, subgroup, and isomorphism. Yet if students spend most of the course developing an understanding of these new classes of groups, they are similarly unlikely to develop sufficiently rich and abstract concepts. Dr. Benson

solved this dilemma by covering less material than is covered in some abstract algebra courses. What kind of balance makes sense?

Main Themes

In the preceding discussion of the students' concept images of group, isomorphism, and related concepts, two main themes emerge. First, the students' use of language and notation was often imprecise, as they blurred distinctions between closely related signifiers and the ideas they are intended to signify. In particular, the students confused associativity with commutativity and inverse with identity, and they did not distinguish among various operations called addition. Furthermore, although their informal understandings were often rich, the formal definitions they provided often lacked quantifiers and were otherwise imprecise.

Second, much of students' reasoning and the procedures they used were based in operation tables. Tables served to mediate abstraction in that the students could work with a concrete representation in order to gain access to abstract objects and their properties. Using tables allowed the students to develop procedures for checking the group axioms and for constructing subgroups. They could see isomorphisms from the patterns present in various tables. Tables served a metaphorical role in the sense that thinking about the table helped students think about the group it represented.

Sometimes the students' reasoning seemed to be largely external and based *in* the table rather than in thinking about the processes underlying the operations, suggesting that the students' internal representations were rather limited. In such cases, the table often served a metonymic role in that it *was* the group, rather than a representation. This kind

of thinking was limited in that it hindered the students' abilities to see subgroups composed of nonadjacent elements in the tables and to see the operation independent of the order in which elements were listed in the table. As the students gained experience renaming and reordering tables, they began to overcome these limitations by separating the table from the group—the signifier from the signified—thereby developing concepts of abstract groups.

The theme of language use is an example of a larger issue of making distinctions between related ideas and being precise. The theme of reasoning from the table is an example of the larger idea of managing abstraction. These themes are further developed in the chapters that follow.

CHAPTER VI

HOMOMORPHISMS, COSETS, AND QUOTIENT GROUPS

The chapter details the students' concept images of homomorphisms, cosets, quotient groups, and related concepts, which were the focus of the third and fourth interviews. The organization is by mathematical content. The bottom-up analysis of these interviews revealed a number of themes that are presented in this chapter. The most prominent among these themes was the issue of naming—that is, the relationship between a concept and its name. A related theme was the students' use of notation, particularly regarding the distinction between a set and one of its elements. In several episodes, notational issues were central, and results from these episodes are collected together as a separate section following the section on cosets. Another prominent theme is that the students had developed considerable proficiency with many of the concepts. Their calculations were often guided by correct expectations, they continued to use operation tables to support their thinking, and they were comfortable with many of the processes and objects.

Because these themes are well illustrated by Carla's interviews, the chapter is based on detailed analyses of her concepts, names, and notations for the topics of homomorphism, coset, and quotient group. The chapter also includes a number of shorter episodes that amplify and clarify some of the issues raised by the analysis of Carla's interviews. In many of these episodes I intervened during the interview, trying to encourage standard language or notation. Thus, the analysis of these episodes provides insights into the

222

learning that is required to make new distinctions and to change one's use of language or notation.

Homomorphisms

The common core of the third interviews consisted of determining whether a particular function was a homomorphism and finding the cosets of its kernel. This section begins with a short analysis of the concept of homomorphism, which is followed by an analysis of Carla's concept of homomorphism. The discussion is then broadened to include other key participants, focusing on their definitions and the manner by which they verified that a function was a homomorphism.

Conceptual Analysis

A *homomorphism* is a function f from a group G to a group G', with operations $*$ and $*'$, respectively, such that for all a and b in $G, f(a*b) = f(a)*'f(b)$. The idea is that the function preserves the group operation (and hence some of the group structure) in the sense that it does not matter whether the operation occurs in G and the result is sent through the function or, alternatively, the elements are sent individually through the function and their images are combined under the operation in G'.

In the class that provided the context for this study, the concept of homomorphism was introduced as a generalization of the formal version of isomorphism, accomplished by dropping the requirement that the function be one-to-one and onto. An isomorphism completely preserves a group's structure. A homomorphism, in contrast, may preserve some of the structure and collapse the rest. Structure is collapsed by mapping elements in the domain to the identity in the codomain. The set of these elements is called the *kernel*

of the homomorphism. The structure that remains is the range of the homomorphism, which is the *image*, $f(G) = \{f(g) \mid g \in G\}$, of the entire domain. Because such structural relationships may be further explored via the concepts of coset and quotient group, the concept of homomorphism provided some of motivation and context for these more advanced concepts.

A typical task involving the concept of homomorphism is verifying that a given function is a homomorphism. For the third interview, I chose a function f from U_8 to Z_4, given by $f(1) = 0, f(3) = 0, f(5) = 2$, and $f(7) = 2$, where U_8 is $\{1, 3, 5, 7\}$ under multiplication modulo 8 and Z_4 is $\{0, 1, 2, 3\}$ under addition modulo 4. When the function is given formulaically, the verification that it is a homomorphism is an exercise in symbol manipulation. Without a formula here, however, the students needed to verify that $f(a*b) = f(a)*'f(b)$ for all 16 pairs of elements a, b from U_8.

Carla and Homomorphisms

I began the third interview by asking Carla to explain what homomorphism means. She gave a reasonably complete definition:

4   Carla: Okay. A homomorphism is a function from one group to another, may not be the same group, and what makes it a homomorphism is, say, you have a and b in the first group. Then $f(a*b)$, which is just whatever function makes that a group, that has the same value as $f(a)*'f(b)$, and $*'$ is the operation that makes the second group a group.

Although the necessary quantifiers for a and b were not explicit, it became clear later that she intended them to be "arbitrary" (line 16). In saying, "whatever function makes that a group," she may have meant "whatever operation" and merely misspoke. On the other hand, she may have been confusing f with the group operation, a possibility that seems

more likely after analyzing more of the interview. I asked her to give an example of a homomorphism.

8   Carla: Well, we were just talking about in class a few minutes ago, f of Z to Z_4.

10   Carla: We know that the integers are a group under addition, so our * is going to be equivalent to adding. And we know that Z_4 is a group under addition mod 4, so we know that *' is addition mod 4. So if we pick an a and a b in the set of integers—let's say a is 4 and b is 10—then $f(a*b)$ is the same thing as saying $f(4 + 10)$, which equals $f(14)$.

11   Carla: So when we find $f(14)$, because we are going from Z to Z_4, we want to separate Z into little subsets, to know what maps to what. And we know $4n$, where n is an integer, always maps to 0, $4n + 1$ maps to 1, $4n + 2$ maps to 2, and $4n + 3$ maps to 3. So if we look at 14, we can say that that is the same as $4n + 2$ where n is 3. So I will just write $f(14)$ as $f(4n + 2)$ just so that.... It's easier to find what that equals. Since it's $4n + 2$, it equals 2. So by definition of a homomorphism we should also get 2 when we do 4 of a, I mean, $f(a)*'f(b)$.

12   Carla: So that is the same as saying $f(4) + $ (in mod 4) $f(10)$. We can rewrite $f(4)$ as just $f(4n)$ where n is 0, and we can write $f(10)$ as $f(4n + 2)$ because ... when n is 2, $4n + 2$ is 10, and that helps us to look at which little subset we are dealing with, so we know what it maps to. So $f(4n) = 0$, when you add mod 4 with $f(4n + 2)$ which equals 2 and $0 + 2$ in mod 4 is 2. So we have $2 = 2$, and we know it is a homomorphism.

Several issues are raised by Carla's statements here. First, Carla specified the domain and codomain of the function but not the function itself. From what follows, however, it becomes clear that she was assuming the canonical homomorphism from Z to Z_4 given by $x \mapsto x \bmod 4$, a function that is suggestive of the group operation in Z_4.

Diane, in her third interview, also specified homomorphisms implicitly in ways that suggest that her concept of homomorphism was severely constrained by her concepts of function and binary operation. The phenomenon of implicit homomorphisms is explored in detail in chapter 7. A related issue is that this particular homomorphism mirrors the construction of Z_4 and reflects Carla's understanding of modular arithmetic, which is also explored further in chapter 7.

Finally, in specifying the subsets as $4n$, $4n + 1$, $4n + 2$, and $4n + 3$, Carla was taking generic values to stand for whole sets, yet she allowed n to take several different values

during her verification. Thus, she was not distinguishing sets from elements; neither was she distinguishing different elements from each other. This issue is discussed further below in the subsection entitled "Wendy and $N + N$." It is not clear whether Carla saw the partitioning of the domain into subsets as *defining* the function or as a description of a function she had already defined via the ambiguous statement in line 8.

With either interpretation, and ignoring the seemingly insignificant error in lines 11 and 12 stating the wrong value of n, by the end of line 12 Carla had given the gist of the verification of the homomorphism property for a specific case. I asked her whether she had proven that it was a homomorphism.

14 Carla: No, this is an example of.... If I was told.... Well, this isn't proving that it is a homomorphism, it's just saying that.... It's an example of why it is a homomorphism.

15 Brad: Okay. In order to prove it what would you need to do?

16 Carla: Um, you would have to take an arbitrary, two arbitrary elements in Z, and you have to prove that—say, they are a and b—$f(a*b) = f(a)*'f(b)$.

Here she demonstrated a good sense of both the quantifiers in the definition of homomorphism and also what would be involved in proving that the function f was a homomorphism.

I then asked Carla to describe the relationship between her function from Z to Z_4 and $f(x) = x$ mod 4, as it had been defined in class.

19 Carla: $f(x) = x$ mod 4 is just telling you that you are going to change whatever x is to a mod 4, so you are always going to get 0, 1, 2, or 3 for your answer. And when I changed the integer to $4n$, $4n + 1$, $4n + 2$ or $4n + 3$, that was just because then I could look at whatever integer I am adding to $4n$, and I know what my answer is. This is how I simplify. For me it makes it simpler.

21 Carla: Well, if it's small numbers, obviously, I don't have to go through that but.... Or especially if I was trying to disprove something or prove that it's a homomorphism or anything, I would need my $4n$, $4n + 1$, $4n + 2$ or $4n + 3$. So, it just makes ... I think in explaining it, it just makes it a little bit more understandable ... because you are dealing with things in the same kind of context.

Carla did not really answer my question, but her description of the process seems to indicate that she thought of her method as a simpler version of the function from class. It is not clear to what extent she would have been able to consider an entirely different function from Z to Z_4.

I next asked Carla to identify the kernel of the homomorphism.

23   Carla: The kernel is $4n$, where n is an integer because any integer, any multiple of 4 is always going to map to 0, and 0 is your identity in Z_4 under addition.

Because Carla did not write anything to accompany this statement, I have guessed that she meant n rather than N, on the basis of what she had been writing previously. With either interpretation, however, her statement is incorrect. There are at least two ways to correct the statement: "The kernel is the set $\{4n$ where n is an integer$\}$," or "The kernel is $4N$ where N is the set of integers." In any case, Carla was not distinguishing between a set and its individual elements. This issue has arisen several times in the results reported in this chapter and is treated in detail below in the section entitled "Notational Issues."

To determine whether f was a homomorphism, Carla first listed the elements in the groups and then explained her plan:

33   Carla: So first I am going to try an example where I am going to pick two elements that map to the same thing, and then if that isn't a counterexample, I'm going to try two elements that map to different things.

34   Carla: So first one is going to be.... I want to show that f of.... Let's see U_8 is a group under ... [pause].... I am trying to think if it is a group under addition or multiplication. But it must be multiplication, because if it was addition then 0 would be in there.

35   Brad: Why?

36   Carla: Because 0 is the identity for addition. So it must be under multiplication.

Although Carla was momentarily unsure of the operation in U_8, she was able to determine the operation quickly by reasoning from the group axioms, demonstrating some proficiency with the groups involved. She then verified the homomorphism

property for two specific cases, but rather than have her verify all the cases, I chose to move on.

I asked Carla to identify the kernel of the homomorphism.

58   Carla: The kernel is 1 and 3 because $f(1) = 0$ and $f(3) = 0$ and zero is the identity in Z_4.

59   Brad: How do you write that down?

60   Carla: The kernel? [Mm.] It is a set. [Okay.] The kernel is the set that contains the elements 1 and 3. [Writes $\{1, 3\} = \ker(f)$.] And that contains the kernel of f.

Carla was correct, although her language suggests that she may not have been distinguishing between the elements 1 and 3 and the set $\{1, 3\}$.

On the surface, the concept of homomorphism seemed relatively unproblematic for Carla. By not explicitly describing her homomorphism from Z to Z_4, however, she indicated that her concept image of homomorphism may have been tied, in potentially limiting ways, to her concepts of function or binary operation. That turned out to be a major obstacle for Diane, as is detailed in chapter 7.

Other Students and Homomorphisms

Almost all the key participants were able to recite a definition of homomorphism and describe the concept in several different ways. Furthermore, they were able to check that the property held for specific elements and were able to use the homomorphism property in proofs of other ideas, although quantifiers were often not explicit. These general observations are elaborated below in a description of students' definitions of homomorphism and the ways that they verified that a particular function was a homomorphism.

Definitions. In the students' definitions of homomorphism, the equation $f(a*b) = f(a)*'f(b)$ was ever present, but few students were careful to state what a, b, and f were intended to be. In the interviews, the students demonstrated varying degrees of sophistication in their concept images around this central equation. Carla's definition, given above, was nearly complete. In contrast, consider the following responses by other students to the question "What is a homomorphism?" from the third interviews:

7   Diane: I think it's like a law of functions. It's a rule.

8   Brad: Tell me more about what you mean by that.

9   Diane: It's a property that we can use for functions and stuff. So if something is a homomorphism, then that means that those elements can follow the rule that it's defined by.

10   Brad: And what's the homomorphism rule?

11   Diane: That $f(a*b) = f(a)*'f(b)$.

12   Brad: Okay. And what kind of function does f have to be in order ...

13   Diane: I think that would be defined later. Like function, it could be any function it wants, but as long as it satisfies this, then it's a homomorphism.

4   Robert: It's a function that takes one group to another group and preserves the group operation.

5   Brad: What do you mean by "preserves the group operation?"

6   Robert: Well, I understand it better in symbols as like $f(a*b) = f(a)*'f(b)$.

7   Brad: Okay, what does that mean, sort of in words?

8   Robert: It means it doesn't matter whether you compose the functions first ... compose the elements first and then take the function of it, or whether you take the function of each element individually and then compose them.

Thus, for Diane homomorphism was a rule or a law. Robert, on the other hand, had effectively three different definitions of homomorphism: a structural one, a symbolic one, and a verbal procedural one. Of the key participants, only Carla gave a reasonably complete characterization. The results were similar on the final exam, in that all the key participants gave definitions of homomorphism that included the equation $f(a*b) =$

$f(a)*'f(b)$, but only Carla's definition was complete, using quantifiers and specifying that

f was a function between groups.

Verifying homomorphisms. In the third interview, all the key participants were given the

same function as Carla had been given and were asked to determine whether it was a

homomorphism. Like Carla, Wendy verified the property for a few examples, described

what would be necessary to prove it in general, and then was comfortable assuming that it

was indeed a homomorphism when I suggested that we go on. Diane had considerable

difficulty with the verification because of her understanding of functions, as mentioned

above. Below I describe the ways that Robert and Lori used the operation table to

complete the verification, but first, I provide some general observations.

While verifying that the function was a homomorphism, the students had minor

difficulties, such as arithmetic errors or forgetting to send $a*b$ through f. Mistakes such

as these are easy to make, of course, when the elements in the domain and codomain look

the same. The students were often able to avoid or overcome such difficulties by relying

on their proficiency with the groups involved, much as they paid attention to the names of

elements and the group axioms to determine which group they were in. The distinction

between $*$ and $*'$ in the homomorphism equation seemed to support this process.

Robert was confused at first because he did not "really understand what f is" (line 16).

After I suggested that he set aside that concern, he was able to try an example. I then

asked whether operation tables might help, intending to explore Robert's understanding

of a method that had originally been suggested by another student in class. The method

simplifies the tedious process of checking the 16 pairs of values a, b from U_8 by using

two tables to organize the calculations of $f(a*b)$ and $f(a)*'f(b)$. If the tables match, then the function is a homomorphism. The method is especially appropriate for small finite groups in tasks such as this one where the homomorphism is not given by a formula. Upon completing the tables, Robert concluded, "You get the same table we got the first time. So that convinces me that this is indeed a homomorphism" (line 109). He was able to explain how specific verifications were represented in the table. And he again demonstrated a geometric perspective with his observation, "You could take this table and stick it right on top of the other one, and you have the identical table" (line 124).

Lori, in contrast, decided on her own to use operation tables to verify that the function was a homomorphism, although she stated ahead of time that she was not sure she remembered how to do it. As it turned out, she applied the method incorrectly so that both her tables recorded $f(a)*f(b)$, making equality automatic even if the function had not been a homomorphism. When I pressed her to describe the process, she believed that that she was computing both $f(a*b)$ and $f(a)*f(b)$.

I constructed another function that was not a homomorphism and asked Lori to check whether it was a homomorphism. Much of the rest of the interview was spent trying to help her make explicit connections among the specific pairs a, b that she checked by hand, the results that she had recorded in her two tables, and the results in a third table that I suggested. By comparing the three different tables, she was able to see what was wrong with her previous process and was able to articulate more clearly the intended and procedural differences between the two tables.

Summary

The students' concept images of homomorphism were dominated by the equation $f(a*b) = f(a)*'f(b)$ and not always with sufficient attention to specifying a, b, f, and the operations. Given a function between groups, they were able to check that the property held for specific pairs of elements and to talk about what would be necessary to verify the property in general. There was the slight cognitive difficulty of keeping track of which group an element was in, but the students often relied on their proficiency with the particular groups to manage this difficulty.

These results suggest that the concept of homomorphism is particularly susceptible to procedural approaches. Typical tasks involve rule-bound symbol manipulation, but even the nonstandard task used in this study was vulnerable to the creation of procedures that can be misapplied, as Lori demonstrated. The concept of homomorphism itself did not seem to raise many issues that impeded the students' progress with the tasks at hand. Instead, issues arose in connection with other concepts, particularly the concept of function. This observation arises again below, because the students' concepts of cosets were sometimes limited by their association with homomorphisms.

Cosets

With the concept of coset, the students' concept images blossomed once again with linguistic, notational, and conceptual issues that arose and were sometimes addressed during the interviews. After a brief conceptual analysis below, the stage is set once again by Carla, who demonstrated both considerable proficiency with the concepts and procedures and also nonstandard language and notation. The section continues with

analysis of an episode from an interview with Diane and Lori, who demonstrated similar nonstandard language. Finally, the analysis is broadened to include other students. Notational issues are elaborated separately in the following section.

Conceptual Analysis

Just as a subgroup provides structural information about a subset of a group, the set of cosets of a subgroup reveals information about how the subgroup fits within the structure of the group as a whole. Through a generalization of modular arithmetic (see chapter 7), the idea is to categorize elements of a group according to their relationship with the subgroup. The categories are called *cosets*. Realization of the structural power of cosets requires also Lagrange's theorem and the concepts of normality and quotient groups. Here I focus only on the concept of coset.

To simplify the language and notation in the following discussion, I leave the group operation implicit, as in *ah*, and call the group operation *multiplication*. The ideas and results hold for a group with any operation via simple translation. This is the power of the abstract concepts of group and binary operation. I in no way intend, however, to trivialize the cognitive requirements in making the translation to a group with an operation that is called something other than multiplication.

If H is a subgroup of a group G, and $a \in G$, then the *left coset* of H containing a is defined by the formula $aH = \{ah \mid h \in H\}$. Right cosets are defined analogously. Computing a particular coset aH requires multiplying a particular value a from the group by each element in H. Computing all the cosets involves completing such calculations for all elements a in the group. Either of these processes may be infinite, and such

situations carry the additional requirement of being able to describe the results without actually completing the processes. For finite groups, it turns out that the cosets of a subgroup partition the group into subsets that each have the same number of elements. Lagrange's theorem—that the order of a subgroup must divide the order of a (finite) group—follows immediately from this result.

Because the kernel of a homomorphism is always a subgroup of the domain, the definition of coset can be specialized as follows: If K is the kernel of a group homomorphism $f: G \to G'$ and $a \in G$, the (left) coset of K containing a is given by $aK = \{ak \mid k \in K\}$. Students proved on the second midterm exam that $f(a) = f(b)$ if and only if $aK = bK$. In other words, a and b have the same image under the homomorphism if and only if they are in the same coset of the kernel of the homomorphism. As discussed below, this specialization of homomorphism was quite salient in the students' thinking.

Carla and the Left Coset

Carla's work with cosets followed immediately from her work with homomorphisms. Given the function f from U_8 to Z_4, given by $f(1) = 0, f(3) = 0, f(5) = 2$, and $f(7) = 2$, Carla had identified the kernel of the homomorphism as the set $\{1, 3\}$. I moved on to cosets.

62   Brad: Okay, Let's call that set K, and what I want to do is investigate these sets aK.

63   Carla: Okay. Left cosets. All right. So to do aK, we see what happens.... Okay. The set $\{1, 3\}$ is always going to be on the right, and we want to work with every single element that is in.... Let's see.... Every ... $\{1, 3\}$ is a subset, in this case it is a subset of U_8.

66   Carla: So you are going to take out every element of U_8, I think it is. The definition is.... I think you take every element of U_8, and you multiply, in this case you multiply because that's the operation. [Okay.] Is that right that it's U_8? I'm not sure.... We just said this today, but ... [inaudible]

67   Brad: Well show me what it is you are going to do here, and then maybe you can answer your own question.

68   Carla: Well I am going to multiply 1, 3, 5, and 7 each individually by the, by K. So 1×{1,

3} gives me {1, 3}. 3×{1, 3} gives me {1, 3}.

For Carla, the notation aK evoked the term *coset* and seems to have supported her reasoning: a is an element that varies through U_8 and K is a set. The individual elements of K required little of her attention, as evidenced by the fact that she did not say or write down intermediate calculations and quickly considered {3, 1} to be the same set as {1, 3}. The lower-level process of iterating through the elements of K was essentially automatic. She computed the other cosets similarly and quickly (see Figure 16), concluding "So we have two elements in the left coset" (line 70).

Figure 16. Carla's cosets of {1, 3}

$$\{1, 3\} = \ker(f) = K$$
$$aK$$
$$1 \times \{1, 3\} = \{1, 3\}$$
$$3 \times \{1, 3\} = \{1, 3\}$$
$$5 \times \{1, 3\} = \{5, 7\}$$
$$7 \times \{1, 3\} = \{5, 7\}$$

From the ease with which Carla performed these calculations, it seems that she had a firm understanding of the concept of coset, but her statement "we have two elements in the left coset" was nonstandard syntax, suggesting unusual thinking. I asked her to explain:

73   Brad: Two elements in the left coset. What do you mean?

74   Carla: Because {1, 3} and {5, 7} are each of, are the two elements that result from aK.

75   Brad: Are the two.... They are elements ...

76   Carla: Well, they are sets, but if you look at them as one.... If you rename {1, 3} to a and {5, 7} to b then there are two elements in aK.

77   Brad: Oh, okay. So you are saying, on the one hand, we can talk about the set of cosets. Does that make sense? [Yup.] So how many cosets are there?

78   Carla: Two.

79   Brad: Two. But now does aK refer to the set of cosets, does it refer to all of them, or does it refer to one specific one?

80   Carla: aK tells you how you find the cosets. So that was none of your choices. What were your choices?

81   Brad: Does aK refer to all of the cosets, or just one of them?

82   Carla: aK is the general formula that gives you all of the cosets. Then you choose specific as for whatever K you are given, and you will find all the specific cosets.

Thus, Carla had an efficient and reliable procedure for computing cosets, and the procedure was supported by the notation aK. Nevertheless, she did not maintain a clear distinction between a single value of a and the set of all of such values, and thus did not distinguish between the particular coset aK and the set of all cosets of K. Instead, a varied as part of a procedure specified by the formula aK, which gave all the cosets. Psychologically, it seems to be an easy step to then imagine that aK is all such cosets, without needing to distinguish between a nonspecific one and the collection of all of them.

Analysis. Carla was happy to let a vary through all the elements of G and to construct the collection of cosets aK that result. What she did not see, however, was a need to distinguish notationally between the particular coset aK and the collection of all such cosets. By calling the cosets *elements*, Carla appeared to have little difficulty seeing cosets as objects, suggesting that she had encapsulated the process of coset formation. Furthermore, her language "the left coset" for the collection of cosets suggests that she had further encapsulated that collection as an object, a set of sets, a point of view that is helpful in order to see the set of cosets as itself a group under the appropriate coset arithmetic.

Figure 17 shows schematically the objects and processes involved in coset computation and delineates the two levels of processes. Carla's language suggests that her thinking was in the transitional process between two objects: the particular coset $aK = \{ak \mid k \in K\}$ and the set of all cosets of K, which might be written $\{aK \mid a \in G\}$. In the midst

of the process, a is varying, so aK denotes neither a particular coset nor all of them. For Carla, aK denoted and specified the process. Thus, my question about the distinction (line 81) was not relevant to her.

Figure 17. Objects and processes in coset computation

A particular a in G
The subgroup K $\longrightarrow$ The coset aK $\longrightarrow$ The set of all cosets of K

Process:
By letting k vary
through K, calculate
all the products ak.
Collect them into a set.

Process:
By letting a vary
through G, calculate
all the cosets aK.
Collect them into a set.
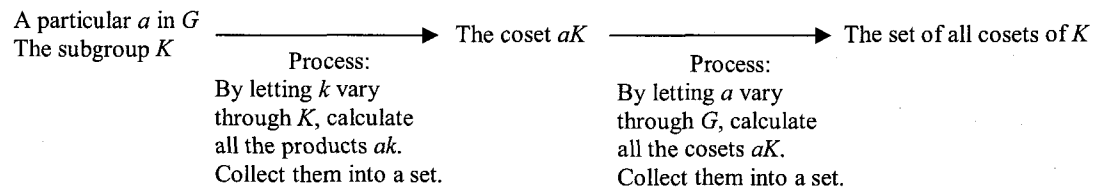
Being immersed in the process allowed Carla some flexibility in her thinking. On the one hand, she could consider a particular coset by stopping the process for a moment. On the other hand, she could consider all the cosets by completing or imagining she had completed the process. From within the process she could broaden her viewpoint slightly and see both a particular coset and the set of all of them as two aspects of the concept of "coset." Thus, "aK tells you how to find the cosets," and together there are "two elements in the left coset." Carla maintained this dual role of the term *coset* through the fourth interview, and even maintained her process orientation, saying, for example, "The left coset *gives* the set of (23) and (132)" (line 239, emphasis added) rather than the left coset *is* that set.

Diane, Lori, and Cosets

Although Carla's ambiguous use of the word *coset* was unusual and idiosyncratic, Diane and Lori used similar language. During their fourth interview, I asked them to find the cosets of a subgroup in D_3. They each drew a triangle and labeled the vertices 1, 2, and 3 to help themselves write down the elements of D_3. Based on the geometric interpretation,

they used the standard term *rotation* and the nonstandard term *flip* to distinguish (1), (123) and (132) from (23), (13), and (12).

They had initial difficulties with the coset formation process, first multiplying each element of D_3 by only (12) rather than the whole subgroup, {(1), (12)}. Diane explained, "The cosets are only going to be using the element (12) because (1) does not matter, because that's the identity" (line 42). To overcome these difficulties, I then asked them to find the cosets of the subgroup generated by (123), and the bulk of the interview dealt with this subgroup. For simplicity of notation in what follows, I call the subgroup H, although Diane and Lori always wrote it out in full as {(1), (123), (132)}, sometimes without the enclosing braces.

Diane explained the process for calculating cosets: "You'd have to take this [subgroup] and multiply it by each of the elements in D_3" (line 67). Lori, on the other hand, was unclear on the question and asked, "Coset of D_3?" (line 68). I told her I wanted a coset in D_3 of the subgroup that she had found. She was still unclear on the process:

74 Lori: All right, I have a quick question. Do you take every element in D_3 and multiply it by the singleton (1), then take every element in D_3 and multiply it by (123), and every element …?

As they continued with their calculations, Diane expected two get "2 different cosets … because the order of this [subgroup] is 3, the order of D_3 is 6, so 6 over 3 is 2" (lines 82-84). After computing (1)H and (123)H and seeing that they both resulted in H, Lori concluded, "The left coset is probably just going to be this alone" (line 114). But then Diane pointed out that they should try the calculation with something other than rotations.

Before they began those calculations, I asked Lori what she was calling the coset.

120   Lori: Well, I don't know what the *whole* coset is.

124   Lori: I still have to run through all of these. I don't know what the whole left coset is yet. Like what Diane is saying, if we do (132), it's just going to give us this again, so maybe we can skip to (23). We are going to get ...

125   Diane: If you are only looking at two elements of the coset, and this is one, then if we take any one of these we should come up with the same.

This is the first indication of unusual language. It became clear as the interview progressed that Lori was calling the collection of cosets "the whole coset." Because she had not completed the calculations, she did not yet know what the whole coset was. Diane had earlier predicted that there would be two cosets (line 82). Here she called them "elements of the coset," demonstrating language similar to Lori's. Diane was also predicting here that calculations with any of the flips would yield the same "element."

Diane computed $(23)H$, and Lori computed $(13)H$, and they got the same result, $\{(12), (13), (23)\}$, though Lori did not include braces. I asked Lori how she had computed the coset.

134   Lori: Um, I took the element, the cycle (13) in D_3, and I multiplied it by the set [the subgroup] to get this coset.

137   Lori: For the whole left coset, it's going to be this [H] and this [$(13)H$], and that's it.

Thus Lori was at this point clear on the process. Furthermore, she was calling $(13)H$ a coset, but H and $(13)H$ together were "the whole left coset."

When I asked them how they would write it, they listed the elements in both cosets, but Lori merely listed the six elements from D_3 without distinguishing between the two cosets. I asked them how that was different from D_3. Lori said, "It's not" (line 143), and Diane said, "It is D_3" (line 144). Then Lori provided a connection to the computations:

146   Lori: You do kind of have to distinguish that these are all one when you multiply it by these elements, and that these are all one, like Diane did. [She adds parentheses.]

147   Brad: Okay. So explain your notation here?

148   Diane: This is the coset, the left coset, and these are the elements in it.

They had both then written ((1), (123), (132)), ((23), (13), (12)). Lori explained:

151   Lori: I know how to say it. You are always going to get like this first parentheses when you multiply the first three rotations in D_3 times the cycle you wanted us, not (12), the cycles that you wanted us discussed by this one. So you are always going to get those first three. So that's like the first left coset, all in one. And the second one all in one is…. The last three elements in D_3 multiplied by that set.

153   Lori: So there is only two left cosets. Like…

155   Lori: Two left coset values. How do you, what do you call it?

156   Diane: Elements.

157   Lori: Yeah, two left coset elements.

158   Diane: There are two elements in the left coset.

During this passage, Lori was trying to adopt Diane's language, which was to call the set

of cosets "the left coset" and to call each of the cosets "elements." Much earlier, when

Diane had predicted that there would be two cosets, she had called them cosets (line 82).

But this may have been a slip of the tongue, for other than this single sentence, her

distinction between "the left coset" and "elements" was entirely consistent. Lori, on the

other hand, seemed to be comfortable with her ambiguous language, and she was also

willing to adopt Diane's language.

Exploring the language. At this point in the interview, I had a clear understanding of the

ways that Lori and Diane were talking about cosets. I chose to explore what it would

take to change their language:

172   Brad: What if I told you that … this that you have computed here is actually *a* left coset. That's one left coset right there.

173   Diane: Wait a minute, so this is *a* left coset, and this is *a* left coset, and together they make *the* left coset? So there is one …

Diane was uncomfortable with my suggestion because adopting it would have been

confusing alongside her language. Apparently, she had not noticed the ambiguity in

Lori's language, for she was describing Lori's usage. Lori, on the other hand, was

perfectly comfortable with this ambiguity.

> 181 Lori: Yeah, that was what I was going to say. There's six left cosets because every single element in D_3 is going to produce a different left coset. So there is actually six left cosets to this subgroup, but they are the same. Like what Diane and I did, all of the rotations are the same, and then all of the stationary ones, or whatever you want to call them, are the same. So then it only comes down to two. Then we call the whole thing one left coset. So it's only one set.

Lori was focusing on the fact that there were six coset calculations, but because some of

them were the same, the result was two cosets. Then the two cosets together, collected as

a set, were called one left coset. Diane continued to insist that H was an element, not a

coset: "This whole set here is one element" (line 187). Furthermore, she did not like

Lori's claim of six cosets. Because the sameness of the results of the calculations, there

really were only two sets.

Diane and Lori agreed that what they had written as ((1), (123), (132)) was in fact a set. I

asked whether they could use braces to call it a set, and they both changed their notation

to {{(1), (123), (132)}, {(23), (13), (12)}}. I asked them what that was. Diane said, "A

set of sets" (line 196), and Lori agreed. Furthermore, it was a set with two elements, each

of which were sets with three elements.

I asked Diane what she meant by the word *coset*. She responded, "This whole thing"

(line 214), pointing to the set of cosets. Lori agreed that she had been using the word

coset to describe the individual calculations. But then she backed off:

> 224 Lori: I generate six sets by multiplying D_3 by this set [H]. And then I see which is similar and which are not, and it gets put into one coset set. Maybe, yeah.
>
> 226 Lori: It seems like every one I have computed, there's always ones that are similar.
>
> 227 Brad: And then you find that you have really only ...
>
> 228 Lori: Have one big set with all of the different set elements that are not similar in it.

Thus, despite my attempt to impose the standard language, Diane's language continued to prevail, influencing even Lori.

I tried again to impose standard language.

229   Brad: What if I told you each of those things that you have computed is called *a* coset.

230   Lori: That's what I was thinking. Like, I computed a coset. There's another one. Just because it's the same doesn't mean that it's still not one. It is a coset; I just didn't write it. That's what I was trying to say.

Thus, Lori was comfortable with the standard language for an individual coset. But after completing the calculations and collecting the cosets, she explained, "There is only one whole set that's called *the* left coset" (line 234). Lori continued to be willing to use the term *coset* ambiguously.

Diane, on the other hand, did not agree:

238   Diane: Yeah, I think we are having trouble with vocabulary [Okay] because I am thinking that this is *the* left coset. It's complete. You know, that's why you call it the left coset. This isn't complete yet. And this isn't complete by itself; you have to put them together for them to be complete.

Diane was focusing on the completed process and was not concerned about individual cosets. She later agreed that it would be useful to have a name for the sets H and $(13)H$, and Lori suggested that the name should be coset. When asked whether it made sense to call both individual ones and all of them coset, Lori responded, "Sure" (line 243). Again Diane disagreed: "It's just a little confusing though because if you were to talk to somebody about it and explain it to them, … they are not going to know what you are talking about" (lines 244-246). Diane was not comfortable with the ambiguous language.

Thus far, my interventions had been unsuccessful. Next I tried to make a clear distinction:

247   Brad: Okay, well it seems to me that there are two possibilities. One is to call the big thing, that's everything together, the left coset, which is what you have been saying, right? And come up with a different name for the individual elements here of what you are calling the coset.

248   Diane: And then we are going to need a name for all the different things ... [inaudible]

249   Brad: Yeah, so we could come up with a different name for that. An alternative is to call this, each one of these things, a coset.

250   Diane: A coset. And come up with a name for this [the set of all of them].

Diane seemed comfortable with the distinction and with the two options I was putting

forward. I continued:

257   Brad: So it seems to me we just need to decide which of those two approaches to take. Does that kind of...? Am I putting words in your mouth?

258   Diane: No, no. I know.

259   Brad: I'm sort of boiling down what the disagreement is.

260   Lori: Oh, yeah, I definitely see that.

261   Diane: Yeah, I've been having problems with calling everything the same thing, and I want to separate it.

So it appears that Diane and Lori saw the distinction I was trying to make. Furthermore,

Diane made clear that she had been uncomfortable "calling everything the same thing." I

again tried to impose standard language.

262   Brad: Well, here is the place where I am going to just tell you. These things here, the smaller things are the cosets. Each of those is a coset. So here we have a bigger thing which contains what?

263   Lori: All of the left coset.

264   Brad: Sss.

265   Diane: Cosets.

269   Lori: Containing the left cosets.

Thus, after emphasizing the crucial distinction between a coset and the set of all of them,

I gave them the standard language, and they seemed comfortable with it. It remained to

be seen, however, what influence that intervention would have on the language they used.

We next began talking about right cosets.

279   Diane: Those are, a right coset. And you have the set of right cosets.

280   Lori: So this is a normal subgroup in these are its cosets. You don't have to distinguish left or right because we are going to know they are the same.

It may be that Diane's first sentence indicated her previous nonstandard language. Her second sentence, however, is syntactically correct, as is Lori's statement.

Nonstandard language continued to surface. In predicting the number of right cosets, for example, Diane used the formula suggested by Lagrange's theorem:

296   Diane: If you get a group that's six elements, divided by a group of three elements, you should get two groups of three elements each.

297   Brad: Two?

298   Diane: Subgroups of three elements each.

So Diane called the cosets first groups and then subgroups. After some discussion, Diane pointed out that one of the two cosets was not a subgroup: "It can't be; there's not identity" (line 318). Lori then concluded, "You don't get two subgroups" (line 322). So Diane changed her language: "You get two sub-, cosets" (line 323). Thus, despite the false start, Diane was able to find the standard language. Lori, on the other hand, sometimes still used ambiguous language:

349   Lori: And there's going to be two cosets in our whole set. We're just calling this a set [inaudible] or *the* left coset.

Conclusion. This case demonstrates that Carla's unusual language was not so unusual. Lori's dual usage of the term *coset* was quite similar to Carla's. Although Lori was not as clear about the process as Carla, neither of them made a clear conceptual distinction between a particular coset and the set of all of them. Diane's language, on the other hand, indicated that she was making the standard conceptual distinction but had attached names in nonstandard ways.

Continuing the comparison, Diane's and Lori's work with cosets was not guided by the symbolism aK, as it had been for Carla, but rather involved working with particular groups. It is not clear what meaning Diane and Lori would have attributed to the symbolism, but symbolic issues were prominent and problematic in other interviews, as is illustrated below in the subsection entitled "Robert and What Varies."

This case also illustrates some of the cognitive difficulties in making new distinctions and renaming concepts. My initial attempts at imposing standard language were unsuccessful. Diane's nonstandard language continued to prevail, and Lori's language retained its ambiguity. The third attempt was partially successful, it seems, because I first clearly set out the conceptual distinction I wanted Diane and Lori to make. Yet even then, Diane and Lori sometimes slipped into nonstandard usage. The theme of making new distinctions arises frequently in this chapter.

Other Students and Cosets

In characterizing the concept images of coset for Carla, Diane, and Lori, I found the central issue to clearly be one of language. By broadening the analysis to include other key participants, a conceptual issue rises to prominence: Many students preferred looking for a homomorphism and its kernel before they were willing to compute cosets, suggesting that the concepts of coset and kernel were closely related in the students' thinking. Carla, for example, initially thought that to compute cosets she would need the kernel of a homomorphism. Wendy displayed similar expectations, as did Robert in both his third and fourth interviews. This connection is not surprising because, as mentioned above, cosets require a subgroup, and kernels are always subgroups. Nonetheless, the

concept images of cosets and kernels were so strongly connected that the students were sometimes obstructed in their progress on the interview tasks.

In Robert's third interview, I asked him whether he could find cosets of the subgroup generated by 3 in Z_{12}. He calculated the subgroup as $\{0, 3, 6, 9\}$, called it *theta*, and then looked for cosets.

278   Robert: All right, so I've got to find the kernel though of Z_{12} in order to get cosets. H is the kernel. Over here it's anything that maps to 0 would be in the kernel ... anything in here that maps to 0 in Z_{12}.

Thus, Robert thought he needed to find a kernel. He soon realized that in order to talk about the kernel he needed a homomorphism, which he said would "Take Z_{12} to theta" (line 284). A good portion of the interview was consumed constructing a mapping and considering whether its kernel was a subgroup and even whether the mapping was a function in his sense of the term.

Robert's association between coset and kernel presented a significant stumbling block that was preventing me from learning about his concept of coset. Eventually, I intervened:

370   Brad: But now let's just say we could design a function that would be a homomorphism and had the kernel as this thing that I have said here. The kernel would be 3, 6, 9, and 0. Let's say we could set up a homomorphism from.... You know, I am not even going to tell you what group we are going to. We are going to send it from Z_{12} to some other group—I am not telling you what—but the kernel of that's going to be this set $\{3, 6, 9, 0\}$.

371   Robert: And then you want me to find cosets.

My suggestion was sufficient, for Robert then calculated the cosets quickly.

Furthermore, by noticing patterns and making and revising conjectures during the calculations, he decided, without doing all of the calculations, that there would be only three cosets.

Robert also showed a connection between coset and kernel during his fourth interview. I had asked him to find cosets of a subgroup in D_3. He calculated the subgroup and then began to consider cosets.

114   Robert: All right. So I take any.... I got to know which things.... See, I am still not sure about the coset thing. You.... When you talk about h being in H, you are talking about things in the kernel, right?

115   Brad: Well, in order to be talking about the kernel what do you need?

116   Robert: A homomorphism.

117   Brad: Do we have a homomorphism here?

118   Robert: No. I think that this means I can just take.... I can take all of these elements ...

120   Robert: ... of D_3, and I can operate them on the left with the elements in my set here, and that'll give me the right cosets. And I can operate them on the right with the elements in the set and have the left cosets.

Thus, during this interview, with the support of improved proficiency with the concepts and processes, Robert was able more quickly to overcome his association between cosets and kernels.

In both interviews, Robert demonstrated an insufficiently general concept of coset, suggesting that the students' experiences with cosets of kernels of homomorphisms overpowered the more general tasks. For Robert, overcoming such a constrained concept image of homomorphism seemed to depend on Robert's proficiency with the processes and related concepts. It seems that when a students' concept of coset depends upon having a kernel, it is not easy to generalize to cosets of arbitrary subgroups.

Summary

The foregoing analysis characterizing the students' concept images of cosets supports two results. First, the students used nonstandard language, which indicated nonstandard concepts that were dominated by the process of creating the cosets and which failed to distinguish between an individual coset and the set of all of them. Furthermore, the

students were resistant to attempts to impose standard language. Second, the students' concept images were insufficiently general and were dependent on their proficiency with the concepts, processes, and examples. Other results about the students' understandings of cosets are presented below. Results concerning the ways the students used notation are presented in the following section. The analysis has already described ways in which the students considered cosets to be objects, but a deeper discussion of the process/object distinction in the students' concept images of cosets must take into account not only these notational issues but also the ways the students used cosets in constructing quotient groups.

Notational Issues

There was considerable variation among the key participants in the ways that the notation supported and inhibited their concept of coset and the related processes. Carla's intermediate position in the process between an individual coset and the set of all of them was strongly supported by the notation aK. Diane and Lori, in contrast, did not use such symbolism during their interview but seemed similarly process oriented. In this section, I present three kinds of notational difficulties: confusion between set and element, trouble managing processes, and losing track of the objects. Each of these is illustrated by an episode related to the concept of cosets, but the issues are more general, as is suggested by these descriptors.

Carla's failure to distinguish between a particular coset and the set of all of them might be characterized more generally as mixing up statements about sets with statements about elements (see, e.g., Selden & Selden, 1978). The characterization that she was immersed

in the process is more compelling, however, because there is no indication that Carla was confused. Furthermore, this notion of being stuck in the process can also explain a statement Carla had made earlier in the interview, but which I had not pursued: "The kernel is $4n$ where n is an integer" (line 23). As mentioned above, whether Carla meant N or n, she was not appropriately distinguishing between a set and one of its elements. A more compelling explanation, however, is that Carla saw $4n$ as the process for generating the kernel, which might be described as an intermediate position similar to that for her concept of cosets. The subsection entitled "Wendy and $N + N$" explores almost identical language and notation during another interview.

The second notational difficulty is related to the set/element distinction but has to do with managing the process of coset formation, thereby providing insight into the relationship between symbolism and processes. During her third interview, Wendy described her coset calculations as follows:

253 Wendy: It's going to be all the elements times.... Like 1.... All the elements, we are going to call them h. h are all the elements in U_8. And k are going to be elements in the kernel. So h times $\{1, 3\}$ is how you are going to find it. So it's going to be 1 times $\{1, 3\}$, and it's going to be 3, 5, 7 times $\{1, 3\}$.

To accompany her statement, she wrote $h \in U_8$ and $k \in K$. Neither her words nor these symbols distinguished the different roles played by h and k. Nonetheless, she maintained the appropriate conceptual distinction, for she also wrote $h\{1, 3\}$ and performed all the calculations correctly. Robert, on the other hand, demonstrated a similar use of notation but had considerable difficulty establishing a sufficient conceptual distinction between the set and the element to manage the processes of coset formation. This is illustrated below in the subsection entitled "Robert and What Varies."

Finally, Carla herself showed a marked contrast between her work with particular groups, such as U_8 and Z_{12} and her symbolic reasoning about cosets generally, where the group remained unspecified. In particular, she seemed to forget that aK and bK were cosets and instead treated the letters like numeric variables from high school algebra. This point is demonstrated in the subsection entitled "Carla and $aK = bK$."

Wendy and $N + N$

This subsection explores a failure to distinguish between a set and an element and provides insight into the kind of learning that occurs as students begin to make such a distinction. During her fourth interview, Wendy described a function from Z to Z_4 where x goes to x (mod 4)—the same function that Carla had described at the beginning of her third interview. Then Wendy listed the cosets of the kernel of the mapping:

50   Wendy: Well, we can take the generator group now, and we can find cosets a lot more easy because if you take ... $4x$ ... every element ... every multiple of 4 is going to get mapped to 0 in x mod 4 ... in Z_4 ... so $4x$ is going to equal the kernel. So the cosets are going to be $4x$. Okay. $4x + 1$, $4x + 2$, and $4x + 3$, because $4x + 4$ is just going to be [the same as $4x$].

52   Brad: Okay, and what's x?

53   Wendy: x is going to be an integer.

54   Brad: Okay, now is x...? In these 4 cosets, is x a specific integer?

55   Wendy: No.

56   Brad: What do you mean?

57   Wendy: It can be any.... Like any integer you put in here will give you.... Any integer you put in for x will give you 0 ... will give you 0 mod 4.

58   Brad: So like you could put in 2 for x.

59   Wendy: Yeah, you could put in anything for any of them, and this [$4x$] is going to equal 0. This [$4x + 1$] is going to equal 1, this [$4x + 2$] is going to equal 2, and this [$4x + 1$] is going to equal 3.

Like Carla, who had used $4n$, Wendy was using $4x$ to denote both a particular multiple of 4 and the set of all of them. My language "a specific integer" did not help her make the distinction I was trying to make. Of course, x was not a *specific* integer. The question is

whether she was imagining that it was a *particular* integer or the set of all of them. Yet,

perhaps even this distinction would not have helped, for she was focusing on the images

of these integers in the codomain, Z_4. And for that purpose, $4x + 1$ was going to map to 1

whether it was a particular value or any set of values.

68   Brad: Okay. Now.... But now you are talking about x as being ... you can put anything in there for x. Any integer, right? But now is the coset then any integer? Or is it all of them. Or is it one specific one? Or...? Do you understand what I am asking?

69   Wendy: No, not really. Is the coset one specific.... Like should we name this something? [She writes N.]

71   Wendy: Yeah. If I call it N, it's always, it's going to be congruent to ... any integer N that maps to the identity.

73   Wendy: So, it's not going to be a specific, it's going to be any multiple of 4.

75   Wendy: So it is not specific. You know what I mean? Like, we are going to call this coset.... If we call this coset N, N is infinite.

76   Brad: Oh. Is it a set then? Or is it a specific number?

77   Wendy: It's a set.

The letter N denoted what she had previously called $4x$ and was "any integer that maps to

the identity," which was not specific but "any multiple of 4." Simultaneously, N was a

set. Thus, Wendy was not distinguishing between any multiple of 4 and the set of all of

them.

When I mentioned again that the set was infinite, she said, "That is why I was calling it

$4x$" (line 81). Because she could not list all the elements, I asked her to show the pattern

of N. She began with positive multiples of 4 and then included 0. She included negative

multiples of 4 only after I asked explicitly whether there would be any negatives in the

set.

88   Brad: So then, when you write this thing $4x$, you mean this set of all the things together, taken as a whole. Or do you mean individual specific ...?

89   Wendy: Uh huh. Taken as a whole. Take as a whole.

So just like N, $4x$ was simultaneously any multiple of 4 and the set of all of them. From the fact that Wendy used both N and $4x$, a glance at her work might have suggested that she was using the standard convention of using capital letters for sets and lowercase letters for individual elements. Wendy was making no such distinction, however. The two notations were identical in meaning.

I asked what could be done with the cosets. She reaffirmed that "$N = 4x$, where x is all integers" (line 98) and renamed the other cosets to $N + 1$, $N + 2$, and $N + 3$. She struggled for a moment over whether to add or multiply the cosets but then decided to add them because "integers are only a group under addition" (line 112). Then she tried to calculate $N + N$ and $N + (N + 1)$.

115   Wendy: All right. So, you are going to have, if you have $4N$.... $4x$ plus, because $N = 4x$. It's going to equal $8x$. All right? And $8x$ is congruent to $4x$, which is congruent to 0 mod 4. So this is congruent to N.

116   Wendy: So N is.... You can tell here that N is the identity element. So, you know that ... [inaudible] when you add N plus, and $N + 1$ you are going to get $2N + 1$ and that's going to give you.... That's the same thing, that is congruent to $N + 1$. Now I have to figure out why. [Writes $2N + 1 \equiv N + 1$]

It seems that Wendy knew that $N + N = N$ and that $N + (N + 1) = N + 1$. Her reasoning was flawed, however, relying on algebraic procedures that work for symbols that stand for numbers but not symbols that stand for sets. Her symbol manipulation was guided more by her expectations about the results than about the meaning of the symbols.

I asked her what $N + N$ meant.

120   Wendy: You are adding the same set together, so it is going to be the same set. You know like.... If you add two of the same sets together you are just going to get all elements ... the same elements in the set. Like if you add 1, 2, 3 ... the set of $\{1, 2, 3\}$ and the set of $\{1, 2, 3\}$ you are still going to have the set $\{1, 2, 3\}$. Your elements aren't going to change any?

Wendy saw $N + N$ as a sum of sets. When I asked her how to compute $\{1, 2, 3\} +$

$\{1, 2, 3\}$, she did not remember at first how to do it but eventually decided that the sum

would be $\{2, 3, 4, 5, 6\}$. She was not happy with the result, however, because it did not

fit her expectations.

134   Wendy: But if you add any multiple of 4 and any multiple of 4, you're going to get another multiple of 4.

136   Wendy: So in this case it's different. So I know that it looks like. I was like, oh well that disproves what we are saying here, but it doesn't because ...

137   Brad: So then is it right to say $4x + 4x = 8x$? Is it right to even call it this thing $2N$?

138   Wendy: Well, $4x$.... You can say $4x + 4x$. I don't know about this [Crosses out $2N + 1 \equiv N + 1$]. But if you say $4x + 1 + 4x$, you are going to get $8x + 1$, which is going to be congruent to.... This is still a multiple of 4, so it is still going to be equal to—congruent to; I don't want to say equal to—a multiple of 4, plus 1, which is what $N + 1$ is.

Wendy clearly had some thinking that was not reflected in the symbols. Furthermore, she

had not understood what I was implying by my question about whether $4x + 4x = 8x$ was

an appropriate calculation to verify that $N + N = N$. I suggested that $8x$ looked more like

multiples of 8, not multiples of 4.

140   Wendy: That's true, it doesn't include every multiple of 4. But we're talking about mod 4. And if you are talking about mod 4, we got this because it is $4x + 4x$. All right? So if you have a multiple of 4 here and a multiple of 4 here, and then you add one, it is still going to be.... If you add two multiples.... Like, if you add.... 8's a multiple of 4, and 8's a multiple of 4, and that equals 16. 16 is still a multiple of 4.

141   Brad: Do the two multiples of 4 have to be the same in this way you are writing $4x + 4x$?

142   Wendy: No. $8 + 12$, okay? That equals 20 and that is still a multiple of 4, so I don't think so.

143   Brad: Okay, so.... But now does the way you have written it, $4x + 4x$, does that handle both of these cases? When they are the same and when they are different?

144   Wendy: Uh huh.

Thus, Wendy's work with the symbols depended upon her thinking. Furthermore, she

did not see a need to distinguish notationally between the two multiples of 4. I pursued

this directly in her notation.

145   Brad: So by $4x$ you just mean ...

146   Wendy: A multiple of 4.

147   Brad: *A* multiple of 4. And by this $4x$ you mean *a* multiple of 4.

148   Wendy: $4y$.

150   Wendy: This is going to be $4xy$. So it's still going to be a multiple of 4 and then you add 1.

151   Brad: $4xy$ do you mean? Or ...

152   Wendy: Plus y. So, all right. This explains it better. Do you see why this explains it better?

It did not take much intervention to get Wendy to distinguish between the two multiples

of 4, but even after she had made the distinction, her symbolic moves were problematic,

suggesting that there were not yet strong connections between her thinking and the

symbols. From this point on, however, her symbolic reasoning improved dramatically:

154   Wendy: If you have a multiple of 4 and a multiple of, another multiple of 4, but where they're not the same multiple of 4, and then.... But this multiple.... This is.... Okay, this is N. Like if this is your multiple of 4, you add a multiple of 4, which we are calling N, and you are adding it to $N + 1$, you have another multiple of 4, $+ 1$. Okay? But because it is not necessarily the same ... like this isn't necessarily 8, and this isn't necessarily 9. They are not necessarily consecutive numbers. You have to have different values for x and y. Okay? So, we can, because they have a common factor, we can pull it out. [Writes $4(x + y) + 1$.]

156   Wendy: Okay, which is the same thing.... So this is still going to be 4 times.... This is still going to be an integer. An integer plus an integer is going to be an integer. So I am going have to call $x + y = z$ so $4z + 1$, so this is still going to be $N + 1$. That's how I can explain this.

Wendy's reasoning and symbolic representation seemed sound at this point. I then

attempted to learn how she had previously been thinking about the symbols.

158   Brad: Okay. But now here, when you are saying this $4x + 4y$, are you imagining that this is one specific x, for now, and this is one specific y, for now?

161   Wendy: Yeah, but it would work for any x and y.

162   Brad: Okay. But are you imagining for a minute that they are fixed?

165   Wendy: It helps me think of it better, but it doesn't necessarily have to be. Because no matter what x or y you put ... any integer you put in there, it will work. So, and these are all the integers. x and y are all the integers for ... are just, are all integers. So in that case and since it works for all integers, you can look at it as the whole set. But yes, you were right, I was looking [inaudible].

167   Wendy: ... But it doesn't make a difference, because.... It did help me clarify at first, but because if you show for a specific case it does work, and then if you take a step back and say it does, it works for every case ... like x, no matter what value for integer x, or y, it will still work.

Thus, Wendy agreed that it helped to think of x and y as fixed, but any integer will work, "and since it works for all integers, you can look at it as the whole set" (line 165). Her thinking could move smoothly from a fixed value, to any value, to all values, and finally to the set of all values without ever a need for a clear distinction. Nonetheless, at this point, she had begun to articulate a distinction in her thinking: first showing that it works for a specific case and then stepping back to see that it works in general.

I asked her to reflect on possible differences in meaning between $4x$ and N, and she asserted that "this set and this set are going to be the same" (line 171).

173   Wendy: I was calling x and y different, but they are not. Because x is going to be all integers. In this set, y are all the integers. Like this is $4z$, and this is $4z + 1$. This is the set $4z + 1$ ... all integers ...

174   Brad: By z there you mean ...

175   Wendy: All integers in z.

Thus, the distinction in Wendy's thinking was not yet reflected in her interpretation of the notation. Wendy's written work through this point included only a single lower case z, but her statement "All integers in z" seems to suggest that she was thinking about the set of integers.

176   Brad: Oh, do you mean the big Z that means all integers?

177   Wendy: By calling it x, I think it is like making me think towards ... by saying x ... like we usually use that for a specific value. But.... So if you write Z ... if you write $4Z + 4Z + 1$ ... $4Z$ is.... Go ahead. [She writes capital Zs.]

178   Brad: This Z ... is that the same as this z?

179   Wendy: Uh huh.

181   Wendy: The set Z.

Thus, just as Wendy was not distinguishing between a single value of x and the set of all

of them, she had not been distinguishing between a single value of z and the set of all of

them. After I asked whether she meant the set of integers, however, she changed her

notation from what appears to be a lowercase to the capital that is typically used to denote

the set of integers (i.e., $\mathbb{Z}$).

183   Wendy: All right, so if we think of that away from x, a specific case.... I think we, when we have x, I tend to think of it ... and I think that is what you are trying to get to up there ...

185   Wendy: ... So if we call it $4Z + 4Z + 1$, $4Z + 4Z$ is going to still equal $4Z$, $+ 1$. See I think that makes it more clear.

186   Brad: Although, now how do you explain to someone why $4Z + 4Z$ is just $4Z$?

187   Wendy: Because it's.... Because of the fact that it's infinite, you're taking every ele-, every multiple of 4.... Well, you can look at that specific example. If you take an x in, and a y, a $4x$ from $4Z$ and you add it to $4y + 1$, well x and y weren't equal. 4.... You can show what I was showing up there that x and y, because they are integers, because you took them from the set of integers, x and y will be an integer, so 4.... This is still going to be an integer.

It seems Wendy was still somewhat uncomfortable about x being a particular value.

When I asked her to explain the sum $4Z + 4Z$, she started talking about the whole set but

then resorted to using x and y to illustrate a particular case. I asked whether she was

beginning to see a distinction between $4x$ and $4Z$.

190   Wendy: Uh huh. Yeah. I am a lot more so than I was when I was up here.

191   Brad: Is that helpful to make that distinction?

192   Wendy: Yeah. I don't like calling it x now because it does, here looks like a more specific value, except it is not. But I think in a sense I was thinking of it, even though I didn't think I was.

193   Brad: You were thinking of it in which way before?

194   Wendy: As a more specific value. Although I knew that I should keep in mind that it was a set, but I think I was still thinking of it too specifically. I think I was right to show ... to explain it using a specific example, but I think it was important to go back, to begin and to end, showing that it was the whole set.

195   Brad: So before when you were talking about this $4x + 1$, were you trying to imagine both ways at the same time?

196   Wendy: Yes. I was definitely trying to.... I definitely knew that it works for all values of

x and *y*, but I wasn't really thinking of it as a whole, as like a *Z*, like integers.

Wendy was seeing a conflict in her earlier work: thinking of *x* as a specific value while also keeping in mind that it was the whole set. She stated that she had been thinking "too specifically" and had not really been thinking of *x* as the "whole set." Her language again suggests an intermediate position between a particular element and the set of all of them, because "it works for all values of *x* and *y*."

She began to focus on the notation:

198   Wendy: For some reason just visually, if you look at this 4*x* + 2 [points on paper] and 4*Z* + 2, visually, if you look at the two, it's easier to see one being a set and.... It's easier to see 4*Z* + 2 being a set over 4*x* + 2.

200   Wendy: 4*x* + 2 looks more like a value than it does a set.

202   Wendy: But I think it is clearer to write, you know, like *Z* as, like an upper case *Z*, like writing the notation as a set for the integers. So 4*Z* + 2 is going to equal, obviously be a set, and it is not going to be a value.

203   Brad: So it is useful, then, you think now, to distinguish between those things that are sets and those things which are sort of generic values.

204   Wendy: Like I use *N*, an uppercase *N* here. I think it makes more sense to use like the.... I think.... Do we use uppercase values for sets, rather than lower case?

So by the end of the interview, Wendy saw that the standard convention of using uppercase letters for sets could be useful in making a conceptual distinction. It seems she had much earlier developed a sense that uppercase letters were usually sets, though perhaps she had never before been in a situation where she felt a need to make a clear distinction between a set and an element.

Wendy seems to have learned to distinguish between a set and an element. Because this interview took place after the final exam, I can make no claims about the stability or durability of this learning. Instead, I would like to point out what seems to have been required. First, Wendy did not make any distinction between a particular value and the

set of all values until she saw that her equation $4x + 4x = 8x$ did not support the idea that the sum of any two multiples of four would be a multiple of four. Yet even after she began to make the distinction verbally, she still did not make it notationally. And even after she made the distinction notationally, she had to revisit her intermediate conceptual position of something that "works for all values of x." Then, by reflecting on the notation, she was able to see how the notational distinction could support her emerging conceptual distinction. Furthermore, it seems that I provided cognitive support by asking Wendy whether she was imagining x as fixed, suggesting that this metaphor was not available to Wendy at the beginning of the interview.

Robert and What Varies

Of the key participants, Robert had the most trouble negotiating the processes involved in computing the collection of cosets and, in particular, in keeping track of what kinds of entities he was dealing with and where they were situated. During his third interview, he was working with the same homomorphism as Carla above and had just determined its kernel. I asked him to find the cosets of the kernel.

140   Robert: Oh boy. Cosets ... equals the set of all ahs such that h is in H. [Writes $ah = \{ah \mid h \in H\}$.]

141   Brad: And what's H here?

142   Robert: Yeah, that's what I was wondering. Good question. I am not so good with these cosets yet.

143   Brad: Well, is the kernel a subgroup here?

144   Robert: Yeah, it's a subgroup of U_8.

146   Robert: Isn't that one of the things we proved on the take-home exam? That if f is a group homomorphism then the kernel of f is a subgroup of G.

147   Brad: Okay. So what would be the cosets of the kernel?

148   Robert: See, I am not really sure, like you say, what this H is. These are.... Okay. H would be 1 and 3 because those are the things that are in the kernel. So it would be the set $\{1, 3\}$, and each little h would be 1, 3. So then the question is the a. Which side of

the equation does this come from?

149   Brad: What do you mean "which side?"

150   Robert: Where do these as live? Do they live in U_8 or do they live in Z_4?

Robert quickly set down a symbolic definition of coset but did not know what to do with

it because he was uncertain as to what H and a represented in the problem at hand. This

first uncertainty was quickly resolved, but determining what a was proved to be more

problematic. The homomorphism made Robert uncertain about whether a was in the

domain or the codomain. Furthermore, his uncertainty was deeper than where a was

located; it also involved what kind of entity a was.

153   Robert: Well, essentially what I would do is I would add 1 to all of the elements in a. I
       want to think that like a is like the generator of Z_4 or something.

154   Brad: What do you mean?

155   Robert: I don't know exactly. Could a just be all of Z_4? If a was all of Z_4.

156   Brad: What do you mean by "all of Z_4?"

157   Robert: 0, 1, 2, or 3. The set.

166   Brad: And is a this whole set, or just one of them at a time?

167   Robert: Well it's just one of them at a time but…. Oh, no, no, it's not. It's the whole set.
       I think it's the whole set.

In the first sentence above, Robert talked about "all the elements in a," suggesting that a

was a set, but in the second sentence, a denoted a "generator" of Z_4. The latter statement

suggests that Robert may have been considering some intermediate role for a, where it

was neither an individual element nor the whole set. But while he considered whether a

could be all of Z_4, I may have pushed him toward the set interpretation by my question,

"Is a this whole set, or just one of them at a time?" which implicitly excluded an

intermediate role.

I then tried to encourage him to take advantage of the common notational distinction

between a set and its elements, using an uppercase letter for the former and the

corresponding lowercase letter for the latter. I pointed out that he was not being

consistent:

168   Brad: Here you have distinguished between big H, which is the set $\{1, 3\}$, and little h, which equals 1, I guess, or 3.

169   Robert: Uh huh. But there is no big A down here. So, a is the whole set. So I would have two cosets.

170   Brad: Now, here you have written a little h.

171   Robert: Yeah, should this be a capital H up here? [Fixes previously written equation to read $aH = \{ah \mid h \in H\}$.]

172   Brad: I am asking you the question.

173   Robert: I am sure that's the way that they define it. Okay. So that would make sense. So, I would have two of them. This one would be 1, 0, 3 ...

Thus, on the one hand, Robert adopted a notational distinction between uppercase and

lowercase letters, but on the other hand, he continued to see a as the whole set.

175   Robert: I did it wrong anyway. I am adding the 1 to each element. One of the little hs from H to each element in this set. And I get 1, 2, 3, 0. Same set. Not really anything different. And h, if I am doing the addition mod 4.... I get the same thing. [He writes $a = \{0, 1, 2, 3\}$, $ah_1 = \{1, 2, 3, 0\}$, $ah_2 = \{3, 0, 1, 2\}$.]

In making these computations, Robert took $h_1 = 1$ and $h_2 = 3$. He was concerned that the

two calculations resulted in the same set, but I chose to focus on the fact that he was

adding elements from different groups:

178   Brad: Okay, let me ask you.... We are looking at these things that you have written ah, right? Where do the things a live?

179   Robert: I believe they live in Z_4, but I am not sure. That's one thing I'm really not sure about.

180   Brad: And where do the things h live?

181   Robert: h live in U_8. The h live in U_8?

182   Brad: Is that a problem? You are talking about multiplying a thing in Z_4 by a thing in U_8, or operating on, somehow.

184   Robert: Yeah that's a problem, isn't it? Well, so maybe a lives in U_8 too. I mean, this is, H is the kernel of U_8, which is 1 and 3. I am pretty sure of that. So a could be the things living in U_8, and if they were, then we would be multiplying 1 times the set of $\{1, 3, 5, 7\}$. And 3 times $\{1, 3, 5, 7\}$.

So Robert decided that the calculations should take place in U_8, but a remained a set in his thinking. Both of these calculations again resulted in the same set, this time U_8. I asked him to explain the notation $aH = \{ah \mid h \in H\}$.

194 Robert: Well, they are saying that ... this is like the name of the set. And this tells you what all the sets ... the set is composed of, ahs such that h is in H. So for all of the as, you operate them with every h in H, and you form a set.

197 Brad: Now is that what you have done?

198 Robert: Well, my big question, I guess, is what is a? Yeah, I believe that's what I have done. 1, 3. Yeah here's the hs: 1 and 3 were the only two hs we had. And I guess is a.... I don't know why I think that a is a set.

205 Brad: So would you write then aH, a big H, is equal to well ... [inaudible] how you write it.

206 Robert: [Laughs.] a big H is equal to.... This is what, I mean, without doing the calculations, that's what it kind of looks like. But it doesn't seem to make sense to have all of these brackets. [Writes $aH = \{ 1\{1, 3, 5, 7\}, 3\{1, 3, 5, 7\} \}$.]

207 Brad: Okay. But in a sense it looks like aH is a set containing two sets.

208 Robert: Yes. Even though they are the same set.

There are two points to make here. First, if $a = \{1, 3, 5, 7\}$, then Robert's calculation was correct, as was his notation for it. Second, although he could write appropriate set notation, he did not like "all of these brackets," suggesting that the set notation did not support his thinking. At this point, I asked Robert to consider that a was not the set $\{1, 3, 5, 7\}$ but instead just one of its elements.

210 Robert: So you are saying that a would be each individual element. Then we'd have ah_1, ah, ... a_1h_1, a_2h_2 ... h_1, a_3h_1, a_4h_1; aH. You could pick 1. $1 \times 1 = 1$. $3 \times 1 = 3$. This is kind of a trivial one.

211 Brad: This is for h ...

212 Robert: Equaling 1, $h_1 = 1$.... This would be 7.

213 Brad: So how would you write it now...? Okay, so that's for a_1 little h_1, what would a big H look like? How did you do that?

214 Robert: It would just be the set of all of those numbers. [Writes $aH = \{1, 3, 5, 7\}$.]

Despite my suggestion that a be one of the elements rather than the set, Robert still let a vary through all of U_8 and fixed h to be one element in H. This is similar, perhaps, to the

idea he was considering at the beginning of this episode when he suggested that a might be a generator of Z_4. Both of these ideas—letting a vary through all its possible values and letting it be a "generator"—are, in a sense, intermediate between considering a fixed value and considering the *set* of all possible values, yet there is no conventional notation for this idea. Furthermore, because Robert's calculations explicitly included only one of the elements in H, it is not clear to what extent he was distinguishing between H and h_1.

At this point in the interview, I took a different tack, bringing in an idea that was more familiar:

224   Brad: What would $3H$ be?

225   Robert: $3H$ would be ... $3h_1$? Or $3h_2$?

226   Brad: 3 big H.

227   Robert: 3 big H would be 3, 1. [Writes $3H = \{3, 1\}$]

228   Brad: Now how did you do that?

229   Robert: Well, H consists of h_1 and h_2, which are 1 and 3, respectively, and I operated 3 with each one of those things, and I made a set. 3×1 is 3 and 3×3 is 9, which is 1 mod 8.

230   Brad: What would $5H$ be?

231   Robert: $5H$ would be 5, 5×1, and 15 which is 7, $5H$, 5 and 7. [Writes $5H = \{5, 7\}$]

232   Brad: Okay. Now, what you are saying to me then is that this 3 is fixed, but you are letting H do what?

233   Robert: Go through its possibilities.

Robert was able to do these more familiar calculations with ease, suggesting that his difficulties were largely notational and arising out of the fact that in aH there are two things that can vary. The notation aH represents a particular but unspecified calculation of which $3H$ is a specific example. Clearly, Robert had not made this connection. But after I provided supporting language about what was fixed and what was varying in $3H$, he was able to describe his previous calculations as fixing h_1 and letting the as vary. Then I took advantage of this distinction:

243 Brad: So I guess there is a choice here, that you can either fix the h_1 and let the as vary, or you can fix the as and let the hs vary.

244 Robert: Okay, I see what you are saying now. And if you were to ... if you let the as vary and fixed the hs you just end up with the same set you had, so that seems, like, inconsequential.

245 Brad: What if you do it the other way?

246 Robert: Then you get a bunch of different sets. 4 sets, I don't know what that would be ... 3 would be $1H$, $3H$ we have there; $5H$; $7H$ is the only one we don't have. 3 and 5, so I put the $5H$, and so we get 1, 2, 3, 4 different sets this way, but really only two different sets.

247 Brad: What do you mean?

248 Robert: Well, these two sets are the same, and so are these two sets. [Draws a line from $3H = \{3, 1\}$ to $1H = \{1, 3\}$ and a line from $5H = \{5, 7\}$ to $7H = \{7, 5\}$.]

With my distinction between fixing either the a or the h, Robert was able not only to

characterize what he had been doing but also to envision an alternative. At this point he

had completed the calculations and had noticed that the four calculations yielded two

cosets. To see whether Robert could see the appropriate procedure in the notation, I

asked him to reflect on the notation:

254 Brad: Built into this notation, aH is equal to the set ah such that h is in H.... Which do you think is implied by this notation? I said, we can fix the as and let the hs vary, or fix the hs and let the as vary, right?

255 Robert: It seems to be let the hs vary. I don't know. That notation, a times h such that h is in H. It seems as though you have to let the hs vary. But if a was an individual element in the original group, then you let both of them vary, essentially. See, but there's nothing there that tells you to let a vary, but there is certainly something to tell you to let h vary.

256 Brad: And yet when you first did this, which one where you letting vary?

257 Robert: I let the hs vary. Over here. But I considered a to be the whole set.

258 Brad: As opposed to ...

259 Robert: The big question for me now is, Is a the whole set, or is it the individual element in the set? Which, if a was the individual element in the set, it would mean that a was varying too. But you take each a and operate it with the varying hs versus taking the whole set of a and operating it with the varying hs.

Thus, despite having completed the correct calculations and being more satisfied with the result of two cosets than he had been previously with the calculations that yielded {1, 3, 5, 7} twice, Robert was still unsure about the interpretation of a in the definition of coset.

This episode demonstrates that notational distinctions do not necessarily create conceptual distinctions. From a mathematical point of view, the appropriate procedure for computing cosets follows directly from the symbolic definition. Yet the symbolism did not support Robert's reasoning and seems to have been the cause of some of his confusion. Nonetheless, in the interviews and on the final exam, Robert had no further procedural difficulties computing cosets, suggesting that some learning had occurred during this episode. It is unclear, however, to what extent he had connected his procedure with the notation.

In computing the set of all cosets of a subgroup, there are processes and objects at two levels, as described above. At the lower level, h varies, creating a particular aH, which is an object. At the higher level, a varies, creating the set of all cosets aH, another object. Robert did not see aH as being a description of $3H$ but instead focused on the fact that a was supposed to vary, thus merging or perhaps inverting the two levels. Robert did not make a clear conceptual distinction between what was fixed and what was varying. Neither did he make a notational distinction between uppercase and lowercase letters. Furthermore, from Robert's first response above (line 255), it seems that he had previously reflected little on the notation. Thus, it may be that Robert did not distinguish aH from ah or even AH, which would partially explain his merging of the two levels.

In order to encourage Robert to make the appropriate distinctions, I emphasized conceptual distinctions between sets and elements and between what was fixed and what

was varying, and I emphasized notational distinctions between uppercase and lowercase letters. It appears, once again, that both conceptual and notational distinctions need to be emphasized.

Carla and $aK = bK$

After Carla had completed the core tasks of the third interview, I chose to investigate her understanding of some of the results for which she had given correct proofs on the second midterm exam (see Appendix B). Because all of the other key participants took considerably longer than Carla to complete the other interview tasks, comparative data are not available. Here I provide a few observations that shed light on Carla's concept image of coset and her use of symbols.

Problem 5 on the midterm began "Let K be the kernel of the group homomorphism $f: G \to G'$ and suppose a and b are elements of G." The students were asked to prove several results, culminating with "$f(a) = f(b)$ if and only if $aK = bK$." In other words, a and b have the same image under the homomorphism precisely when they lie in the same coset of its kernel. I began quite generally:

133   Brad: What if, in this example here, we know that $aK = bK$. What can you conclude?

134   Carla: If $aK = bK$, we know $f(a) = f(b)$.

135   Brad: Okay. Now why do you know that?

136   Carla: 'Cause we did it on the take-home. The definition of aK is that aK, uppercase K, is the same as the set of a lowercase ks such that f.... Actually we wrote this two different ways. It's.... Okay, one way to write it is ak such that k is a member of K, which is actually the kernel. But another way we wrote it was that it is all xs in G such that $f(x) = f(a)$. So if we write bK as that.... If we write the definition of bK, we have bk such that k is a member of K. We'd also have x is a member of G such that $f(x) = f(b)$. If we said that $aK = bK$, then for the two sets to be equal, $f(a)$ has to equal $f(b)$ because they both equal $f(x)$.

Carla was almost correct. She first stated the major result from the exam and then presented, as definitions, two descriptions of aK, though she expressed concern a few

moments later that something might be missing from one of these definitions. The first

definition, $aK = \{ak \mid k \in K\}$, is correct and complete, essentially independent of where a

and K come from, requiring only that K be a set and that the product ak be defined for

each $k \in K$. The second characterization, $aK = \{x \in G \mid f(x) = f(a)\}$, on the other hand,

was not a definition but an intermediate result from the exam. Furthermore, her

explanation that "$f(a)$ has to equal $f(b)$ because they both equal $f(x)$" ignored the fact that

in the second description of aK, and likewise for bK, the variable x was not a particular

value but rather was to vary through all possible values for the purpose of finding those

that satisfied $f(x) = f(a)$.

In order to get a sense of the meaning Carla was associating with the symbols, I asked her

how she would write an element in aK.

142   Carla: Well it depends on what we are dealing with. Are we dealing with permutations or sets or what? Or just integers, or ...?

Carla had freed herself from the example we had been working with, which implies that

her earlier statement that $f(a) = f(b)$ required some specifications about f, a, b, and K. But

she wanted to attach her thinking to something.

143   Brad: Well, let's say we want to try think of it sort of in the abstract, independent of any particular representation. Okay? And let's say we have some element of aK.

144   Carla: All right. If it's an element of aK, it's also.... Well you know that k is in the group, because it's in the kernel. So if it's in the kernel it has to be in the group. You know that a is in the group because when we do left cosets we take elements from the group, and since the group is closed ak is going to represent an element of the group.

146   Carla: I don't know if that is answering your question.

So Carla took k to be in the kernel, just as she had assumed f to be a homomorphism,

suggesting that notational conventions supported her thinking in implicit ways that might

have been hard for her to articulate.

From this point on, the interview gradually took on a noticeably different character, becoming more directive and less exploratory. More of my questions required short, symbolic answers. During this time, Carla provided shorter answers and focused mostly on manipulating the symbols, similar to the work that she had done on the midterm exam. For example, she wrote both $z = ak$ and $z = bk$ (lines 148-150), not realizing that the ks might need to be different. At another point, she attempted to divide (line 179) rather than to multiply by an inverse. Her most surprising statement came after I asked her to verify some of her symbolic results for the particular case of the homomorphism and cosets from earlier in the interview. In response to some difficulty with the symbols, she said:

209   Carla: See, I don't like this business of multiplying integers times sets. I don't think that is very good.

210   Brad: Why?

211   Carla: Just 'cause you don't you usually do it. You usually do sets times sets.

Clearly at this point in the interview, Carla did not recognize that her earlier calculations of cosets were elements times sets (see Figure 16). All of these surprising statements, coupled with the fact that Carla did not use the word *coset* again in the interview (lines 147-246), suggest that she had was not conscious of the fact that aK and bK were intended to represent cosets like those she had computed earlier in the interview. Furthermore, she was making no substantive connection between the first and the second half of the interview or, equivalently, between her example-driven procedures for generating cosets and her symbolic calculations related to the proofs on the midterm exam. One possible explanation is a profound compartmentalization between two activities involving the symbol aK. On the one hand, there were activities where aK specified the process for generating cosets, and on the other hand, there were proof

activities that involved manipulating symbols such as *aK* according to certain rules. A less extreme explanation is that she would not have characterized one of the coset calculations in Figure 16 as an element times a set because she did *all* such calculations, which might be described as a set times a set. As for her discomfort with multiplying an element times a set, she softened her position the next day: "I don't know why I said that, because I was thinking about it later, and 4*Z* is an element times a set" (Field notes, May 3, 1996).

This episode further supports the position that that, for Carla, *aK* denoted a process if it denoted anything. Despite the fact that the process yielded cosets and a set of cosets that Carla considered to be objects, the notation *aK* did not denote such objects but was somehow separate from them. This result suggests that at least two encapsulations are required in learning about cosets: one for specific groups and another for reasoning about cosets generically.

Summary

The common theme in the above episodes is that of insufficient connections between notation and thinking. The notations 4*x*, *aH*, and *aK* did not always support a student's thinking even when that thinking was sound. Although Carla, Wendy, and Robert demonstrated different degrees of success with the tasks involving cosets and very different understandings, they were similarly imprecise in their use of the notation. Carla and Wendy, on the one hand, were clear on the concepts and the processes, yet the notation was not strongly connected to their thinking. Robert, on the other hand, was unclear on the concepts and processes and was not able to read the intended processes from the notation *aH*.

These episodes suggest an explanation for the problem of distinguishing between a set and an element. What seems to be missing in all of these episodes is the ability to think about generic objects. Reasoning with the notations as $4x$ and aK as generic representatives of a class of objects seems to require imagining, metaphorically, that x, a, and K are fixed. Then a kind of encapsulation is required to see $4x$ and aK as denoting particular but unspecified objects—a multiple of 4 and a coset. I say a different kind of encapsulation because there is every indication that the students could conceive of multiples of 4 and cosets as objects, suggesting that something like encapsulation had already occurred.

These episodes also suggest the hypothesis that proper usage of notational distinctions requires first the creation of a need for a conceptual distinction. The conceptual distinctions and notational distinctions are not automatic, but rather each requires learning. They are neither simultaneous nor consecutive but rather dialectic. Furthermore, there is good reason to believe that the metaphors "suppose x is fixed" and "now let x vary" can provide cognitive support for learning to make such distinctions.

Quotient Groups

As mentioned above, the set of cosets of a subgroup can reveal structural information about how the subgroup fits within the structure of the group as a whole. Thus, given a collection of cosets, a guiding question is whether those cosets can form a group—a quotient group. This section characterizes the students' concept images of quotient groups. It begins with a conceptual analysis that details the relationships among the concepts of subgroup, coset, homomorphism, normality, and quotient group, describing

important aspects of the students' curricular experiences regarding these concepts. The majority of the section consists of a detailed analysis of Carla's concept of quotient group, based largely on interviews in which she again demonstrated unusual language alongside well-established concepts and procedures. The analysis is then broadened to include other students.

Conceptual Analysis

A *quotient group* or a *factor group* is a group whose elements are cosets and whose operation is given by extending a group's operation to its cosets. Assuming the group's operation is called *multiplication*, the product of two cosets involves multiplying, in the appropriate order, all possible pairs of elements, one from each coset. Not all cosets can form a group in this way, but when the left cosets of a subgroup are the same as its right cosets, the subgroup is said to be *normal*, and the cosets will form a group under the set operation described.

For example, the set of left cosets of $\{1, 3\}$ in U_8 is $\{\{1, 3\}, \{5, 7\}\}$ (see Figure 16). Because multiplication in U_8 is commutative, the set of right cosets is the same, and thus $\{1, 3\}$ is a normal subgroup. The product of $\{1, 3\}$ and $\{5, 7\}$ is computed as follows:

$$\{1, 3\} \times \{5, 7\} = \{1 \times 5, 1 \times 7, 3 \times 5, 3 \times 7\} = \{5, 7, 7, 5\} = \{5, 7\},$$

where the products inside the braces are taken to occur in U_8. All the products of pairs of cosets may be organized in an operation table (see Figure 15), where it is possible to see that the cosets form a group with two elements.

Figure 18. An operation table for $\{\{1, 3\}, \{5, 7\}\}$ from U_8

$\times$	$\{1, 3\}$	$\{5, 7\}$
$\{1, 3\}$	$\{1, 3\}$	$\{5, 7\}$
$\{5, 7\}$	$\{5, 7\}$	$\{1, 3\}$

To analyze the interviews involving quotient groups, it will help to step back from particular objects and processes and consider them more generally. Most of the interviews may be described as instantiations of the general coset activity represented in Figure 19. How does one make sense of such an activity? What does one take from it? Figure 19 serves not only as a representation of an activity but also as an analytical tool constructed to investigate these questions. It supports semiotic analysis by distinguishing the processes from the names and the symbols from the objects in order to discern meaning. Thus, making productive sense of a coset activity requires distinguishing among the various components in Figure 19 and then assigning names appropriately.

The standard description of Figure 19 is as follows: Given a group G and a subgroup H, the left cosets of H are calculated and collected. The cosets are then considered elements in a new structure, a set of cosets. Coset products are computed, and results may be organized in an operation table. Carrying out theses two processes requires some conceptual flexibility: first conceiving of a set or an element as fixed and letting other sets or elements vary, and then letting the fixed set or element vary to carry out the higher processes. Managing such a process, after all, was precisely the problem Robert had in computing cosets, as described in the subsection entitled "Robert and What Varies."

If the set of (left) cosets constitutes a group under this operation of coset multiplication, then the set of cosets is called a *quotient group*. The construction depends upon both the group and the subgroup. Thus, the resulting group is called the quotient of H in G, denoted G/H and read "G modulo H."

Figure 19. Coset activity

Group

↓   Selecting

Subgroup

/ / | \\

           Calculating

↓   ↓   ↓   ...   ↓

Left Cosets

↓   Calculating

Products of Cosets

↓   Naming

Quotient Group
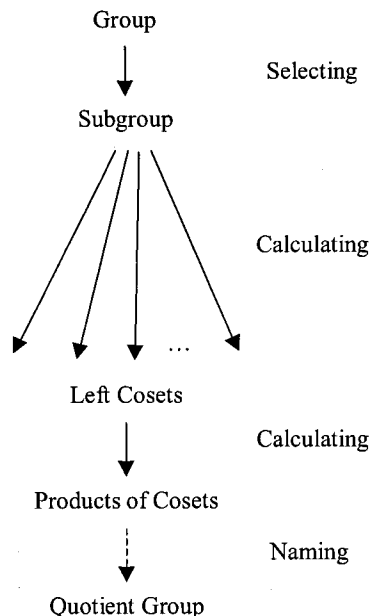
Of course, the set of (left) cosets does not always constitute a group, and the arrow is dotted in Figure 19 to indicate this potentiality. It turns out that the key issue is one of closure: whether the product of two cosets is again a coset. When this criterion is satisfied, all the other axioms follow: In G/H, the identity element is the coset H; the inverse of the coset aH is the coset $a^{-1}H$; and associativity follows from the associativity of the operation in G. It turns out that equality of left and right cosets is both a necessary and sufficient condition to guarantee that the product of two cosets will be a coset. In other words, the cosets of a subgroup form a group precisely when the subgroup is normal.

Independent of whether the subgroup is normal, and thus even when set of cosets does not form a group, structural information is provided via Lagrange's theorem, which says, once again, that the order of a subgroup must divide the order of the group. An immediate corollary is that the number of cosets is the missing divisor. In other words, if

H is a subgroup of *G*, then $|H|$ divides $|G|$, and the number of cosets of *H* is given by $|G|/|H|$. There is an obvious symbolic similarity with the notation for the quotient group, *G/H*. The similarity is significant, in that the elements of *G/H* are precisely the cosets, which implies that $|G/H| = |G|/|H|$.

This equation provides obvious potential for confusion. On the right side, the slash denotes a quotient of natural numbers, whereas on the left, the slash denotes a quotient group. Furthermore, this equation is true in the completely different but familiar context of high school algebra. If *G* and *H* are real numbers with $H \neq 0$, then the equation is an identity for the absolute value function, whose notation is that same as that for the order of a group. This analogy may reinforce students' common but mistaken impression that *G/H* is division.

At about the time of the third and fourth interviews, the students had been investigating cosets of subgroups and cosets of kernels of homomorphisms. The guiding question, which built upon the students' early work with set arithmetic was, "Is it possible to create a group with these cosets." The students also explored cosets more generally, proving, for example, that the cosets partition a finite group into equal-sized pieces (see "Problems to work on the week of April 22," Appendix B).

The students also had been investigating relationships between a group and its image under a homomorphism. Given a group homomorphism $f: G \rightarrow G'$, the students investigated f^{-1} of elements in the range of the homomorphism, which turn out to be cosets of its kernel. This fact follows from the proof on the second midterm exam that $f(a) = f(b)$ if and only if $aK = bK$, where *K* is the kernel of the homomorphism. The

students considered the operation table of the range of f and the operation table for f^{-1} of each element in the range (see "Notes on cosets, April 29," Appendix A). It turns out the two groups are isomorphic. This is the essence of the first isomorphism theorem, which says that the structure of the image of a homomorphism is identical to the structure of the cosets of its kernel.

Preceding all of the fourth interviews and some of the third interviews, Dr. Benson had introduced Lagrange's theorem and the term *quotient group* to give standard names to ideas that had emerged from the students' work on these kinds of tasks.

Carla and the Normal Group

During Carla's third interview, she computed the cosets $\{1, 3\}$ and $\{5, 7\}$ in U_8. I asked her whether she could use these cosets to form a group.

85   Carla: Well, let's see.... We will create a table with $\{1, 3\}$ and $\{5, 7\}$, of course, on the top and on the right. Actually, we talked about this in class. If we know that the right coset is equivalent to the left coset, then it does create a group. I think it's called a normal group, and these will.... I think we will have a group because I think that the right coset equals the left coset, because it really doesn't matter if you multiply on the right or on the left.

Citing results from class, Carla was convinced that the set would be a group because, as she correctly observed, the left and right cosets were the same. She was essentially correct, except for her use of the words *coset* and *normal*. I have already discussed her use of the word *coset* to denote a set of cosets. As for her use of the word *normal*, it seems she had lost track of the subgroup that gave rise to the cosets and was applying the word to the quotient group rather than to the subgroup. Furthermore, her left coset and normal group both named the same set of cosets, suggesting that the set of cosets was a very salient object for her.

Carla completed the coset products easily and quickly, taking advantage of the commutativity of the underlying group and organizing them in a table (see Figure 15). From the table, she noticed almost immediately that the group was isomorphic to the group on the set $\{e, a\}$, where e is the identity. Later she saw that it was also isomorphic to the range of the homomorphism, which was the subgroup $\{0, 2\}$ in Z_4, as detailed in chapter 5.

Carla had participated in an activity depicted structurally in Figure 19 and had created written records similar to Figure 15 and Figure 16. She admirably negotiated the various kinds of processes involved in the activity and seems to have had a good sense about the kinds of entities (e.g., sets or elements) she was dealing with. On the other hand, Carla seems to have attached the names *coset* and *normal* to the objects and processes in nonstandard ways. This hypothesis is further explored in a detailed analysis of her fourth interview.

Cosets in Z_{12}. Carla's fourth interview concentrated on examples and nonexamples of quotient groups, which required computing cosets and paying attention to whether the left and right cosets were equal. When I mentioned to her, as a preface to the interview, that we would be discussing quotient groups, she showed discomfort:

6   Carla: I have got to figure out what they are. Quotient group.... This is one of the things I haven't.... I've needed to study it, but I haven't done it yet. [Okay] What is the other name for a quotient group, that we gave? It wasn't a normal group right, that's different.

Thus, although Carla had computed a quotient group in the previous interview, she had not yet attached the name to the idea. I suggested we start with an example and return to the term later.

To get started, I asked Carla to find the subgroup generated by 3 in the group Z_{12}. She was momentarily confused about the operation in Z_{12} but quickly resolved that issue and soon determined that the subgroup generated by 3 was $\{0, 3, 6, 9\}$. I asked her to find the cosets of that subgroup.

22   Carla: Okay. First we will do left cosets. Probably the left coset and the right coset will be the same, but we'll start with the left coset. So, do you want me just to use the kernel as the ...

23   Brad: Do it whatever way you think is best.

24   Carla: Well, I'll start with the kernel at least because I am most comfortable with that. The kernel is 0 because 0 goes to 0. 0 is the identity.

25   Brad: You said 0 goes to 0. What does 1 go to?

26   Carla: No. 1 is not in there.

27   Brad: Oh. Okay. So what does 3 go to?

28   Carla: I don't think.... We haven't really.... Well, we know 0 goes to 0 because the identity always goes to the identity in two groups. But we haven't really defined what the others [inaudible]. The other, you know, function that we are dealing with [inaudible]. I am doing it with 0 and that's it ... because I know I can do it with the kernel because the kernel is.... Oh, wait a sec-. We do cosets.... We have to have a subgroup. And we have to perform an operation of the members in Z_{12} on the left, for the left coset. So you have members of Z_{12} * with the subgroup. And I already know what the subgroup is, so it's 0, 3, 6, 9. Took me a while, but ...

Carla momentarily thought that she needed a kernel to talk about cosets but then reasoned that only a subgroup was necessary. It seems from line 22 that she was still using the term "left coset" for the set of cosets, just as she had in the third interview.

Carla called the subgroup S and listed the elements of Z_{12}. In describing how she knew the operation was addition mod 12, she talked at first about a group G', suggesting she had not completely abandoned the idea that there was a homomorphism involved. Then she went on.

35   Carla: Let's just keep it simple and call S is a subgroup of G. So S is 0, 3, 6, 9, as I said before. And G is Z_{12}. And I am going to tell the truth here: The reason I chose addition mod 12 is because that's the only real operation we have right here. [Laughs.]

Carla began the coset calculations without hesitation. Her language evolved in the course of doing the calculations, becoming abbreviated: "If you are doing the operation of 0 on 0, 3, 6, 9 ..." (line 37), "If you do 1 star (which is adding mod 12) with 0, 3, 6, 9 ... 2 with S" (line 41). Soon she stopped doing the calculations directly and instead reasoned about the results, based on a pattern.

43   Carla: So we can see that there is a pattern here. 0 and 3 have the same left coset. And so 6 will and 9 will, because they are all multiples of three, because I can see a pattern of 3 here. So 1, the left coset, containing 1 is 1, 4, 7, 10. And that is the same as the left coset containing 4. So 1, 4, 7, and 10 are going to have the same left coset. So then it will be 2, 5, 8, and 11 that will have the same left coset.

The results of her calculations are shown in Figure 20. At first she had not recorded an operation in some places, such as between 2 and S, but when I asked her about the notation, she inserted an asterisk (*) and said that it was addition mod 12.

Figure 20. Carla's cosets calculations for {0, 3, 6, 9}

$$S = \{0, 3, 6, 9\}$$

$* = +_{12}$	lt coset		
	0*{0, 3, 6, 9}	=	{0, 3, 6, 9}
	1*{0, 3, 6, 9}	=	{1, 4, 7, 10}
	2*S		2, 5, 8, 11
	3*S		0, 3, 6, 9
	4*S		1, 4, 7, 10

Carla went on to compute right cosets. She first listed the calculations she intended to make, and then stopped after computing $S*0$.

55   Carla: And right away I see that the right coset is going to the same thing as the left coset because we are adding mod 12 and it's commutative. So it doesn't matter which side your single element is that changes.

57   Carla: So we have a normal group because.... I think it's a normal group that says that the left coset equals the right coset.

58   Brad: So what is the normal group here? When we have the left cosets and the right cosets being the same we have this thing called normal, but I want to know precisely what is normal here.

59   Carla: So the normal group is a set of sets. So one of the smaller sets will be 0, 3, 6, 9. [Okay.] Another of the sets will be 1, 4, 7, 10. [Okay.] Another one will be 2, 5, 8, and

11. So our normal group consists of three sets.

Two points are to be made here. First, Carla did this work with little interaction from me, and she noticed quickly that right coset calculations would yield the same result as the left coset calculations she had just completed. Second, her usage of the terms *coset* and *normal group* was consistent with her usage in the third interview. For Carla, the construction of the normal group required that the left coset equal the right coset. When I asked in what sense it was a group, she began by stating that the identity would be {0, 3, 6, 9}. To explain this, she decided to construct a table (Figure 21).

Figure 21. Carla's normal group with {0, 3, 6, 9}

	{0, 3, 6, 9}	{1, 4, 7, 10}	{2, 5, 8, 11}
{0, 3, 6, 9}	{0, 3, 6, 9}	{1, 4, 7, 10}	{2, 5, 8, 11}
{1, 4, 7, 10}	{1, 4, 7, 10}	{2, 5, 8, 11}	{0, 3, 6, 9}
{2, 5, 8, 11}	{2, 5, 8, 11}	{0, 3, 6, 9}	{1, 4, 7, 10}

While filling in the table, Carla at first performed the set addition by listing aloud all the pairs to be added. As she continued, she filled in the table according to what she believed it should look like. She used abbreviated procedures, partly to check her expectations, and she described much of her thinking aloud:

79 Carla: All right. So I know that this is a normal group, so I.... Because I know it's a group, I know it's going to create a group table. So I can.... I have a pretty good idea what this table is going to look like. It's going to have 1, 4, 7, 10 in the first column second row, and it's going to have 2, 5, 8, 11 in the position below that because 0, 3, 6, 9 is the identity, so it needs to just reflect what row it's in because I am working on the first column.

81 Carla: Now I think that the middle spot in the table is going to give me 0, 3, 6, 9. So I am going to.... Well, but I can see that it's not going to do that. So then I know that it's going to be 2, 5, 8, 11. The reason that I knew it wasn't going to do that is because I saw 1 + 1 is 2, and that's not in it. So that cancels that right there.

85 Carla: Yes, [you should always get another coset in these calculations] with a normal group. And I know that with my 3-by-3 group tables one of the options is to have all of the identities down the diagonal, so that's why I automatically thought that might be the identity in the middle position. But then when I checked it, I realized it wasn't. But it still can be.... I can still have a group table.

88 Carla: Since this is a group, I know that every element only appears once in each row or column, so I see that 0, 3, 6, 9 is left, for the second row. And then I can just check that 2 + 1 is 3, 5 + 1 is 6, 8 + 1 is 9, 11 + 1 is 12, which is 0 mod 12.

92 Carla: Even if you just see one of the elements, you know which coset it is going to belong to because the cosets don't overlap.

98 Carla: It has to contain 3, and that's the only coset that contains 3. So then I have one spot left so I know that has to be 1, 4, 7, 10 because both the third column and the third row lack that set.

Carla's approach, based on patterns and facts, was reasonably efficient and included some redundancy, which she used to catch errors. She knew every pair of cosets would produce another coset (line 85). She knew the result would be a group, so she had expectations about the patterns in the table (line 79), and she used the fact that in group tables every element appears exactly once in each row and column (lines 88, 98). She knew the cosets did not overlap (line 92), so she used representative calculations to determine which coset should appear in a cell and to check errors. She believed, for example, that, as a 3×3 group table, it would have the identity along the diagonal (lines 81, 85). That is incorrect, although both 2×2 and 4×4 tables can have the identity along the diagonal. This error of memory did not cause much trouble, however, because on the basis of a representative calculation she quickly realized that her belief was not correct.

Exploring the language. Carla had previously stated that the result would be a normal group, but I wanted to get some clarity on how she was using the phrase. When I asked her what the resulting table was and what it had to do with the word *normal*, she responded that it was a group table that had "everything to do with the word *normal*" (line 102). She continued:

105 Carla: Well, the thing that it has to do with is that this table is a demonstration of the group of those three sets—that they are a group.

109 Carla: The normal part of it just says that you got it because you had left and right cosets that were equal to each other.

111   Carla: That's how you know. That's how it differentiates from just any old group.

So the normal group was a group, as demonstrated by the table, and it was normal because there were left and right cosets that were equal to each other. These associations of the words *normal* and *group* are essentially correct. In particular, when the left and right cosets are equal to each other, the word *normal* applies. But the term is supposed to point back (see Figure 19) to the subgroup that led to the particular set of cosets not forward to the quotient group.

I next explored whether Carla could make any connection to the subgroup:

112   Brad: But these were left and right cosets of what?

113   Carla: Of S.

114   Brad: Of S. So does the word *normal* have anything to do with S?

115   Carla: I am not sure what you mean. Normal group, like just as far as words in the English language, doesn't really have any meaning to me, that's just what it's called.

116   Brad: Okay, what what is called? That's what I really want to get at.

117   Carla: Oh. The normal group is the group of cosets where the left and right coset of S are equal.

I take this last statement to be Carla's definition of normal group. She recognized that it was the subgroup that led to the cosets but did not see any reason to attach the word *normal* to the subgroup. Furthermore, she stated moments later that if the left and right cosets of S had not been equal, she would not have gotten a group.

At this point in the interview, it was clear that Carla had the right ideas but was using the words *coset* and *normal* in unconventional ways. I then explored what she would do with the conventional language:

121   Brad: Does the phrase *normal subgroup*, would that mean anything?

122   Carla: My guess would be that a normal subgroup would be a subgroup of a normal group.

126   Carla: If that's true, then a normal subgroup of this normal group that we are talking about could be this set {0, 3, 6, 9} because that's in the kernel. That is.... Well, actually I am

mixing things here. I know it could be 0, 3, 6, 9 because that's just the identity, so it's obviously closed and has its own inverse and ...

In response to my suggestion, Carla held on to her notion of normal group and applied her concept of subgroup to that. As the interview continued, she proposed {0, 3, 6, 9} as a one-element subgroup of her normal group because it was the identity of that group. She went on to legitimize this group by noting that there is no "rule against" a one-element group and that the group axioms were satisfied. I then decided to be more explicit about the term *quotient group*:

164   Brad: Now what if I were to tell you that the thing you have actually created here, this table with an operation table—I mean a group table for these three cosets—that's a quotient group.

165   Carla: Oh. Maybe the quotient group is.... My only guess would be that the.... The quotient.... I don't know if we ever defined it, actually. But my guess would be that the quotient group would be.... Well in this case the quotient group equals the normal group.... Do you see what I'm saying? I mean, it's the same ...

166   Brad: What if I were to tell you that the thing that you have been calling the normal group is the thing that *is* called the quotient group?

167   Carla: [Laughs.] Oh, okay. What is a normal group then?

Carla was willing to accept the term *quotient group* but was reluctant to let go of her usage of *normal group*. Despite my direct statement, she at first was still trying to figure out what a quotient group was (line 165) and only gradually came to decide that it was the same as her normal group, though perhaps only in this case. She had not yet detached the term *normal group* from her previous meaning. My next statement (line 166), however, seems to have caused her to consider a different meaning for *normal group*.

As the interview continued, Carla said, "I guess I have a problem with names" (line 171). I suggested that the idea was already there and that the issue was sticking a name on it. Then I encouraged her to connect the word *normal* to the subgroup:

178   Brad: But in order to talk about cosets you need to have what?

179   Carla: A subgroup.

180   Brad: Right, so this word *normal* tells us something about the subgroup that led to those cosets. Now let's see if *that* can make any sense.

181   Carla: Maybe it's when the subgroup is the kernel? I don't know. Because that's the only thing that I can think of that we've been talking about that kind of sets the subgroup apart. Like that's the only more specific characteristic of a subgroup that I can think of right now that we have talked about. The other subgroups were just subgroups, but if you have a subgroup that contains only the elements that map to the identity, then that's more specific.

182   Brad: Okay. But in order to have ...

183   Carla: We don't even have a homomorphism here. So, there goes that theory.

Even when I told her directly that the word *normal* applied to the subgroup (line 180), Carla was not able to make the appropriate gluing. The suggestion she made about the kernel, it should be mentioned, is correct in principle, in that the kernel of any homomorphism is a normal subgroup and any normal subgroup is kernel of a homomorphism. This is a pretty sophisticated view, however, and one that would ordinarily have required consideration of some major results, such as the first isomorphism theorem, that were not yet available.

Cosets in D_3. I next turned to the group D_3, which has both normal and non-normal subgroups. To begin, I asked Carla how she would write down the elements of D_3. She acknowledged the possibility of representing them as rotating and flipping triangles but chose to use permutation representations instead, which she also called the "123 way." She wrote down the six elements quickly. I asked her for the subgroup generated by (12), and she immediately responded that it would be just (1) and (12)

202   Carla: Because any subgroup has to have the identity, so that's (1). That's why (1) is in there. And then if you.... And (12) obviously has to be in there because it's generated by it. So if you operated (1) on (12) you just get (12). So it's in there so we are all set. And (12) operated with (12) just gives you (1). Another way you can think of it is that the order of (12) is 2. So if you call alpha (12), then the subgroup is only going contain alpha to the 0 and alpha to the 1.

Here Carla described two approaches and provided a hint of the extent of the proficiency she had developed both with the group D_3 and with the creation of subgroups more generally. She carried out six coset calculations (see Figure 22), making one small error that was quickly corrected. She saw that she got only three cosets as a result and suggested that the right cosets would be different because here "the order that you do that in does matter" (line 222). When computing the right cosets, her calculations were guided by expectations that grew from noticing which elements were paired with other elements in the cosets, an approach that implicitly took advantage of the fact that cosets do not overlap.

Figure 22. Carla's cosets of {(1), (12)} in D_3

$S = \{(1), (12)\}$

Lt Cosets			Rt Cosets		
(1)*{(1), (12)}	=	{(1), (12)}	S*(1)	=	S
(12)*S	=	S	S*(12)	=	S
(23)*S	=	{(23), (132)}	S*(23)	=	{(23), (123)}
(13)*S	=	{(13), (123)}	S*(13)	=	{(13), (132)}
(123)*S	=	{(123), (13)}	S*(123)	=	{(123), (23)}
(132)*S	=	{(132), (23)}	S*(132)	=	{(132), (13)}

Upon completing the calculations, Carla noted that the left cosets were not equal to the right cosets and claimed that with these cosets, an operation table would be a mess. I asked her to try to make an operation table with just the left cosets. Before computing the product of {(1), (12)} and {(23), (123)} she predicted the result would be a "four element thing" (line 246), demonstrating some proficiency with set arithmetic with non-normal subgroups. When her prediction turned out to be correct, she noted, "It's not closed" (line 256), because the product of the two cosets was not another coset. (See Figure 23 for the partial table.)

Figure 23. Carla's table for the cosets of {(1), (12)} in D_3

	(1),(12)	(23),(132)	(13),(123)
(1)(12)	(1)(12)	(23)(132)(123)(13)	
(23)(132)			
(13)(123)			

Note: Carla included no braces for the sets and omitted commas below the top line.

I next asked Carla to consider the subgroup generated by (123). She responded

immediately:

263   Carla: The subgroup generated by (123) is (1), (123), and (132). The reason that I know that's because I remember (123) and (132) are inverse of each other, from working with them before.

She seems to have known by recall that (132) and (123) are "inverse of each other," and

yet moments later she did not remember what (123) squared was. She called the

subgroup S again and computed the left cosets (123)*S and (132)*S quickly (see Figure

24), reasoning from the fact that S is closed. For the other cosets, her calculations were

guided by expectations based, for example, on the fact that cosets do not overlap. By the

time she was computing the right cosets, her procedure had become quite abbreviated,

demonstrating increasing proficiency with coset calculations and with D_3.

Figure 24. Carla's cosets of {(1), (123), (132)} in D_3

$S = \{(1), (123), (132)\}$

Lt Cosets			Rt. Cosets		
(1)*S	=	S	S*(1)	=	S
(123)*S	=	S	S*(123)	=	S
(132)*S	=	S	S*(132)	=	S
(12)*S	=	{(12),(23),(13)}	S*(12)	=	{(12),(13),(23)}
(13)*S	=	{(13),(12),(23)}	S*(13)	=	{(13),(23),(12)}
(23)*S	=	{(23),(12),(13)}	S*(23)	=	" "

After completing the calculations, Carla noted that the left and right cosets were the

same. She then reflected on the fact that, while order mattered in D_3, the order did not

seem to matter in these coset calculations:

303   Carla: Right, but we had the right combination of things so that.... Where the order mattered.... Because the elements of the sets can be in any order.... So where the order mattered it covered up for it because you had the right combination of elements.

314   Carla: Well, we just see that the subgroups in case of the subgroup generated by (12), we didn't have the right combination of elements ... or "right" as in to make the cosets equal, the right combination of elements so that they were equal. And in this case we did have the right combination of elements.

318   Carla: Well, one thing that might have helped us is that this subgroup has more elements in it. So it might cut down on the error.

Rather than just ascribing the differing results to the original subgroups, Carla looked inside the subgroups to determine what might have caused the results.

Exploring the language again. Equipped with these additional examples, I pursued the language again:

320   Brad: Do you suppose we could use the word *normal* to help us out here?

321   Carla: [Laughs.] I don't know.... Well, which one is normal? Can you tell me that? I don't, I have no idea what normal is. I thought I knew about it. I have no idea.

322   Brad: Well, you were saying before that you have a situation kind of called normal when you have left and right cosets being the same, right? You said something like that, didn't you? So where are left and right cosets here?

323   Carla: Well, left and right cosets are the same here. Maybe it has something to do with their inverses. Because this one (123) and (132) were inverses of each other, whereas in this one each element is its own inverse. That probably didn't answer your question. I was kind of half listening. What was it?

Carla recognized that she was not sure about the use of the word *normal*, and she recognized that the two subgroups were somehow different in that one led to left and right cosets that were the same and the other did not. Rather than merely using (and not using) the label *normal* to distinguish between them, she was looking for something deeper: some differing characteristics of the subgroups that would explain the differences in the cosets. As a result, she did not hear my question, so I repeated it:

324   Brad: Before you said that normal had something.... You used the word *normal* as having something to do with when the left cosets and the right cosets are the same.

325   Carla: Yeah. But I wasn't right on that 'cause that was quotient group.

326 Brad: Ah, well now, but what was quotient group?

327 Carla: Quotient group is one by, the set of cosets where the left and right cosets are equal. I guess.

This last statement is almost identical to Carla's previous definition of normal group (line 117), suggesting that she had substituted the new name for the old idea. As for the differences between the subgroups, she was still trying to find a reason:

339 Carla: Well, yeah, but the only thing that I see ... the only difference that I really see between, at least right now, between them, is the thing about the inverses. I'm not seeing any other ... unless it is just normal because it's generated ... [inaudible] doesn't make sense. I don't know. It doesn't.... If a quotient group is the set that you get when, from the cosets, if the left and the right cosets are equal, then maybe the normal group is the group that you use to get the left and right cosets in that case.

340 Brad: What do you mean?

341 Carla: Is the normal group the group that is always going to be your S, when you find your left and right cosets in the way we do the quotient group?

342 Brad: Oh, so you mean the subgroup we started with?

343 Carla: Yeah.

351 Carla: So I would say that this subgroup is a normal group. If, well, if my guess is right, this subgroup is a normal group, and that one isn't. [She points first to the subgroup {(1), (12)} and then to {(1), (123), (132)}.]

It seems that at this point in the interview she had at last made the desired connection between the word *normal* and the subgroup that determined whether the left and right cosets were the same. The only remaining difficulty, it seemed, was that she was calling it a normal group rather than a subgroup.

352 Brad: What about saying "normal subgroup"?

353 Carla: [Laughs.] Well, that kind of would be doing left and right cosets of left and right cosets, I would think. You know. Go another step into it.

354 Brad: What do you mean?

355 Carla: Well, I would think a subgroup is usually a smaller group. So I would think that a normal subgroup, you have to get that from a quotient group. And that it would have to give you a different quotient group.

Carla gave meaning to term *normal subgroup* in a way that indicated that for her the term

normal was still strongly connected to the quotient group, maintaining some of its

previous meaning. Yet a few moments later, she reconsidered:

363   Carla: Right. So, well, then maybe we were all ... we only talked about normal subgroups and not normal groups. Is that what you're saying? Because I'm fine with calling it a subgroup. I really don't care. Because if you don't really talk about normal groups and you only call them normal subgroups, you're just focusing on the fact that they are subgroups.

While thinking about the terms used in class, Carla relaxed the tie between the word

normal and the quotient group and was willing to consider alternatives. She realized that

she may have been remembering the wrong term and that if the correct term was *normal*

subgroup then perhaps it made sense to apply it to the subgroup that led to the cosets. I

closed the interview by checking on how she labeled the various objects and the

connections among them.

364   Brad: So, which is the subgroup here? Which is the normal subgroup?

365   Carla: S.

366   Brad: And it's a subgroup of what?

367   Carla: D_3.

368   Brad: And it's normal because ...

369   Carla: It is what you use to get left and right cosets which are equal.

370   Brad: And this subgroup here, generated by (12) ...

371   Carla: It's not a normal subgroup.

372   Brad: Okay. But it is a subgroup of ...

373   Carla: D_3.

374   Brad: D_3 still. So here we have one subgroup of D_3 which the left and right cosets were different. So we say its ...

375   Carla: Not a quotient group. And S is not a normal subgroup.

376   Brad: Well, you don't get a quotient group.

377   Carla: Right. This set of these is not a quotient group.

From Carla's correct usage of the term *normal subgroup* in the above exchange, it is clear that some learning had occurred. Because this exchange required only short answers from Carla, however, it is not clear to what extent she had changed her meaning and usage of the word *normal*. On the final exam, she did use the term correctly, writing, for example, "*H* is not a normal subgroup of *G* because lt. cosets ≠ rt. cosets" in response to the question, "Is *H* a normal subgroup of *G*? Explain" (problem 9d). It would be interesting to know how she would have responded to a more open-ended question such as, "Compute the left and right cosets of *H*. What can you say on the basis of your calculations?"

Analysis. The schematic diagram in Figure 25 is an expanded version of Figure 19, showing a few more of the processes and objects involved in activities related to quotient groups. The core of the activity is depicted vertically along the center of the diagram. Given a group and a subgroup, compute the left and right cosets. Compare the set of left cosets with the set of right cosets. If they are the same, then designate the original subgroup as a normal subgroup. Calculate the various products of left or right cosets. If the sets of left and right cosets are the same, identify the set of cosets (with their products) as a quotient group.

This core activity may be enlarged in two ways. First, the group and the subgroup may be identified as the domain and kernel, respectively, of a homomorphism between two groups. This is similar to Fraleigh's (1989) treatment, in that the concept of quotient group is introduced as the group of cosets of the kernel of a homomorphism. The second way to enlarge the core coset activity is to ask, once the quotient group has been calculated, whether the quotient group is isomorphic to a familiar group. Several

students did this naturally on the basis of the patterns in the table, as discussed in chapter 5.

Figure 25. Quotient group activity
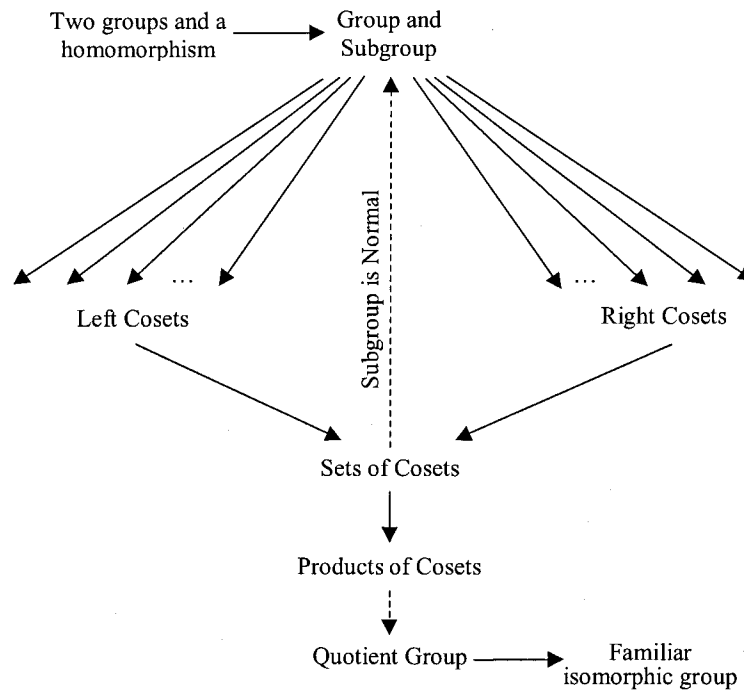
Note: Arrows indicate processes. The dotted arrows denote designation processes that are legitimate only if the left and right cosets are the same (the result of an implicit comparison process on the sets of right and left cosets).

These extensions are useful in the analysis because for Carla these extensions were sometimes natural parts of the activity. In particular, in one task she at first wanted the subgroup to be the kernel, demonstrating a connection between the core activity and the concepts of kernel and homomorphism (Carla 4, line 28). And after completing the group table for {1, 3} and {5, 7}, she went on to show, without any prompting from me, that the group was isomorphic to a familiar group with two elements.

The extensions are also useful for mathematical reasons because the entire diagram can be tied together with the first isomorphism theorem, which says that, given a group

homomorphism, the quotient of the domain and the kernel of the homomorphism is isomorphic to its image. This is illustrated in the isomorphism that Carla noticed between the quotient group $\{\{1, 3\}, \{5, 7\}\}$ and the subgroup $\{0, 2\}$ in Z_4.

Building on the analysis above, it seems that for Carla, the entire center of Figure 25, from just below "Subgroup" to "Sets of Cosets," was essentially one concept: coset. Although within that concept she could distinguish between left and right cosets, she did not distinguish vertically between the processes and objects. Regarding her use of the word *normal*, she correctly associated the word with the comparison of the cosets, but rather than reaching backward in the activity to attach the label to the subgroup, she reached forward and attached it to the quotient group. This labeling is not surprising when one realizes how distant the subgroup is when one is comparing the sets of right and left cosets. Furthermore, once the name *normal* was attached to the quotient group, it was very difficult for Carla to make any connection between it and the subgroup.

What is important in the diagram in Figure 25 is not the names of the processes and objects but the structure—the relationships among the various components. Just as learning group theory requires abstracting from the particular names of the elements and the operations, thinking about the learning of group theory requires abstracting from particular names of the objects and processes. This conclusion alone is not very surprising, as it is obvious that different languages use different terms for concepts. The above analysis, however, suggests more. First, to learn group theory requires not only attaching names but also carving the activity into concepts. Second, making new distinctions and changing one's language both require accommodation in the sense of

reorganizing one's conceptual structures. These ideas are elaborated in the next subsection.

Diane, Lori, and Quotient Groups

Diane and Lori's fourth interview took place the same day that Dr. Benson had introduced Lagrange's theorem and the term *quotient group*. Diane and Lori were able to compute an operation table for a quotient group, but whereas Carla had a nonstandard name for the quotient group, Diane and Lori had no name at all. Instead, their concepts of coset and quotient group were aided and obstructed by vague connections with Lagrange's theorem.

While computing cosets of a subgroup in D_3, Diane stated that there would be two cosets, "because the order of this [subgroup] is 3, the order of D_3 is 6, so 6 over 3 is 2" (line 84). The calculation is correct, by Lagrange's theorem (see the subsection Conceptual Analysis above). I asked them for justification.

88   Lori: Probably the definition came up in class.
90   Lori: That's like Lagrange's theorem, wasn't it?
92   Lori: I don't remember it, because I just learned it [inaudible].
93   Diane: Well, if this is a subgroup ...
95   Diane: ... then we know that it's a normal group of this and we got the subgroup of that.
96   Lori: It's the order of the group divided by the subgroup. Like order G divided by H.
97   Lori: Yeah, it was order G divided by H, but that was G mod ...
98   Diane: Well, they're similar; they are the same definition G mod H.

Rather than a clear justification, Diane and Lori provided vague associations with a number of phrases and symbolizations: Lagrange's theorem, normal group, order G divided by H, and G mod H. Many of these words could be part of a correct justification, but some of the phrases are nonstandard. The data are insufficient to provide a sense of

what Diane meant by the term *normal group*, but because she does not use the term later to describe a group of cosets, it seems unlikely that she was using the term as Carla had.

The phrase "order G divided by H" is ambiguous because of the lack of parentheses in everyday speech: Did Lori mean $|G|/H$ or $|G/H|$? The former is mathematically problematic; the latter is standard notation only when the quotient group G/H is defined, which requires that the subgroup be normal. In that case, it is numerically though not logically correct as a way to count cosets because G/H and the set of cosets of H are identical as sets.

The issue of the meaning of G/H came up again later in the interview. Diane had again called upon the calculation 6 divided by 3 to justify that there would be 2 cosets. Recalling that we had proved this result in class, I asked them what about the proof implied that 6 divided by 3 made sense?

291   Lori: That was the quotient group definition; Lagrange is a little different. Lagrange, like, stems from it. Is that correct? You actually had G divided by H under the whole quotient group definition. I remember that.

292   Brad: Right, but remember G is a group, and H is a group....

294   Brad: So a group divided by a group is a little different than a number divided by a number. I mean, you are doing 6 divided by 3.

295   Lori: Well, we are doing the order of G.

296   Diane: If you get a group that's 6 elements, divided by a group of 3 elements, you should get 2 groups of 3 elements each. [Lori writes G/H.]

Diane's phrase "divided by a group of 3 elements" suggests that her work with cosets and with Lagrange's theorem had become something like quotitive division: 6 elements ÷ 3 elements per group = 2 groups. This statement would be almost correct if she had used the term *coset* rather than *group*. But again she did not provide justification.

From what Lori wrote and said, it appears that she was considering only the orders of the groups in her division, so she did not see a problem of meaning in either "G divided by H" or "the order of G divided by H." As the interview continued, she seemed more concerned with what to write down than with its meaning:

326   Lori: Yeah, isn't it like the order of G divided by H? [Okay.] Or the order of both, which is it? I am not clear on the definitions.

327   Brad: Well, clearly if we talk about G the group divided by H a subgroup, and write it that way ...

328   Diane: If you're taking the orders ...

The way Diane interrupted me suggested that she too was thinking mostly about the orders of the groups.

I again distinguished between dividing groups and dividing their orders and stated that the quotient group was written as G/H, which Lori described as "no order" (line 332).

334   Lori: [Writes "quotient: G/H."] Did you even right this on the board or am I jumping ahead? Just because I remember this. Did we talk about this today?

336   Diane: We did something with order.

337   Lori: Then this must be Lagrange's theorem. It must be like this. [She changes earlier G/H to $|G|/H$. See line 296.]

338   Diane: No, it was on the left board. It wasn't near Lagrange. Cause Steve wrote Lagrange, and Steve wrote something else. He had something like G, order G, and something like order of K, I think it was. [pause] Yeah. Because K is the subgroup and this is a subgroup. So you take G, which is this, and divide it by the order of K. I think that was it. [Lori changes $|G|/H$ to $|G|/|H|$.]

Thus, for Diane and Lori, the concepts, processes, and notations for Lagrange's theorem and quotient groups were mixed together in a jumble of vague memories from class. Moreover, the way that Lori moved flexibly from G/H to $|G|/H$ to $|G|/|H|$ suggests that the vertical bars did not carry much meaning for her and reinforces the point that the symbolizations were about the orders of the groups and not the groups themselves.

After we had determined that the subgroup was normal, we were ready to talk about

quotient groups. I asked what could be done with normal subgroups.

368   Lori: Maybe figure out something with quotient groups?

373   Lori: So quotient group is just G divided by H.

374   Brad: Well that's the notation for it.

375   Lori: That's not the definition.

376   Diane: No.

377   Lori: Can I see the definition? I am not too clear on it.

378   Diane: No, I am not too clear.

Thus, for Lori and Diane, *quotient group* had little meaning beyond the symbolism G/H,

further supporting the idea that they had been dividing not groups but orders of groups all

along.

I asked several questions about what they typically might do with the set of cosets but did

not get much of a response until I asked whether it would be possible to create a group.

Diane began making some calculations. Lori did not see what she was up to, so I asked

Diane to explain.

405   Diane: I am multiplying each element ...

407   Diane: ... of one set with each element of another, of the coset. Take (1) divided by each
        element, (123) divided by each, multiplied by each element and then (132) multiplied by
        each element. I think we're going to get a [group].

They completed their calculations and saw that it was a group with two elements that

were sets, where one of the sets acted like the identity. I asked them what to call it.

443   Diane: A group of cosets? No. A group generated by cosets?

445   Lori: I am sure it has another name, though.

446   Diane: ... [If] you say a group of cosets, then you can say any cosets you absolutely want
        to. It has to be a little more.

448   Diane: A group of cosets under D_3.

450   Diane: The elements of D_3 generated by ...

452   Diane: ... generated by (1), (123), (132)?

Diane's suggestion of "a group of cosets" was a good start, but she had a sense that she

needed to provide a more precise description. Lori later likened it to a subgroup but saw

that it was not a true subgroup, because "things are like sectioned off" (line 460).

They continued suggesting names:

462   Diane: Probably something so obvious. A group of cosets.
463   Lori: Group coset. Coset groups?
464   Diane: Normal subgroup. Normal group? Normal coset group.

Again, they provided good suggestions, but they were not connecting with the language

that had been introduced that day in class. I asked them what other words had been used

in class and they suggested *isomorphism* and *kernel*. With that, they noticed that their

group was isomorphic to the $\{e, a\}$ group. I reminded them that the left and right cosets

had been equal, and they pointed out that H was a normal subgroup.

499   Diane: It's two [cosets] together, and since this is the table generated by it [the normal subgroup] maybe this is the normal group.
500   Lori: Normal subgroup group.

The strong connection with normality provided some reasonable yet nonstandard

language that was, in fact, identical to Carla's language, suggesting once again that

Carla's unusual language was not so unusual after all.

I decided to intervene:

524   Brad: What about quotient group? What do you suppose a quotient groups is?
525   Lori: This is a "divided by" it's not like a contained in some little subgroups [inaudible]. It's G divided by H.
526   Brad: Well, that's the notation for it.
527   Lori: Okay. Steve said he liked to think of it as a remainder, like in the Z_4 case.
528   Diane: I think this would probably be more of a quotient group, because you want to say that this group is only generated by this element, and you would only get this table right

here because you're not taking into consideration this element.

529   Lori: Right, and when I said that this normal subgroup, when I said that this generated this table ...

531   Lori: ... it didn't really, because if you just had it how would you come up with this set? So this is the normal subgroup, and this must be the quotient group, generated by that normal subgroup.

They both were uncomfortable about the idea of saying the normal subgroup "generated" the other cosets. Nonetheless, they eventually both pointed to the operation table they had created for the cosets and decided that it must be the quotient group. Lori, for example, said, "I'm thinking, by definition, this must be the quotient group" (line 540).

Thus, in the end, both Lori and Diane were able to compute and recognize a quotient group. They had no trouble making the calculations, once they got started, but they associated neither the process nor the result with the term *quotient group*. While trying to find a name for the group of cosets they had created, they had several good suggestions, including Carla's term *normal group*, but they decided to call the group a quotient group only after I brought up the term.

Other Students and Quotient Groups

An overriding theme in the analysis of the interviews above was the problem of attaching names and notations to processes and objects with which the students had developed some proficiency. In this section, I elaborate on this theme. The section opens by first returning to the previous discussion of the process/object distinction in the students' understanding of cosets and then broadening the discussion to include the concept of quotient group. There was a strong sense in which the students understood cosets and quotient groups as both processes and objects, and yet there were ways in which their

conceptions were incomplete, as evidenced by significant linguistic and notational difficulties.

The linguistic difficulties manifested themselves primarily in the strong connection between the concepts of normality and quotient group, because, after all, they both depend upon left and right cosets being the same. Thus, I next discuss the students' understanding of the concept of normality followed by a presentation of the various ways that students tried to name the quotient group. The section closes by bringing together the issues of naming, notation, processes, and objects by considering the metaphor of gluing names to ideas.

Processes versus objects. Although a few of the key participants experienced initial difficulties computing cosets, they all were eventually able to manage the process, some of them with considerable proficiency. Thus, all the students developed process conceptions of cosets. After computing the cosets of a subgroup, all the students were willing to talk about cosets as being elements of a larger structure, a set of sets. They could compare left and right cosets and notice whether they were the same or different. They could perform coset arithmetic, even in the case of D_3, and could talk about whether a pair of cosets produced another coset. Thus, metaphorically, cosets were also objects. Nonetheless, the students had difficulty using notation to support their thinking, and they sometimes used language ambiguously in ways that suggested that their thinking was immersed in the process of computing all of the cosets.

To begin their computations of quotient groups, the students often needed a direct question such as, "Can we make a group out of these cosets?" With that question, the process for creating such a group seemed obvious and natural to everyone except Lori,

although she was quick to latch on to what Diane suggested. Operation tables served both to organize the students' calculations and to help them see that the result was or was not a group. Operation tables helped them think of cosets as objects that can have inverses or that can function as the identity. With the operation table they saw isomorphisms between the quotient group and other familiar groups. Thus, metaphorically, quotient groups were objects as well. Nonetheless, the students had trouble with the term *quotient group* and with the notation G/H.

A brief characterization of the students' concept images is roughly as follows: Cosets were objects, but aH was a process. The term *coset* applied both to an individual coset and to the set of all of them. When the left and right cosets were the same, the "left coset" could form a group, which was called "the normal group." Both of these were objects. Coset arithmetic was a process, and the resulting operation table was an object, yet the terms *quotient group* and the notation G/H referred to neither of these.

Clearly the hierarchical process/object distinction is insufficient to explain these results. What does it mean to say that a student thinks of quotient groups as objects when the student does not call those objects quotient groups? What does it mean to say that a student thinks of cosets as object when the notation aH is not part of that understanding? I reconsider the process/object distinction in chapter 8.

Normality. The students' concept images of quotient group were very closely tied to their concepts of normality. The common thread among the students' uses of the word *normal* is captured by Carla's statement, "You had left and right cosets that were equal to each other" (Interview 3, line 109). This characterization applies both to standard and

nonstandard uses, including Carla's "normal group." All of the students gave such a characterization of the word *normal* at some point during their interviews.

What was different among the students' conceptions was the object to which the word *normal* applied: to the group of cosets or to the subgroup. When one determines that the left and right cosets are the same, the standard language points both back to the subgroup and forward to the group of cosets, attaching the names *normal subgroup* and *quotient group*, respectively. All of the key participants except Carla were able to point back to the subgroup to call it normal, although they demonstrated varying degrees of certainty. Diane and Lori were willing to point forward while searching for a term to describe the group of cosets, suggesting terms such as *normal group* and *normal coset group*, thereby supporting the reasonableness of Carla's language. The terms *quotient group* and *normal* and the ideas they represent are so strongly connected perhaps it should not be surprising that the terms are sometimes confounded.

The students also demonstrated strong connections between normality and commutativity, often concluding correctly that the left and right cosets were the same when the group was commutative. Robert's concept image of normality was strongly connected to the word *Abelian*, a label applied to groups in which the operation is commutative. In his fourth interview, for example, when comparing left and right cosets of a subgroup, he said, "It does not look like this thing is Abelian as I predicted" (line 148). In his explanation, he focused on whether the elements themselves commuted with each other and then whether the left and right cosets were the same. Later, he seemed to be thinking that when the left and right cosets were the same, the resulting quotient group

should also be Abelian, a prediction that is true for all of the groups the class investigated in detail but false in general.

Naming the quotient group. All of the key participants demonstrated some difficulty attaching the standard name *quotient group* to the group that they created from cosets. For example, after completing an operation table for some cosets, Robert said, "I'd describe it as a normal subgroup. But actually, no. It's a group somehow that's like generated by a normal subgroup" (line 351). Later, he provided another formulation: "I formed a group of … an Abelian group using cosets generated by (123)" (Interview 4, line 391).

During her third interview, Wendy called the result "the coset group … because it's a subgroup of coset elements" (lines 309-313). I asked her for clarification on the word *subgroup*, and she rephrased her response: "So it's the group of all of the elements in the, all of the cosets elements" (line 315). By her fourth interview, Wendy had attached the term *quotient group* in the standard way. She explained that "a quotient group is the operation table of the cosets … elements" (line 6) and later, after computing an operation table for the quotient of $4Z$ in Z, she characterized the quotient group as "the group containing all the cosets" (line 219).

Thus, all the students came up with names for quotient group that indicated reasonable conceptual connections and that demonstrated a good informal sense of the concept. The connections to words *normal* and *coset* were particularly strong. Wendy was the only key participant who was able to give a characterization of the term *quotient group* without my intervention. Wendy's fourth interview took place after the final exam, however, so I hesitate to draw conclusions from this comparison with the other students.

In any case, Carla's insistence on calling a quotient group a "normal group" is not really surprising.

Gluing. The diagram entitled "Quotient Group Activity" (Figure 25) does not paint a sufficiently detailed picture of the conceptual complexity of what there is to learn. If the figure were to be augmented to include various symbolizations and Lagrange's theorem, however, it would become unwieldy and defeat its purpose. Thus, rather than constructing another figure, I state more simply that the overriding issue regarding the concept of quotient group was the confusion among the following:

- names
    - Lagrange's theorem
    - quotient group
    - normal [group or subgroup]
    - coset group
- symbolizations
    - G/H
    - $G \bmod H$
    - order G divided by H
    - $|G|/H$
    - $|G/H|$
    - $|G|/|H|$
- processes and objects
    - counting cosets by dividing the orders of the groups
    - calculating the products of the cosets
    - the resulting group [given by its table]
    - the subgroup that gave rise to the cosets

Each of the above is a signifier of some aspect of the students' activity regarding the concept of quotient group. Some of them are standard signifiers; others were invented by the students. The list is not meant to be exhaustive but rather is intended to demonstrate the complexity of making the standard connections among the various signifiers. Clearly learning the concept of quotient group is more complicated that is suggested by the metaphor of gluing names to ideas (see, e.g. Hewitt, 2001).

Perhaps some of the confusion was caused by the fact that the standard names, notations, and processes were introduced at about the same time. This explanation is probably insufficient, however, for the notation G/H seems to be very easily interpreted as division of natural numbers, and that was not the only confusion. Because the students were able to carry out the processes adequately and appropriately and were able to talk about the results as objects, a more plausible explanation is that there is considerable cognitive work to be done in attaching the names and notations to the objects and processes.

Summary

Once again, the overriding issue regarding the concept of quotient group was one of language and notation. The students were happy to consider the cosets as elements of a larger structure. They often knew that in order for the cosets to form a group, the left and right cosets needed to be the same. Despite the fact that many of them needed some prompting about coset arithmetic, they seemed to regard the procedure as reasonable and even natural. By organizing their calculations in an operation table, they were able to see whether the result was a group. Thus, the students conceived of cosets and quotient groups as both processes and objects. Attaching language and notation to these concepts and calculations was problematic, however, and the students were aware of these difficulties. Regarding the term *quotient group*, other terms such as *coset group* and *normal group* seem to be more natural. More generally, some of their difficulties seemed to arise out of the strong procedural, conceptual, notational, and experiential commonalities between the terms, making it difficult for the students to manage the connections among them.

Main Themes

Regarding the concepts of homomorphism, coset, and quotient group, the main themes developed at the end of chapter 5 also serve well for organizing the results of this chapter. Once again the students' use of language and notation emerged as a central issue. Once again, the students used operation tables and other conceptual tools for managing their relationships with abstract ideas.

With each of the topics in this chapter, the students' language and notation was sometimes nonstandard and often imprecise. Furthermore, the students sometimes had trouble with the concepts and process because of confusions about the language and notation. And even when they seemed to understand the concepts, they sometimes had trouble using language and notation in standard ways to support their thinking. They often left off quantifiers in their definitions of homomorphisms and began reasoning symbolically without adequately specifying their symbols. Nonetheless, at other times, the students' language was nonstandard yet precise. Carla, for example, consistently called a quotient group "the normal group." Furthermore, the students' seemingly idiosyncratic language was often not idiosyncratic, as there were commonalities across students.

The students often did not distinguish adequately between a set and an element, particularly regarding notation but even regarding their use of the word *coset*. The notation *aH* was sometimes a particular coset, sometimes the set of all of them, but mostly a process for generating cosets, although the notation did not always support the students' understanding of the process. The data suggest the problem of distinguishing between a set and an element might be better described as a difficulty conceiving of a

symbol as representing a generic yet particular object. For some students, symbols such as $4x$ represented neither a specific nor a particular multiple of 4 but rather any multiple of 4. And if it can represent any multiple of 4, it is a short cognitive leap to imagine that it represents all multiples of 4. This kind of reasoning was characterized as being immersed in the process of generating all such elements.

The metaphor of gluing is clearly unsatisfactory to describe the cognitive requirements of making connections among language, notation, processes, and objects. Instead, learning to make standard linguistic and notational distinctions seems to require first conceptual distinctions and then a dialectic that supports connections among them. Thus, both creating and changing a student's use of language and notation require building cognitive structures that support and fit the linguistic and notational distinctions.

To manage their relationships with abstract ideas, the students often used operation tables to support their reasoning. Operation tables served to organize the students' calculations for determining whether a function was a homomorphism. Organizing their coset calculations in an operation table seem to help them see whether a set of cosets formed a group and also helped them see the result *as* a group—an object—with elements that were sets.

With the concepts of coset and quotient group, the students also supported their thinking through proficiency with the concepts, examples, representations, and related facts. For example, when creating tables of quotient groups, they demonstrated considerable proficiency with abstract groups and their representations, using facts about the tables in order to support their calculations and catch errors.

The process/object distinction was helpful in characterizing the students' thinking but was insufficient for making developmental distinctions. The students could conceive of cosets as objects, yet the notation aH specified a process. They conceived of quotient groups as objects, but confused the notation G/H with Lagrange's theorem. The data suggest that two encapsulations might be required for cosets: one to see a specific coset (say in Z_{12} or D_3) as an object and another to see aH as representing such an object, which harkens back to the problem of imagining generic particular objects. These issues and themes are further elaborated in chapter 8.

CHAPTER VII

PRELIMINARY MATHEMATICS

This chapter includes discussion of the relationships between the concepts in group theory and preliminary mathematical ideas that became prominent in my analysis of student understanding. While success in abstract algebra clearly requires broad and strong background knowledge, it was not clear a priori what background concepts would be implicated. The two concepts for which there were sufficient data for an analysis are functions and modular arithmetic. Because neither of these topics was explicitly investigated in the interviews, it is not possible to provide a thorough analysis of the students' understandings of either of them. Instead, I provide an analysis of particularly salient episodes that raise some interesting issues and then discuss how those issues played out more broadly with the other students.

Before presenting my analysis of the students' understanding of functions and modular arithmetic, however, I make brief comments about two other concepts: exponents and zero. These preliminary mathematical ideas likely played a significant role in the students' thinking, but little detailed data was available. The topics deserve mention here because of their importance in both group theory and school mathematics and because of potential implications for an abstract algebra course that aims to provide opportunities for students to strengthen their understanding of these key ideas.

Issues related to exponents came up periodically when, in unfamiliar settings, students were unsure what g^0 might mean or whether g^{-4} should mean $(g^{-1})^4$ or $(g^4)^{-1}$. In abstract algebra, as in school mathematics, the rules for exponents are initially defined for exponents that are positive integers. Those rules are extended first to allow exponents that are zero and then exponents that are negative integers. The guiding principles behind the extensions are, first, that the rules that work in the original system should continue to work in the extended system and, second, that the original system should be isomorphic to a subset of the extended system. Only with these guiding principles can the conclusions be adequately supported that g^0 should be the identity and that g^{-4} may be either $(g^{-1})^4$ or $(g^4)^{-1}$ because they must all be equal. These extensions are identical to the extensions from the natural numbers to the whole numbers to the integers— extensions that prove to be important and difficult for primary school children (Kilpatrick et al., 2001). Thus, it should not be surprising that extending the rules of exponents requires some mathematical and cognitive work.

A more surprising set of issues that came up from time to time, and likely had some influence on the students' difficulties with exponents, had to do with the ontological status and properties of zero. Some students were convinced, for example, that zero is neither even nor odd. Some were convinced that $0/0 = 0$. Because the status of zero as a number presented considerable obstacles historically (see, e.g., Kaplan, 2000; Seife, 2000), perhaps it should not be surprising that zero continues to be a difficult concept for some advanced undergraduates. (See Nardi, 2000 for an extended discussion of the role of zero in advanced mathematics.)

The concepts of exponents and zero are important in both undergraduate and school mathematics. With the available data, however, I am able to say only that the concepts deserve more attention both in the teaching of undergraduate mathematics and in research about students' understanding of undergraduate and school mathematics. Regarding the concepts of functions and modular arithmetic, this study provides more evidence and insight.

As mentioned in chapter 6, Diane's concept of function became a significant obstacle in her understanding of homomorphisms. An analysis of her third interview revealed similarities with the thinking of students who were able to complete the interview tasks successfully. Regarding modular arithmetic, several students used the word *mod* with unusual syntax, but the analysis showed Carla's syntax to be consistent and mathematically insightful. Other students demonstrated conceptual difficulties that may be explained by distinct but related uses of the word *mod* in standard mathematical language.

Functions

The concept of function plays a role in the abstract algebra concepts of isomorphism, homomorphism, and binary operation, though it is easy for its role to remain implicit for the concept of binary operation and for informal versions of isomorphism. The concept of function figured prominently when the students were dealing with homomorphisms, as they tried to manage the relationship between the function provided as a potential homomorphism and the binary operations in the domain and codomain of the function. Before discussing these relationships in the students' understanding, I provide a brief

discussion of two observations that suggest that some of the difficulties identified in the literature on the learning of functions continue to be difficulties for mathematics majors in advanced courses.

Based on the literature on the learning of functions, it is not surprising that some of the students in this study wanted functions to be given by formulas (see, e.g., Vinner, 1992; Ferrini-Mundy & Graham, 1994). For example, when I gave the students a potential homomorphism by specifying the image of each element in the domain, both Diane and Robert wanted to know what the formula was, and Carla also strongly identified the function with its formula. They were willing to assume that there was a formula, however, and thus were able to continue with the interview and with the tasks I had proposed.

The students' concepts of function were also marked by confusion between closely related terms and ideas. Carla, for example, temporarily confused the roles of x and y regarding the real-valued function described by $y = x^2$ when she said that when x is 3 the y-value would be $\sqrt{3}$ (Interview 2, line 32). A number of students in the class interchanged the terms *range* and *codomain*, an issue I chose to explore with Carla in her second interview. In the interview, she displayed connections and understandings built up around the terms. For example, she talked about the idea of restricting the range of a function, analogous to the way the domain of a function is often restricted in mathematical discourse. Carla's "equal-handed treatment of x and y" (Lauten, Graham, & Ferrini-Mundy, 1994, p. 233) seemed to contribute to a robust yet problematic concept image. Thus, it seems unlikely that her confusion could have been fixed by merely telling her she had the terms reversed.

The remainder of the discussion on the concept of function is intended to illustrate some ways in which the students' understandings of functions mattered when they approached concepts in abstract algebra. I base this discussion on an episode in which Diane demonstrated particularly unusual understandings and then broaden that discussion to illustrate how similar phenomena arose with Lori and Carla.

Diane, Functions, and Homomorphisms

Diane's third interview focused mostly on homomorphisms and functions, and we did not get to the tasks about cosets. The following episode illustrates three aspects of her understanding of functions, which together impeded her progress on the interview tasks and likely obstructed her understanding of the concept of homomorphism. First, Diane's concepts of function, homomorphism, and binary operation were connected and intertwined in unusual ways that led to implicit (i.e., unspecified) homomorphisms. Second, she had trouble attaching the names of and her associations with the concepts one-to-one and onto to the objects and processes under consideration. Third, she had an unusual concept of function that was connected to and supported by the notation. These issues were revealed slowly over the course of the interview.

When Diane provided her definition of homomorphism, she did not mention the word *group*, which perhaps is not surprising, since students in the course almost always were dealing with some group or other. I asked her how groups were involved:

15   Diane: If a group is a homomorphism, that means you can do this $[f(a*b) = f(a)*f(b)]$
        with every single element and have both sides be true and a group is a homomorphism.

This appears to be merely unusual syntax, but, as will become clear below, for Diane a group could be a homomorphism and the group operation supplied the function.

I asked her to come up with an example and to show that it was a homomorphism.

19 Diane: So, from like U_8 I have 1, 3, 5, 7. So for a homomorphism I would have 1, 3 would equal $f(1)$ and $f(3)$ like that? [Writes $f(1*3) = f(1)*f(3)$.]

20 Brad: Okay, so what is $f(1)$ then?

21 Diane: $f(1)$ is 1, and then $f(3) = 3$. $f(1*3)$ here is 3, and $f(3)$ is 3. So it's 3, and then 1*3 is 3.

Diane began verifying the function was a homomorphism before she had specified what the function was. From her calculations (lines 21 and following), however, it is clear that the function she had in mind was equivalent to the identity function on U_8, which, as an isomorphism, is necessarily a homomorphism. It seems unlikely that she would have described her thinking this way.

I asked Diane how she had decided that $f(5)$ was 5 and $f(7)$ was 7:

27 Diane: Well the function here is mod 8. So under mod 8, 5 is 5, and 7 is 7. So the function here is U_8.

31 Diane: The function here is U_8. Is the mod 8. Multiplication.

So for Diane, the group U_8 provided a function, and the signifiers U_8, mod 8, and even "mod multiplication 8" (line 41) were synonyms for that function. It is true that a group operation is a function in the sense that it takes an ordered pair of elements from the group and returns an element of the group, but Diane's function was a function of one variable, not two, and hence was not the group operation. Because she seemed to be concerned about doing "mod 8," a better description of her function might be $f(x) = x$ mod 8, but because the domain of her function was the set $\{1, 3, 5, 7\}$, it is impossible to tell the difference between this function and the identity function $f(x) = x$.

Thus, Diane was confusing the concepts of binary operation, homomorphism, and function, all under the vague heading "doing mod 8," an idea that was provided implicitly

either by the binary operation in U_8 or by the construction of its elements. This is the

phenomenon of *implicit homomorphisms* mentioned in the case of Carla in chapter 6 and

discussed further below.

I next asked her to consider the function from U_8 to Z_4 that I had used with all the

participants in their third interview:

34   Brad: Okay. And under this mapping, I want you to send 1 from U_8, I want you to send it over to 0. And I want you to send 3 also to 0, and I want you to send 5 and 7 from U_8, both of them I want you to send to 2.

35   Diane: Using a homomorphism, or just send them over?

36   Brad: Well, send them over. Now I want that to be the mapping, and I want to call it g. Is that a function? Call it little g. How about that?

37   Diane: It's a function only if this is true. If you can send them over using g as a function then it's true.

38   Brad: I am not sure I understand what you mean by that.

39   Diane: Well, U_8 and Z_4 here are.... To bring this over to here is g. So if you can bring this over here, and it's.... I think if you can bring 1 and 3 over to 0, and you bring 5 and 7 to 2. I think g can be a function.

40   Brad: I am not quite sure what it is that you are saying and also what it is that you are worried about. Explain it out loud, as much as you can, what is going through your head.

41   Diane: Well, from here I was saying that U_8 was a function. Mod multiplication 8 was the function. And you are saying that here g is the function. But you have U_8 and Z_4, but I would think that U_8 and Z_4 would be the functions.

In this excerpt, Diane's language is particularly curious and hard to understand. The

explanation lies in her concepts of function, homomorphism, group operation, and the

interaction between them. It became clear later in the interview that Diane's

understanding of function was based largely on a metaphor of "sending over," which, in

this excerpt, forms half of Diane's distinction in line 35. Regarding her concept of

homomorphism, this excerpt reinforces the proposition that, for Diane, the groups

supplied functions, so that where I had suggested a single function g, Diane saw three

functions: U_8, Z_4, and g, which were mod 8, mod 4, and the function that was used to "send elements over."

As Diane considered the function I had defined as g, she said did not know "what the function would be" (line 43), suggesting she wanted it to be given by a formula. She began using the concepts one-to-one and onto to support her reasoning, but she knew she was "fuzzy a little bit on the one-to-one and onto definitions" (line 114).

116   Diane: Well, I am remembering that one-to-one means that each element from G has at most 1 element in G' that it's being mapped to. And then onto means that each element G has at least one element.... No, that's not right; it's the other way around. One-to-one means that G' has at most one element being mapped to it, and onto means at least one being mapped to it. Is that right? Now I am not even sure what the definitions mean. I know one-to-one is at most 1 and onto is at least 1. I am not sure where it starts.

Thus, her concepts of one-to-one and onto were guided by associations with "at most one" and "at least one," respectively, but these associations were not enough to help her determine whether to focus on the domain or the codomain.

When the interview returned to determining whether g was a homomorphism, Diane described how she was thinking about the verification process:

144   Diane: And, well, here you have the function g. I should put g here. And then you take your two elements and put them through g and you get an element here in Z_4. So if you do it on the other side of the homomorphism, you take a here send it through the function and get your element, send b through the function and get your element, and then you put them together in Z_4.

146   Diane: We could take like g and 3 times 1. And in U_8, $3 \times 1 = 3$, and $g(3)$ is 3. And g of.... I don't think I did what I said I'm supposed to do. I think I sent this over to get $g(3)$ in Z_4 is 3 and 3 star 1...

In her description of the process, it appears Diane saw $g(a*b)$ and $g(a)*g(b)$ as two sides of the homomorphism, suggesting that the defining formula for a homomorphism was the homomorphism itself. Furthermore, in evaluating $g(3)$, Diane ignored my definition of g altogether and instead decided that $g(3)$ was 3.

At this point Diane suggested we stop the interview because she was too tired and was

not thinking clearly. I told her that it was entirely up to her but, that I would probably

still find the interview to be useful in my analysis. She chose to continue. We returned

to determining why $g(3) = 3$.

163   Diane: Okay, from here $g(3)$, I took 3 and 1 from U_8, put them together in U_8, sent them over to g, then you take $g(3)$ from here, and that equals 3 because you are in Z_4.

164   Brad: I didn't follow that. I understand that 3×1 is 3 in U_8. But then where is the sending over happening, and where are you writing that? That's what I am not following.

165   Diane: Well, I want $g(3)$, so I am taking the 3, bring it over through g.

166   Brad: Taking the 3 from?

167   Diane: From U_8. Across the function g into Z_4, and so I am now taking 3 from Z_4. And it equals 3.

168   Brad: Okay, because 3 in Z.... Hold off from this for a minute. What is $g(5)$?

169   Diane: First you take the five from U_8. 5 in U_8 is 5. Bring over $g(5)$ into Z_4. So $g(5)$ is 1.

170   Brad: And how did you do that?

171   Diane: You took the element 5 from U_8, and you want to bring it over through g. So $g(5)$ is now in Z_4, and 5 in Z_4 is 1.

172   Brad: Oh, because ...

173   Diane: Because g isn't happening to any element until it passes into Z_4.

174   Brad: And so what does g do?

175   Diane: g maps something over; it sends something over. So g is sending the 5 over. It's not doing anything to the 5 really, it's just sending the 5 over.

176   Brad: And once the 5 gets over there ...

177   Diane: You can do the other operation to it.

Diane then explained that the other operation is "the Z_4 operation" (line 181). Diane had

completely ignored my definition of g and saw g instead as sending elements over from

one group to another. Once an element had been sent over to Z_4, she did the Z_4 operation

to it (i.e., found its remainder, mod 4) to be sure that it was an element of Z_4. It is not

clear from Diane's description, however, whether the 5 becoming 1 was part of g or

something that happened after g had completed its job.

In this episode, Diane's peculiar take on functions and homomorphisms was supported, in part, by the ease with which an element from U_8 could be transformed into an element of Z_4. I then constructed a function h from D_4 to U_8 for which this would be impossible. In response, Diane used my specification without hesitation, saying, for example, "$h(R_0)$, that sends R_0 to 1" (line 209) and later "R_0 is 1" (line 213), suggesting that she had made an identification of elements between the two groups in such a way that h was still doing the sending. When I asked her why she knew that h mapped R_{90} to 3, she responded, "Because you said so" (line 233), implying that she was aware of my role in the specification of h.

I then asked her to think about $g(3)$ from the previous example, and she responded that "g maps 3 to 0" (line 235). I ask her whether that was what she had said before:

239   Diane: Nope. I don't know why this would make more sense. I guess it's because we just did it in class or something. Of course, it shouldn't have made a difference.

246   Diane: That g was just.... I didn't think that g itself was doing anything really. So I thought that if I did this first, that g is actually U_8.

247   Brad: That g is actually U_8? What do you mean by that?

248   Diane: Well, when I said "$g(3)$ is 3."... In U_8, $g(3)$ *is* 3. So I think maybe that's what I was doing.

Thus, at this point in the interview, her thinking had changed, but I still had not reached much clarity on how she was thinking about the function g. I asked her to clarify:

251   Brad: So what are you thinking now g does to ...?

252   Diane: g sends the 3 to 0.

253   Brad: And the 3 is something that exists where?

254   Diane: In U_8.

255   Brad: And the 0 is something that exists ...

256   Diane: In Z_4.

257   Brad: Okay. Then you said $g(3)$.... And whenever you talk about $g(3)$, where is that thing?

258 Diane: Right here in the middle.

259 Brad: What do you mean by that?

260 Diane: I haven't actually…. This by itself, $g(3)$, I haven't exactly done anything to it yet.

Diane imagined three different places where a value might exist: the domain, the codomain, and in the middle. This interpretation becomes clearer in the following explanation:

264 Diane: You take an element in U_8, apply the function g to it, so now you have the function g of something in U_8. So now it's no longer in U_8, because you are applying g to it. [Okay.] Okay? Now as you bring it over, g is not doing anything but mapping it over, just carrying it over. So, if you are stuck—well, you're not going to be stuck—but if you happen to be in the middle, you are not actually doing anything to the element. So it's still going to be g of an element in U_8. And when you bring it all of the way over to Z_4, then you are actually doing something. So $g(3)$ is now 0 because you are actually over in Z_4.

Thus, 3 begins in U_8; $g(3)$ is the process of carrying the 3 over to Z_4, which exists between U_8 and Z_4 but which does not actually do anything until the 3 is brought all of the way over, at which point $g(3)$ becomes 0.

The above episode illustrates three issues, which are treated in different ways in the discussion that follows. Diane's implicit specification of the homomorphism from U_8 to U_8 is discussed below with reference to a similar episode with Carla. Diane's difficulty attaching the terms *one-to-one* and *onto* to aspects of the given functions can be seen as another instance of the issue of naming, which is treated in detail in chapters 5 and 6 regarding other concepts. Her nonstandard interpretation of function notation is discussed next with reference to Lori's understandings and interpretations of functions.

Function Notation

The idea that $g(3)$ denotes a value between the domain and the codomain is a nonstandard interpretation of function notation. Lori demonstrated similar nonstandard interpretations

during her third interview, which took place separately from Diane's. She was working

with the function f from U_8 to Z_4, given by $f(1) = 0$, $f(3) = 0$, $f(5) = 2$, and $f(7) = 2$.

Regarding the notation $f(1)$, she knew that 1 was in U_8. She first claimed that $f(1)$ was in

the codomain because it was 0. Then she changed her mind: "I'm sorry, f(1), and that's

in the domain, ... but then when I take their functional values, $f(1)$ is 0" (line 79).

I asked her again where $f(1)$ was.

83   Lori: $f(1)$. Do you mean $f(1)$ before it's evaluated? Or $f(1)$ after it's evaluated.

84   Brad: Well, explain both.

85   Lori: $f(1)$ is actually in the domain. But then the functional value of $f(1)$ is 0, and that's in the codomain.

Thus, Lori first imagined two senses of $f(1)$: before and after it is evaluated. But then she

made a distinction between $f(1)$, which is in the domain, and its value 0, which is in the

codomain. So at first, the notation was ambiguous, but then she resolved the ambiguity

by making an incorrect notational distinction.

As the interview continued, she decided that "$f(1)$ is like 1" (line 91), which perhaps

prepared her for a different notational distinction. I then asked her to consider another

function g, which I did not specify, and asked her where $g(3)$ would be. She decided

correctly that 3 was in the domain and $g(3)$ was in the codomain and made similar

conclusions about $f(1)$.

I asked her to explain how she had been thinking of it earlier.

115   Lori: Oh, I said that if—it's wrong now—but I said that if, $f(1)$ before it was sent is in the domain, and then after it's sent, it's 0, and it's in the codomain. [Okay.] That's wrong. 1 is in the domain, and $f(1)$ is in the codomain, and $f(1)$ evaluated in the codomain is 0.

121   Lori: 1 is before it's sent, $f(1)$ is after it's sent, and 0 is after it's sent.

Lori's focus on "before and after" suggests that she, too, focused on a process that occurred between the domain and the codomain. For Diane, $f(1)$ was in the middle, whereas for Lori, $f(1)$ was in either the domain or the codomain depending upon whether it was before or after evaluation.

It appears that there is something salient about the process of evaluating a function—something that led Lori to conclude, if only momentarily, that $f(1)$ represented both the before and after of this process. Lori resolved this ambiguity first by placing $f(1)$ in the domain. Diane, in contrast, resolved the ambiguity by creating a new abstract entity that existed "in the middle" between the domain and the codomain.

The idea that $g(3)$ is a value in the middle between the domain and codomain is a reasonable, although incorrect, conclusion to draw from the notation and the metaphors. In standard mathematical usage, when $g(3) = 0$, the 3 is in the domain, and both $g(3)$ and 0 are in the codomain. But by imagining a function as a process, as a machine, or as something that "sends elements over," the process or the traveling will take time, creating a metaphorical need to notate a value in process or in transit. The notation $g(3)$ is an obvious choice because the way it is written seems to suggest that the 3 is still inside the function and has not yet emerged as 0. For Diane, $g(3)$ denoted this value in the middle.

This interpretation of the notation and the metaphors seems so obvious and natural that it seems likely that other students have come to similar conclusions, but the phenomenon has apparently not been recognized in the literature on the concept of function. Yet, the fact that I was able to observe this unusual conception depended, perhaps heavily, on the particular line of questions I asked. Without the particular question in line 257, Diane's unusual thinking would likely have remained hidden.

Implicit Homomorphisms

Just as Diane's unusual interpretation of function notation was found to have similarities in the thinking of other students, her implicit specification of a homomorphism was an issue that arose in Carla's interview as well, as mentioned above. Both Carla and Diane began verifying that a function was a homomorphism before they had specified what the function was. In other words, they were working with functions defined implicitly and did not see a need to be explicit. On the one hand, for Carla, the function was inherent in the construction of Z_4 and was based on her conception of modular arithmetic, as is explored in more detail below. For Diane, on the other hand, the function was tied to mod 4, which she saw as the group operation. In both cases, the result was that their understandings of functions and modular arithmetic led to unexpected outcomes.

Diane's and Carla's conceptions raise serious questions about the prominence given to the canonical homomorphism $f:Z \rightarrow Z_n$ given by $f(x) = x \bmod n$. On Problem 3 on the second take-home exam (Appendix B), the students were asked to show that this function is indeed a homomorphism, and the example was also discussed in class. These events occurred shortly before third interviews, which may explain why the function was prominent in Carla's mind. The issue is that the function does not seem to *do* very much, giving the impression that the function is already part of the codomain and has little to do with a mapping between two groups. A similar statement can be made about the canonical homomorphism $f:Z \rightarrow nZ$ defined by $f(x) = nx$.

Another way to look at these episodes is to observe that Carla and Diane were not able to distinguish between a function between groups, the group operation, and the construction

of the group. Thus, this is yet another example of a need to encourage students to distinguish among ideas that are closely related.

Summary

Regarding the students' concept images of functions, I have presented a few issues that became prominent in the analysis and that had particular relevance for their understanding of elementary group theory. The issues seemed peculiar to Diane at first, but similarities with other students emerged during the analysis.

Diane's idea that $g(3)$ was in the middle between the domain and codomain was a natural, although incorrect, consequence of the notation and the metaphors. Lori similarly had trouble placing $f(1)$, and although she did not suggest that it was between the domain and the codomain, her language suggests that the process of evaluating the function was psychologically salient. The dominance of the process in function evaluation has psychological similarities with the dominance of process conceptions in students' concept images of cosets.

Diane and Carla both specified functions implicitly when providing examples of homomorphisms. The analysis suggests that their concepts of function, homomorphism and binary operation were intertwined and that the binary operations in U_8 and Z_4, or, more simply "mod 8" and "mod 4," provided the functions that they were thinking about. Thus, it is plausible that some of their nonstandard conceptions were related to their understandings of modular arithmetic, which is the focus of the next section.

Modular Arithmetic

Many students used the word *mod* in unexpected ways in the interviews and in class,

suggesting that their concepts might have differed from those in the mathematical

community. Rather than a thorough analysis of students' understanding of modular

arithmetic, I focus on two episodes that provide grounding for the results. The section

begins with a semiotic analysis of Carla's use of the word *mod* during her first interview.

Then I present an episode that occurred during a class early in the course when many

students were still making sense of modular arithmetic. The conceptual analysis is

postponed until after the episodes and incorporated into the analysis in order to give more

of a sense of the phenomena that the analysis is intended to describe.

Carla and mod

All statements in which Carla used the word *mod* were collected for analysis. Focusing,

in particular, on statements in which her language use seemed unusual, I constructed an

explanation that accounts for most of what she said. The series of statements below is

intended to illustrate the evolution of this explanation. In these statements, the word *mod*

is set in bold to draw attention to those phrases that were the subject of analysis. The

series of statements is interrupted periodically with comments and working hypotheses to

illustrate how the conclusion grew out of the data.

12   Carla: Because the addition wasn't a group **mod** *n*. Something about **mod** *n* isn't a group with addition, because multiples of.... Something about multiples of *n*. Let's see, **mod** *n* with addition. All right. For an identity for **mod** *n*, that would be just zero with addition.

61   Carla: *S* is just **mod** *n* under addition.

79   Carla: So the next thing to check would be associativity. But **mod** *n* is a subset of *Z* because all of your elements in **mod** n are integers, and Z under addition is associative, so therefore **mod** n under addition is associative. So therefore **mod** n under addition is a group.

Hypothesis: For Carla, mod n was a set. In Carla's written work, S denoted a set and she identified that set with mod n (line 61). The notion that mod n is a set, which is clear in lines 61 and 79, also fits Carla's language in line 12.

94   Carla: **Mod** 5 means that $a = 5q + r$. **Mod** 5 means that you have elements called a that satisfy the condition $5q + r$. It is kind of a way to back up from the integer group.

96   Carla: Okay, because if you take an integer, you can ... To me it seems almost like it's simplifying because you are dealing with less elements. If you take an integer, say 18, you can write it in a **mod**, and you can express.... Say you were dealing with **mod** 9. 18 is 0 **mod** 9, but 36 is also 0 **mod** 9, so it gives you a way to kind of group and simplify.

Hypothesis: For Carla, mod n was also an operation. Her statement in line 94 also indicates a connection to the division algorithm, but it is not clear how that relates to her other conceptions.

118   Carla: For example, the integer—I mean the inverse of 2 **mod** 3 would be 1/2 **mod** 3, and 1/2 is not an integer, and it is not in **mod** 3, and the only elements of **mod** 3 are 0, 1, 2.

189   Carla: If you are dealing with Z_3 you are supposed to add them in **mod** 3. So what you are saying is that the operation is different because you are dealing with **mod** 3 or **mod** 6. Yeah, I think I remember something being mentioned about that in class.

Note: For Carla, Z_3 was not the same as mod 3.

196   Carla: If you are dealing with **mod** 3, 2, 3, 4 is the same thing as 2, 0, 1 in **mod** 3. So I don't know. I think that it could be a subset because ...

198   Carla: Well, I mean.... Actually, we already know it is a subset, I think that the **mod** 3 and **mod** 6 wouldn't make a difference as far as being considered different operations because it looks, even though the numbers are different, in the portion of the Z_6 table they are equivalent to the Z_3 table if you are dealing with (**mod** 3). Even though I think that something was said in class about **mod** ... 2 **mods**, different **mods** are different operations.

200   Carla: Initially they seem the like different operations, but.... Oh, well actually we weren't debating really whether Z_3 was a subset of Z_6. We are talking about, are their operations different? I don't know if they are different. I'm not sure. I don't think they are because I remember something being said about that they're not. And I have nothing that would make me think that they are the same operation. Because you are doing $nq + r$, but with **mod** 3 you are doing $3q + r$ and in **mod** 6 you are doing $6q + r$, and $3q + r$ is very different from $6q + r$. So I think they are different.

Conclusion: For Carla, mod n was the set that you get from "doing $nq + r$." This explanation accommodates the set, operation, and process interpretations given above.

On the surface, it seems Carla had a rather odd and problematic concept, but on a deeper level, there were some profoundly mathematical aspects of her thinking. First, Carla was able to think about the concept mod n as being both an object (a set) and as a process (doing $nq + r$). Second, the set is a set of remainders, which results from a process that involves stepping back (backing up) from the integers, and organizing them into equivalence classes, which she elsewhere described metaphorically as packages (Interview 1, line 98). Thus, Carla's mod n describes precisely, without using the standard names, the idea, the process, and the object that are behind the symbol $Z/3Z$, which simultaneously denotes all three. The idea in the standard construction is using the subgroup $3Z$ to organize the integers into equivalence classes, the process is creating the cosets, and the object is the group consisting of the three cosets under the operation of coset arithmetic. In other words, if Carla had instead said "$Z/3Z$ is the set you get from forming the quotient (or the group of the cosets) of $3Z$ in Z," we would have been quite impressed. Of course, this language was not available to her at this early point in the class.

I have emphasized that the students' conceptual grids may be different from the accepted one, thus giving rise to concepts that do not fit the accepted grid. But here we have a case where the portions of the grid fit quite well but it is very hard to see the fit through Carla's idiosyncratic language. Her unusual language, in this case, does not seem to indicate a problematic concept. On the other hand, she did seem to focus too much on the division in the process, making her thinking somewhat slow and laborious.

These excerpts also suggest that some learning had taken place. During the interview Carla came to a determination that, when "doing $nq + r$," the value of n is a

distinguishing characteristic of the related group operation: A different value of n indicates a different operation.

Other students and mod

I attempted similar analysis for other students and was not able to come to clear conclusions. Nonetheless, an episode from class provides helpful ideas for the analysis. During the second week of class, after the students had been working for a while on problems involving modular arithmetic, Dr. Benson asked all the collaborative groups to spend a few minutes writing down every fact they could think of that was related to $a = b$ mod n. They had access to two definitions:

> DEFINITION   Modular Equations
> If a and b are integers and n is a positive integer, we write $a = b$ mod n when n divides $a - b$. (Gallian, 1994, p. 8)
>
> We say $a \equiv b$ mod n if a and b have the same remainder when divided by n. (Problem Sheet 1, Appendix C)

The students used both equality and equivalence, with the corresponding symbols, to describe such relationships.

During the discussions in this particular class period, several issues arose. For example, the statement $15 = 7$ mod 4 created disagreement in some groups of students. Some students thought the statement was wrong; others thought that it just had not been simplified all the way. Wendy pointed out, "The whole point is you're getting a remainder." Another student suggested that the idea was to "Take a big problem and make it smaller." More generally, students wondered whether the following statements were equivalent:

$$a = b \bmod n \quad \text{and} \quad b = a \bmod n.$$

For some students, these were not equivalent because $a = b$ mod n meant that a was supposed to be smaller than b and than n; others thought that b was supposed to be smaller than a and n.

These notational difficulties were accompanied by more general algebraic difficulties. Some students, for example, were not sure whether they should write n divides $(a - b)$ or n divides $(b - a)$, not realizing that the statements were equivalent. In thinking about the common remainder interpretation, one group first wrote

$$\frac{a}{n} = q + r \quad \text{and} \quad \frac{b}{n} = q' + r$$

where q and q' were quotients and r was the remainder. After some discussion, they decided that the correct expression was

$$\frac{a}{n} = q + \frac{r}{n} \quad \text{and} \quad \frac{b}{n} = q' + \frac{r}{n}.$$

The groups of students spent some time considering yet not resolving, many of the issues raised by the class. Dr. Benson said,

> Some of you are likely thinking, "Why don't they just tell us?" Well—and I don't mean this in a sarcastic way at all—we did. The point I want to make is that we did tell you and still there were issues that need to be resolved. (Field notes, Jan. 24, 1996)

A student said, "Wait! Problem Sheet 1." Finally, the class summarized the results on the board:

1. $a = b$ mod n
2. $b = a$ mod n
3. n divides $a - b$
4. n divides $b - a$
5. $a - b = nq$ for some integer q
6. $a = b + nk$, where k is an integer

7. a and b have the same remainder when divided by n

8. $\dfrac{a}{n} = q + \dfrac{r}{n}$ and $\dfrac{b}{n} = q' + \dfrac{r}{n}$, with q, q' quotients, r remainder

9. $\dfrac{a-r}{n} = q$ and $\dfrac{b-r}{n} = q'$, with q, q' quotients, r remainder

Note: q, q', and r are integers and n cannot be zero.

Several points can be made here. First, as Dr. Benson pointed out, giving the students a definition was not sufficient for building adequate understanding. They did not even think of consulting a definition until after Dr. Benson reminded them that we had told them.

Second, a deep understanding of the concept of modular arithmetic should include all of these statements and the connections among them. In other words, all of these (and the reasons for their equivalence) are desirable parts of a sophisticated concept image. But these representations all use traditional algebraic symbolism and say little about other representations that might support understanding. Carla, for example, used the metaphor of packages, as mentioned above, to help her think about the equivalence classes. Wendy used the metaphor of the clock to help her think about the arithmetic in Z_n.

Third, although the nine characterizations are algebraically equivalent, they were not psychologically equivalent for these students. Significant thought was required for the students to decide that some of these were equivalent to $a = b \bmod n$. One might hope that the fact that Characterizations 3 and 4 are equivalent would be obvious from the fact that $a - b = -(b - a)$. But for some of the students, this was far from obvious, perhaps indicating insufficient proficiency with high school algebra.

Analysis

Can all of the difficulties be attributed to reluctance to consult the definition and to insufficient proficiency with algebra? Prompted by these events, I conducted a conceptual analysis. What was most surprising in this episode was the students' disagreements over whether Characterizations 1 and 2 were equivalent. The very notion of an equivalence relation implies that it must be symmetric. Just as $a = b$ implies $b = a$, so also $a = b \bmod n$ implies $b = a \bmod n$. What could be the cause of difficulty with something that should be obvious?

I claim the difficulty is caused by ambiguity in the notation. Specifically, the problem is polysemy: related but distinct uses of the word *mod*. It is not often recognized that the symbol *mod* is both a type of equivalence relation and a binary operation. The equivalence relation interpretation is what Dr. Benson had in mind in the episode described above. In other contexts, particularly in computer programming languages,[15] *mod* is a binary operation: an instruction to divide and keep the remainder. In fact, Gallian (1994) makes this explicit:

> DEFINITION $a \bmod n$
> Let n be a fixed positive integer. For any integer a, $a \bmod n$ (sometimes read "a modulo n") is the remainder upon dividing a by n. (p. 7)

This definition is followed by computational examples such as "8 mod 3 = 2."

These two interpretations of the symbol *mod* create ambiguity in the interpretation of the statement $15 = 7 \bmod 4$. If *mod* qualifies an equivalence relation, the statement is true

---

[15] Both Pascal, BASIC and use *mod* as a binary operator. C and C++ use "%" as the symbol for modular arithmetic, so that "17 % 5" evaluates to 2, for example. The mathematical programming languages Maple and ISETL use *mod* as a binary operator. Mathematica and Mathcad, on the other hand, use function notation, so that `Mod[17,5]` and `mod(17,5)`, respectively, are the appropriate ways to calculate 17 mod 5.

because 4 divides $15 - 7$. But if *mod* is a binary operation, then the right side of the equation evaluates to 3 (the remainder after dividing 7 by 4), and the statement is false because $15 \neq 3$.

At least some people in the mathematical community are aware of this ambiguity. The mathematical typesetting language TeX, for example, provides two different commands: "`\bmod` is to be used when 'mod' is a binary operation, ... and `\pmod` is to be used when 'mod' occurs parenthetically at the end of a formula" (Knuth, 1984, p. 164). Some texts (e.g., Fraleigh, 1989) distinguish between these two uses by adopting the convention that Knuth describes, but this convention is not universal (see, e.g., Gallian, 1994). Dr. Benson and I did not use this notational convention consistently, but other results of this study suggest that, even if we had, the students would not have used the convention consistently, at least until they also had made the corresponding conceptual distinction. The conclusion, then, is that both the notational and conceptual distinctions should be made explicit in instruction, perhaps even making connections between the two uses. For example, the statement $a \equiv b \pmod{n}$ can be translated into a binary operation interpretation as follows:

$a \bmod n = b \bmod n$ (in the sense that they both operations give the same result)

Then, with the help of these distinctions, the students might be more likely to see the equivalence of the various formulations of equivalence modulo n, such as:

$a = b \pmod{n}$ if and only if n divides $a - b$
$a = b \pmod{n}$ if and only if a and b have the same remainder when divided by n

This way, the polysemy of the word *mod* might be used to support rather than impede the growth of conceptual understanding.

In summary, this analysis supports three points. First, the distinctions among the several uses of the word *mod* should be made explicit in instruction. Second, the various formulations provide different conceptual support and sources of meaning, and several should be available in instruction. Third, the connections among the various formulations are not obvious and each connection requires some learning. What is unclear is which definition should be introduced first.

In fact, there is a third use of the word *mod*, as in the expression $a = b$ (modulo H), where H is a subgroup of a group. The relationship between equivalence modulo H and equivalence modulo n is accomplished via the following generalization of the standard definition:

$a = b$ (mod n) if and only if n divides $a - b$

$a = b$ (modulo H) if and only if $ab^{-1} \in H$ (or $a - b \in H$, if the group is additive)

It is difficult to see the meaning behind these definitions. Thus, it is worth asking whether there is an alternative definition that would suggest more meaning and simultaneously make strong connections with equivalence modulo n. The conceptual root for equivalence modulo H is the idea of remainders, but remainders cannot be imposed directly on a group and a subgroup, where there may be no division algorithm. An alternative definition may be found via equivalence classes, but because the binary operation carries no notion of equivalence classes, a better route is through the concept of coset.

$a = b$ (modulo H) if and only if a and b lie in the same coset of H (i.e., $aH = bH$)

The results of this study suggest that the above definition should be available to students for the meaning and understanding it might provide. Furthermore, this definition could

provide opportunities for students to deepen their understanding of modular arithmetic, cosets, and the connections between them. Again, however, it is not clear whether this definition should be introduced before or after the standard definition.

Summary

This section explains some nonstandard uses of the word *mod* that occurred during Carla's interview and during a class early in the course. For Carla, mod n was the set the you get from "doing $nq + r$." Conceptually, Carla's nonstandard usage seems to carry the same meaning as the more standard statement that Z_n is the set you get from calculating the quotient group Z/nZ and choosing representative elements. The class had disagreements about the correctness of several formulations of $a = b \pmod{n}$. Some of the difficulties that Carla and other students had with the term *mod* are explained by the fact that the word is used ambiguously, both as a binary operation and as a type of equivalence. It is suggested that the various uses of the word *mod*, the various formulations of equivalence modulo n and equivalence modulo H, and the connections among them should be explicit in instruction.

Conclusion

This chapter set out to describe the students' understandings of preliminary mathematics as those understandings came into play in their learning of group theory. The students' understanding of preliminary mathematics was not a specific focus of any of the interviews. Nonetheless, there were episodes demonstrating that the concepts of function and modular arithmetic are crucial and that sometimes the students' nonstandard conceptions appeared to obstruct their progress on tasks and concepts in abstract algebra. Regarding the key concepts of function and modular arithmetic, the analysis shows that,

just as in the previous chapters, students' understandings are intimately tied to issues of language, notation, and metaphor. Furthermore, the analysis suggests that students' understandings may be strongly influenced by the particular examples and particular definitions chosen.

CHAPTER VIII

CONCLUSIONS AND IMPLICATIONS

This exploratory study sought to describe students' understanding in abstract algebra in the context of an undergraduate course. Using Tall and Vinner's (1981) notion of a concept image, which is the entire cognitive structure associated with a concept, the study identified prominent characteristics and components of students' concept images for central concepts in group theory, up to and including the concept of quotient group.

The setting for the study was an abstract algebra class for mathematics majors, covering many of the standard topics from group theory but in which lectures were replaced by collaborative and individual work on problem sets designed to promote connections between students' prior and emerging understandings and the concepts of group theory. The analysis and the ensuing results are based largely on interview data with five students from the class. Other data sources, such as field notes and students' written exams, provided corroborating and contrasting evidence.

The research questions were as follows:

- What are the prominent characteristics and components of students' concept images as they are learning the fundamental ideas of group, subgroup, and isomorphism?

- What are the prominent characteristics and components of students' concept images as they are learning the more advanced ideas of homomorphism, coset, and quotient group?

- How do students' understandings of prior mathematics come into play as they are learning elementary group theory?

332

Detailed results for each of these questions, organized according to mathematical concept, are found in chapters 5-7. This chapter presents a synthesis of the findings, along with conclusions and implications.

What was initially most salient in the data and the results was that the students used and interpreted language and representations in nonstandard and unexpected ways that were not readily explained by available theoretical and conceptual perspectives on advanced mathematical thinking. On the conviction that the students' understandings were reflected in their language and actions, I endeavored to describe and explain the meaning behind their utterances and their use of representations. Through these efforts, characterizing students' concept images became a process of theory generation. Specifically, through an analysis of language, the goal was to understand students' representations and to represent their understandings. In this sense, this study was about the interplay among mathematics, language, and representations.

The results of the study derive from three types of analysis: detailed analysis of the interview transcripts, global analysis of the students' use of words and notations, and my own conceptual analysis of the mathematical content. The methods of analysis emerged through the analysis itself, and the research questions evolved as part of the process. The detailed analyses generated preliminary hypotheses that were refined through continuing analysis and synthesis. The global analyses involved searching the data for words and notations to confirm and refute the emerging hypotheses. The conceptual analyses served to make explicit the ways in which the students' use of language was compared with standard usage in the mathematical community. These analyses were performed

iteratively so that emerging results could continually inform other types of analysis. The analysis and synthesis produced two main findings.

The first finding concerns issues of language and notation; it is described below as making the vague more precise. In short, language use that at first had seemed idiosyncratic and ambiguous was found to have common threads across students. The students confused related words and had trouble attaching names to their experiences in standard ways. This finding supports the position that attaching names to experience is not simply a process of gluing labels to pre-existing, "self-identifying" concepts but first requires cutting up experience and organizing it into concepts. Students do not necessarily make the same distinctions as those made by mathematicians and mathematics educators, and thus they cut up experiences in different ways, both indicating and further establishing a collection of concepts that are substantially and structurally different from concepts that are used in the mathematical community.

The second finding has to do with issues of representation and abstraction; it is elaborated below as making the abstract more concrete. In order to gain access to abstract ideas, the students relied on representations, metaphors, and other conceptual supports in order to manage their relationships with unfamiliar abstractions. Representations both provided and obstructed the students' access to abstract mathematical ideas and thus both supported and constrained their understandings. The result is well illustrated by the phenomenon I call *reasoning from the table*, in which the group operation table serves metaphorically as the group, supporting students' thinking and reasoning. The operation table is a representation that mediates abstraction, giving students access to and ways to think about abstract ideas but sometimes also impeding their progress toward an abstract view.

The conclusions continue with the suggestion that the two findings are distinct and fundamental aspects of mathematical activity. Learning advanced mathematics is a matter of making distinctions and managing abstraction, and language and representations are the tools. Mathematics is a complex interplay between logic and intuition, between precision and initially vague abstractions.

Following the synthesis of the findings, the chapter discusses some general implications for mathematics teaching and teacher education. The central theme is the importance of encouraging students to make their thinking explicit as a way to build their understanding and also to identify the distinctions and abstractions they are and are not making. The chapter closes with a discussion of implications for theoretical and empirical work in mathematics education.

Making the Vague More Precise

Learning advanced mathematics involves learning concepts, processes, language, notation, and the relations among them. That learning can be uneven, and what is learned can be connected (and disconnected) in surprising ways. A main finding of the study was the seeming independence between the students' ideas and the language of abstract algebra. Sometimes the students' nonstandard language was close to standard usage, as when they interchanged two closely related terms: using *range* for *codomain, associativity* for *commutativity,* or *identity* for *inverse.* At other times the students used a term more broadly than was appropriate—"the left coset" for *the set of cosets*—thereby introducing apparent ambiguity into their language.

In some cases, such ambiguity may lie mostly in a superficial interpretation of the students' utterances, for they actually used their idiosyncratic language precisely and consistently, as when one student used the phrase *normal group* for the quotient group that may be constructed when a subgroup is normal. Some students resisted attempts to impose standard usage, particularly when it seemed to contradict their own usage. Other students, in contrast, seemed to be much less precise in their use of language and also less bothered by ambiguity.

Of course, the use of standard language is not necessarily an indication of understanding. For example, although many of the students stated on the final exam that a subgroup is (or is not) normal because the left and right cosets are (or are not) the same, some of those same students did not compute the cosets correctly.

Mathematical discourse depends for its effectiveness on subtle distinctions in notation and syntax that are established by convention. These distinctions, however, are not necessarily apparent to students. Furthermore, mathematical discourse is not without its own ambiguity. The word *mod*, for example, is used both as a binary operator and as a modifier of an equivalence relation, thereby creating ambiguity in statements such as $15 = 7 \mod 4$.

One way in which teachers and researchers can deal with such problems of ambiguity may be to focus explicitly on linguistic, notational, and conceptual distinctions, probing beneath the surface whenever possible. This idea is elaborated below along with other implications. In this section, I elaborate the linguistic, notational, and conceptual issues that arose in this study.

Naming Concepts

Although lacking the rigor and specificity of a true instructional method, the "discovery method" exists, in name at least, in the literature and in some mathematics classrooms (see, e.g., R. B. Davis, 1990; Dean, 1996; Mahavier, 1997; Morriss, 1998; Touval, 1997). In fact, the instructor for the class that provided the setting for this study characterized the class this way (Benson, in press). A description of the discovery method would likely include the following features: Give the students a rich problem situation to explore. They will discover patterns and relationships, develop ideas and concepts, and create objects and processes. Then simply give the students the commonly accepted terminology, and with some metaphorical glue (Hewitt, 2001), they will attach standard names to established objects or properties unproblematically. Leron and Dubinsky (1995), for example, suggest that "except for the new name, the students can really feel that the instructor merely summarizes what they have found in their investigations" (p. 238).

This study showed that the final step of naming is not necessarily routine and unproblematic. Learning mathematical vocabulary and its appropriate syntax is sometimes a complicated process with much potential for a misstep. What might explain the difficulties students have with the seemingly trivial process of attaching a name to an idea? I identified three kinds of naming difficulties, each with its own explanation.

The first kind of naming difficulty is that two words are sometimes confused when they involve closely related ideas. A person says one thing but means another, as when the students in this study swapped *identity* for *inverse, commutativity* for *associativity, multiplication* for *addition,* and *range* for *codomain.* Sometimes such errors are mere

slips of the tongue; at other times the ideas themselves have become somewhat muddled. In either case, these results are consistent with the observation from linguistics and cognitive science that such whole-word substitutions occur when words are semantically related (Hotopf, 1980; Stillings et al., 1995). Cognitive science provides an explanation that fits with its models of long-term memory: Metaphorically, closely related words are stored in close proximity and thus are sometimes confused and hard to keep apart (Stillings et al., 1995). This study demonstrates that even when the boundaries between related concepts are relatively easy to draw, such as between the identity and inverse properties, the distinction between the corresponding words is sometimes hard to manage. When the boundaries are harder to draw, the distinctions can become quite problematic.

The second kind of naming difficulty seems to be caused by the name itself. In such cases, learning the name requires building some cognitive structure around the name to support its meaning and use. Sometimes, as in the terms *cycle* and *identity,* the name carries an everyday meaning that is somewhat different from the mathematical meaning (see, e.g., Lajoie & Mura, 2000; Pimm, 1987). Other terms have multiple meanings within mathematical discourse (Zazkis, 1998; Durkin & Shire, 1991). In this study, such polysemy was noted for the words *mod* and *congruent,* but only the former seemed to cause difficulty. With still other terms, the name carries content that begs explanation. One student experienced just such a difficulty with the term *isomorphism* when interviewed early in the course. Perhaps the students' difficulties with the term *quotient group* may be explained similarly.

The first two kinds of naming difficulty are not particularly novel; both appear in the literature. The third is rarely recognized, however, at least in the mathematics education literature. Furthermore, it is of a different character from the other two because it is more epistemological than cognitive. The difficulty is as follows: The linguistic, notational, and conceptual distinctions that students make are not necessarily the same as those made by mathematicians and mathematics teachers. Some of the distinctions that teachers make are neither apparent nor relevant to their students. And students make some distinctions that their teachers do not make. Making distinctions, delineating concepts, and assigning names are, as Foucault (1971) noted, a matter of imposing order on experience:

> Order is, at one and the same time, that which is given in things as their inner law, the hidden network that determines the way they confront one another, and also that which has no existence except in the grid created by a glance, an examination, a language; and it is only in the blank spaces of this grid that order manifests itself in depth as though already there, waiting in silence for the moment of its expression. (pp. xix-xx)

Foucault came to this conclusion through an historical analysis of two great discontinuities in the nature of knowledge in the seventeenth and early nineteenth centuries, yet this perspective seems particularly apt for describing many of the discrepancies observed in this study between students' language and what I take to be standard mathematical usage.

In this study, the students used grids that did not fit with standard mathematical discourse. Some students, for example, did not make clear linguistic distinctions between a particular coset and the set of all cosets. To her, these objects and the process that connected them were all part of a single concept. In contrast, another student, discussing the meaning of $g(3) = 0$, not only distinguished between the value 3 in the domain and

the value 0 in the codomain but also saw $g(3)$ as a value that was "in the middle." Thus, in situations where a mathematician might see two ideas, the first student saw one, and the second saw three.

A problem in naming that involved several related concepts occurred when one student expressed many of the key ideas about normality and quotient groups but had not yet attached the standard names to those ideas. When the left and right cosets were the same, all the students knew that the word *normal* applied in some way. But what was it that was normal? The resulting group, the subgroup, the generator of the subgroup, or the cosets? And from the symbolic equation $aH = Ha$, there was also a sense that the word *Abelian* should apply.

In developing language to describe a particular area of study, constructing definitions and meaning is a matter of carving the area into a collection of related concepts, which requires imposing a structure on the area of study and a grid on the various activities. This imposition is, in principle, arbitrary, although it is guided by historical and cultural precedent and by Occam's razor, a principle that has its roots in Plato's suggestion that we carve the universe at its joints (Plato, 1998, 265e) and that was eloquently interpreted by Einstein: "Make your theory as simple as possible, but no simpler." When students try to make sense of an activity and to construct meaning for the various words, they, too, impose a grid on the activity, but constructing the standard grid requires that they see the joints that are implicit in the standard distinctions, some of which are historical accidents. The results of this study indicate that students do not necessarily see such joints and thus do not use the standard conceptual grid.

Returning to the issue that opened this section, the results of this study imply that the discovery method underestimates the cognitive complexity of naming and, more specifically, that the metaphor of gluing names to ideas is too simplistic. The gluing metaphor implies that the experience is cut up into concepts that are naturally and unproblematically identified and, thus, that naming is a matter of assimilating standard names into well-defined cognitive structures that have been created by the experience. This study suggests, in contrast, that cognitive structures are still being created during the process of naming. I am claiming not that the process of naming always requires building new cognitive structures but rather that sometimes assimilation is insufficient to describe what is involved in learning to use the standard language. And in any case, the resulting concepts may not fit with those of a teacher or researcher, not to mention the possibility of problems of fit among teachers and researchers.

During several interviews, after I had developed a good sense of the student's nonstandard language, I intervened and tried to encourage the student to use standard linguistic conventions and distinctions. In general, these interventions were not successful until the student had made the corresponding conceptual distinctions and thus had developed a psychological a need for the standard linguistic or notational conventions. Even then, learning standard usage seemed to depend upon a dialectic between the conceptual and the linguistic distinctions. These results suggest that learning new language and particularly changing one's use of language may be better seen as accommodation rather than assimilation. Therefore, it is necessary, at the very least, to pay attention to the distinctions that students make and to make standard distinctions explicit in instruction. These ideas are explored further below.

The three naming difficulties described here have distinct explanations. In practice, however, they are by no means distinct but rather operate simultaneously. After one carves one's experience into concepts, the concepts and words may be very closely related and hence difficult to keep separate. At the same time, the available words may carry other meanings or may present other challenges.

> I am not saying that language creates reality. Far from it. Rather, I am saying that what counts as reality—what counts as a glass of water or a book or a table, what counts as the same glass or a different book or two tables—is a matter of the categories that we impose on the world; and those categories are for the most part linguistic. And furthermore, when we experience the world, we experience it through linguistic categories that help to shape the experiences themselves. The world doesn't come to us already sliced up into objects and experiences; what counts as an object is already a function of our system of representation, and how we perceive the world in our experiences is influenced by that system of representation. The mistake is to suppose that the application of language to the world consists of attaching labels to objects that are, so to speak, self-identifying. On my view, the world divides the way we divide it, and our main way of dividing things up is in language. (Searle, quoted in Magee, 1979, p. 184)

Searle's point, along with the results of this study, put a subtle spin on linguistic determinism. Language is simultaneously a social and a personal construction. The language that an individual experiences both influences and constrains the reality that the individual creates. The language that an individual uses influences and constrains that reality more fundamentally. These two languages are unlikely to match and might not even fit, however, for the individual and the community do not necessarily divide things up in the same ways, and thus their meanings can be substantively different.

Learning Notation

As with names, students' use of notation does not necessarily incorporate the same distinctions as in standard mathematical discourse. Thus, much of the above discussion about names applies to students' notation as well. In particular, the students in this study did not always distinguish between sets and elements, between the two different notations

for permutations, between distinct elements in the same set, and between variables and names of elements.

These issues are examined further below as part of the discussion of the ways that notation can help manage abstraction. Here I point out only that many mathematical distinctions are maintained by notational conventions. Sets, for example, are denoted by uppercase letters, elements by lowercase letters. Functions are f and g; variables are x and y. The identity element in a group is e. Experience suggests that these conventions support thinking in the sense that it takes some psychological adjustment to think about a function $x(f)$ or, as demonstrated in this study, to recognize that a letter other than e might represent the identity. The psychological support provided by these conventions is not often discussed in the literature. Yet, students do not necessarily adopt these conventions, making it difficult sometimes to interpret their work. And even when they do adopt the conventions, one cannot necessarily conclude that they have made the corresponding conceptual distinctions and abstractions.

Using Definitions

The students' definitions in this class were a blend of formality and informality. Although the nature and role of mathematical definitions was not an explicit focus of the class, definitions were periodically introduced, and the instructor and I regularly worked with students, both individually and collectively, to help them get better at using definitions. Thus, I took a broad view of definitions in the analysis for this study, allowing both formal definitions that the students provided on exams as well as informal statements that they provided when asked what a word meant. This approach was

intended to increase the possibility of insight into the meanings that the students imagined words to hold.

The students' formal definitions often lacked quantifiers and were otherwise imprecise. Their informal definitions were of varying degrees of correctness and bore varying degrees of resemblance to the formal definitions.

Some of the students' definitions were informal and either vague or missing important features:

- *Subgroup* means a subset that's a group.
- Z_6 means mod 6
- Z_6 means $\{0, 1, 2, 3, 4, 5\}$
- *Isomorphism* means congruent, same form (with renaming and reordering).
- *Identity* means the do-nothing element.

Some of their definitions were informal but largely correct and potentially supportive of a correct formal version:

- *Group* means it's associative, it has an identity, it has an inverse, and it is closed.
- *Kernel* means the elements that are mapped to the identity.
- *Homomorphism* means a function that preserves the operation.
- *Coset* means aK or aH.
- *Normal* means the left and right cosets are the same.
- *Quotient group* is the group of the cosets when the left and right cosets are the same.

Other definitions were similar to formal definitions but were missing quantifiers or specification of notations:

- *Associative* means $(a*b)*c = a*(b*c)$.
- *Homomorphism* means $f(a*b) = f(a)*'f(b)$.
- *Closure* means for all a and b in G, $a*b$ is also in G.

The students had trouble stating some of the formal definitions, particularly regarding the use of quantifiers, but their informal definitions served them well in much of their work and reasoning.

Because informal definitions were often helpful in guiding the students' intuitions, it seems counterproductive to suggest that their informal characterizations should have been discarded in favor of more precise formal versions. Further, as I argue below, both precision and vague intuition are necessary in mathematical activity.

It is worth exploring whether potential confusion arising from the absence of quantifiers and other imprecision can be reduced through instruction that makes explicit connections between carefully chosen informal definitions and the associated symbolic expressions. For example, some informal definitions have quantifiers built in:

- A function is a homomorphism if it preserves the group operation.

Many of the students not only could state this informal definition but also associated the word *homomorphism* with the formula $f(a*b) = f(a)*'f(b)$. Articulating the formal definition, then, is a matter of connecting the informal idea with this formula in ways that make the quantifiers explicit. This seems a promising approach because it acknowledges the importance of intuition as well as the need for precision.

Generalization Versus Distinction

This study has shown through many examples that students cut up experience in ways that do not necessarily fit with standard mathematical discourse, leading to surprisingly novel or confounded concepts and to unconventional use of language and notation. This phenomenon is fundamentally about making distinctions and generalizations, which

involves paying attention to some differences and ignoring others. In this way, making distinctions is the opposite of generalization.

Sometimes the students were too general in their use of language and notation and were insufficiently precise, failing to make important distinctions. Some, for example, inappropriately generalized the notion that associativity is a "global" property—a property that applies to any subset when it applies to a group as a whole—and concluded that addition in Z_n is associative because Z_n is a subgroup of Z. Preventing or overcoming this generalization requires making a distinction between addition in Z and addition in Z_n. Conversely, at other times, the students were not general enough, making unexpected restrictions or distinctions, such as in considering $g(3)$ to be "in the middle" between the domain and the codomain.

It would be disingenuous and unproductive to suggest that students need to pay better attention to differences, because many differences in notation and language are not significant. Sometimes small differences in a signifier denote large differences in the signified, as in the common convention of denoting sets by uppercase letters and elements by lowercase letters. Conversely, large differences in the signifier sometimes denote small differences in the signified. This is, after all, the idea behind the concept of isomorphism: noticing that two seemingly different representations are essentially the same and thus may be considered to represent the same abstract object. The problem is knowing which differences merit attention.

Making such distinctions is a problem both for students trying to learn mathematics and for teachers and researchers trying to understand and analyze students' thinking. Only by paying careful attention to language was I able to learn that for at least one student $g(3)$

was "in the middle" between the domain and the codomain. Similarly, only by paying attention to the subtle distinctions between the words *coset* and *cosets* was I able to conclude that some students' concepts of coset were categorically different from the standard concept, in the sense that they did not adequately distinguish between a particular coset and a set of cosets. Simultaneously, however, I had a sense that other students' concepts of coset were not general enough, because they insisted that the process of calculating cosets must begin with the kernel of a homomorphism. Generalizing the coset idea to subgroups required letting go of something (the kernel) that seemed central to the process as it had been introduced.

Making the Abstract More Concrete

The students in this study demonstrated at least three ways in which they managed abstraction: metaphor, reification, and increasing proficiency. The students gained access to groups via operation tables, using the operation tables metaphorically to support their thinking about the groups. They gained access to cosets, quotient groups, and properties of binary operations by focusing on the processes associated with these concepts. Eventually, through increasing familiarity, some of these processes were reified as objects. Also through familiarity, abstract objects and properties became more concrete as the students developed proficiency with the concepts, procedures, and examples, and gained a better sense of what to expect. These strategies are discussed in turn in the sections that follow.

Philosophically, it is possible to create abstract objects, properties, or categories by using a definition, a description, a process, or a representation, or by noticing a pattern or

common properties. Notationally, it is possible to discuss abstract categories by choosing a generic member of the category. Psychologically, however, it seems there is much more involved.

Many of the results in this study can be viewed as examples of ways student reduce the abstraction level (Hazzan, 1999) of a concept. But what do strategies such as metaphor and reification imply about the psychological processes called *abstraction*? Frorer, Hazzan, and Manes (1997) assert that are three different kinds of abstraction. Sometimes abstraction is about ignoring the pesky details. At other times, it involves thinking in terms of properties. At still other times, it is about one's relationship to an idea (see also Wilensky, 1991). This last kind of abstraction is most helpful in explaining the results of this study. The first two kinds seem to require enough familiarity to be able to imagine generic objects that have particular properties or whose details can be ignored. This problem of imagining generic objects is elaborated below in the section about managing abstraction and trying to be general.

Operation Table as Metaphor

Operation tables served to mediate abstraction for the students in this study in that they worked with a concrete representation to gain access to abstract objects and their properties. A group's operation table makes the group more concrete by making aspects of its form directly visible. Furthermore, by squinting one's eyes or coloring the operation tables by elements or by cosets, the abstract group—of which the particular table is an instantiation—can almost become visible. Abstracting the essence of a group from an instantiation seems a quintessential example of an activity that requires reflective abstraction—abstraction based on action (i.e., operations) alone. With the help of the

operation table, however, perhaps only empirical abstraction is required. Thus, the table becomes both a tool for reasoning and an object of reflection. Under Wilensky's (1991) view of abstractness as a measure of one's familiarity with a situation, the table serves to increase one's familiarity, thereby making the abstract more concrete.

Operation tables served a metaphorical role for many students in that the tables supported their reasoning and helped them think of groups as objects. For some of the students in the study, the table *was* the group rather than a representation—a metonymic substitution of the concrete for the abstract. Their reasoning seemed to be largely external, in the sense that it was based *in* the table and in procedures that required that the operation table be present rather than in reflection on the binary operation. The cancellation laws (e.g., $ab = ac$ implies $b = c$), for example, became embodied in the requirement that each element appears exactly once in any row or column. One student went so far as to describe the geometry of the table, suggesting that for him the table was a geometric object with geometric properties such as symmetry. The geometric object was a metaphor that supported the class's use of the word *congruent* in discussing group isomorphisms. In fact, a congruence between two geometric figures requires specifying a correspondence between parts, just as a group isomorphism requires specifying a correspondence between the elements of groups.

The table served also to heighten the students' sense of anticipation about the way the calculations should turn out, similar to Boero's (1993) observation about the role of anticipation in algebraic manipulation. One student, for example, expected {5, 7} to be its own inverse. Many students came to expect certain patterns in their operation tables

and likened those patterns to cycles, which caused some potentially problematic connections with the cycle representations of permutations.

The students in the study could see isomorphisms by looking at operation tables. Some of them were especially drawn to the squares of elements in the group, as the squares appeared on the diagonal of the table. They were able to see important differences between groups on the basis of the number of elements appearing on the diagonal.

The table as metaphor is not without its limitations, however. First, it becomes cumbersome for large groups, and extending the metaphor to infinite groups requires some sophisticated patterning abilities since it is not possible to write out the whole group table. Second, the students expected subgroups to occupy a corner of the table, probably because of an overly literal Groups-Are-Containers metaphor. Third, writing down a group table requires one to choose an order for the elements, which sometimes made it difficult for the students to recognize isomorphisms and to think of the order as nonessential. Nonetheless, through experiences in renaming and reordering operation tables, the students began to separate from the table from the group—the signifier from the signified—and thus began to develop concepts of abstract groups.

The results of this study suggest that the operation table can play a useful metaphorical role in students' thinking about group theory because of the conceptual support that the metaphor can provide. Still, it is important to make the metaphors explicit and to be aware of their limitations.

Processes and Objects

In the literature, the process/object distinction is usually portrayed as developmental and hierarchical, with an object conception being the more sophisticated (see, e.g., Dubinsky et al., 1994; Sfard & Linchevski, 1994). The present study suggests, however, that this portrayal may be too simple. There were three concepts for which the process/object distinction was particularly relevant: group, coset, and quotient group. This study demonstrated that in all three cases that the emergence of an object conception was not necessarily an indication of a well-developed process conception.

Regarding the concept of group, all the key participants in this study demonstrated a strong object conception, often based in the operation table. Not all of them, however, demonstrated a strong process conception. When the students focused on the operation table, they took it to be an object—a whole with pattern and symmetry. Furthermore, subgroups were imagined as portions of the table—as subobjects in a sense—although this view was sometimes constrained by the arrangement of the elements in the table. When the students focused on the processes involved in carrying out the operation, they were able to see that addition mod 6 is different from addition mod 3, although this distinction was sometimes overwhelmed by the sense that "addition is addition." The data give the impression that the process conception was more powerful and more sophisticated. Was the developmental trajectory reversed in this case?

One possible explanation is that the tables were not objects but pseudo-objects for these students (Sfard & Linchevski, 1994; Zandieh, 2000). When they were focusing on the table, there was a noticeable "externality" in their relationship to the group and its operation in that they sometimes needed to see the whole table before they could reason

clearly, as mentioned above. Finding the inverse of a in Z_n, for example, requires looking for 0, the identity, in the a row of the table. This process is easier to carry out when the a row is complete. Those students whose reasoning was based on the operation, on the other hand, were able to find the inverse of an element without consulting the table, sometimes even coming to the general conclusion that the inverse of a would be $n - a$.

Regarding the concepts of coset and quotient group, the results of this study are consistent with the finding in the literature that in order to construct quotient groups, one must be able to conceive of a coset as an object (Dubinsky et al., 1994). When the students were able to compute quotient groups, they seemed to consider cosets to be objects. Contrary to the literature, however, conceiving of cosets as objects was not very problematic for these students, even for the student who found the process of constructing cosets difficult to manage. The students were perfectly happy to talk about sets of cosets and about a binary operation on cosets, both of which might be taken as hallmarks of object conceptions. They found it relatively easy to compute quotient groups in both Abelian and non-Abelian cases and seemed to see the calculations as rather natural. Furthermore, after completing such computations and organizing them in an operation table, they saw the quotient group, and hence the set of cosets, as being an object, probably because their object conception of groups was supported by the operation table.

Nonetheless, as discussed above, many students found it quite difficult to use standard language to describe what they were doing. Moreover, regarding the concept of coset, there was a sense in which some students were stuck in the process, failing to distinguish either notationally or linguistically between a particular coset and the set of all cosets. For these students, the notation aH signified the process for calculating the cosets.

In short, the students in this study conceived of cosets as objects but did not necessarily think of *aH* as representing one such object; they saw the operation table for a quotient group as being an object, but they had trouble calling it a quotient group. This result adds a subtle distinction to the previous finding that students have trouble seeing cosets as objects (Dubinsky et al., 1994). It is possible that this study yielded a different result partly because the analysis, with its particular attention to language and notation, was more sensitive to such a distinction. It is also plausible to conclude that the different result was caused partly by the introduction of "arithmetic of sets" early in the course. By the time the students were asked to try to create groups out of sets of cosets, performing an operation on two cosets seemed completely natural and unproblematic for many of them. By actually computing products of cosets, the students were able to see that it was desirable that the product of cosets be another coset. They also saw cosets as elements of a new structure, although most of them did not call it a quotient group during their interviews. Leron and Dubinsky (1995) suggest that the computers can support students' calculations with cosets so that they may begin to see cosets as objects (p. 240). I agree and would add that this study suggests that hand calculation is also beneficial. Taken together these two approaches suggest that proficiency with procedures helps turn vague ideas into objects. That is, both approaches support reification.

The inclusion of set arithmetic in the course may be supported for mathematical reasons as well: Coset multiplication is a special case of set arithmetic that has applications throughout abstract algebra. Thus, it makes sense to use these ideas to support each other rather than to keep them separate. Furthermore, by spending some time computing coset products that are not again cosets, students gain some experience with the mess that is

created when a subgroup is not normal. They have an opportunity to appreciate the usefulness of the concept of normality and to explore the concept of closure in an unfamiliar situation. At the same time, they may begin to look for the kinds of regularities that support normality and their relationship with the overall group structure. In other words, set arithmetic in general and coset arithmetic in particular can provide an experiential base on which to build an understanding of the more formal aspects of abstract algebra.

What, then, do the results of this study say about the process/object distinction? Is the process/object distinction truly developmental? One might argue in support of the developmental process/object distinction by stipulating that a student's thinking should be classified not as an object conception but as a pseudo-object conception unless it is an encapsulation of a process, conceived with generality and fullness. This solution seems problematic, however, because it guarantees that the distinction is developmental, in the sense that "object conception" would really mean "object and process conception," which would provide no theoretical room for object conceptions that are weakly supported by the underlying processes.

Alternatively, one might acknowledge that the process/object distinction is too blunt an instrument. After all, even for a specific concept, not all encapsulated processes are the same. Instead, researchers could try to characterize various kinds of process conceptions and also various kinds of object conceptions. Given the results of this study, however, this approach seems the more promising one, although it implies that there is considerable theoretical work to be done.

I do not have a satisfactory solution for this dilemma. I do suggest, however, that researchers remain cautious about attributing object conceptions to students on the basis of object-like language and, more importantly, about making developmental claims on the basis of such language. Certainly an object and process conception is a desirable goal for many mathematical topics, independent of the developmental trajectories that might be followed along the way.

Concept Proficiency

Early in the course, the students were tied to particular representations, such as the operation table, and tied to the processes. As mentioned above, the students' concepts became more abstract and more flexible as they moved away from the operation tables and began to think of processes as objects. These changes demonstrate the usefulness of metaphor and reification in managing abstraction. These phenomena alone, however, do not account for another prominent method of managing abstraction, namely, through increased proficiency with a concept. Again, by *proficiency*, I mean not only fluency with the procedures but also the ability and disposition to use understanding to reason about and solve problems with the concepts (see Kilpatrick et al., 2001).

This characterization was particularly noticeable in the phenomenon I called *operation confusion*. In the early interviews, some students were unsure of the operations in Z_3 and Z_6 and spent much of their time determining the operation. In later interviews, they were still sometimes unsure of the operation, but, in contrast, they were able to determine the operation quickly by relying on the group axioms and familiarity with the elements in the group.

As the students became more familiar with concepts and associated representations, processes, and examples, their concept images became richer and more flexible and efficient. Their work was more often guided by correct expectations, and even when their expectations were incorrect, they were increasingly able to notice errors and resolve inconsistencies by relying on multiple ways of thinking about the concepts. If abstractness is regarded not as property of a concept but rather as one's relationship to the concept (Wilensky, 1991; Frorer et al., 1997), then increasing proficiency is a way to reduce abstractness.

Balancing Precision and Abstraction

Thus far, the results of this study have been discussed under two themes: issues of language and notation and issues of abstraction. Elaborating and synthesizing these themes has led to the theoretical proposition that the themes are not separate but are fundamentally intertwined in mathematical thinking and learning. On the one hand, precise definitions, notation, and language are necessary for the precise thinking that allows careful distinctions to be made among objects. On the other hand, there is a fundamental human cognitive need to reduce the operative models—to abstract and generalize so that many diverse phenomena can be particular instantiations of a single idea. The results of this study suggest that mathematical activity might be considered a carefully orchestrated balance between two opposing tendencies: making distinctions among things that seem the same and blurring the distinctions among things that seem different. In other words, mathematical insight occurs not only when we realize things we thought were different are the same, but also when we realize that things we thought were the same are different.

Advanced mathematical thinking includes not only precise definitions and logical deduction (Tall, 1992) but also significant abstraction and generalization (Dreyfus, 1991). Although teachers expect students to reason with sufficient precision and with sufficient generality and abstraction, in fact these are opposing expectations. Being sufficiently precise requires using rigorous definitions and careful notation to maintain important distinctions. On the one hand, some of the students in this study did not make important distinctions between set and element and between addition in Z and addition in Z_6. Being sufficiently general and abstract, on the other hand, requires blurring distinctions among things that were thought to be different. Building a deep understanding of the group Z_6, for example, requires that -1 be considered the same as 5. The very concept of isomorphism is about blurring representational distinctions in order to gain access to abstract mathematical objects that "lie behind" all of them.

Mathematical learning involves building intuition, creating mathematical objects, and making distinctions among them. Each introduction of a name, symbol, or definition, raises the possibility of making unusual abstractions and nonstandard generalizations and distinctions, for the standard distinctions are not pre-existing in the world of mathematical objects (wherever that is). Clearly there are more and less effective ways of maintaining the balance between precision and abstraction, but the best solutions are rarely clear a priori. Thus, students should be expected to make unusual distinctions regularly as part of the learning process.

My argument about the balance between making and blurring distinctions echoes Poincaré's (1946) observation that logic and intuition play complementary roles in mathematics. Guided by intuition alone, one defines mathematical objects vaguely.

Intuition cannot bring certainty and can even deceive, so rigor is necessary. Rigor requires logic and begins with definitions, but logic can create nothing new. Thus, logic and intuition are each indispensable.

Building on Poincaré's argument, logic is about precision, making careful distinctions, and reasoning from rigorous definitions and axioms. But formal definitions alone are empty. Instead, there must be something to formalize, some intuition that is being made precise. Definitions themselves do not create mathematical meaning. Instead, they allow for distinctions among objects that have been or might be created by other processes. These propositions imply not that intuition must precede definitions but rather that definitions must be populated with mathematical objects before there can be any meaning.

Intuition, on the other hand, is guided by abstractions and generalizations that are well supplied with examples. Yet, intuition alone is vague and ambiguous, and meaning remains confused until the intuitions have been carefully delimited and distinguished from related ideas. That, of course, requires logic and precise definitions.

Effective mathematical communication requires that both the speaker and the hearer (or the writer and the reader) have constructed abstractions and distinctions that are somewhat comparable, in the sense of fit. Effective mathematical learning involves building such abstractions and distinctions. When and how does this happen? It is a wonder that it happens at all, given that so much of it is implicit.

Poincaré's argument suggests that in the history of mathematics, precise definitions were the result of a slow evolutionary dialectic between logic and intuition. That suggests the

hypothesis that one way of improving students' abstraction and distinctions, along with their intuition and use of definitions, is to make this dialectic explicit.

Managing Abstraction and Trying to be General

In elaborating and synthesizing the results of this study, I saw the students' use of notation at first as an issue that was mostly about imprecision and ambiguity. It now appears, however, that notation is simultaneously a tool for mediating abstraction. The claim that notation is used both to manage abstraction and to impose precision may be seen as a rephrasing of a guiding principle of semiotics: that a sign is not meaningful in itself but rather is meaningful within a system of signs. This point becomes particularly apparent when comparing the students' use of notation with standard usage.

In introducing and discussing the concept of binary operation, the instructor used a diamond or a star, intending to denote a generic operation, which might be any familiar or unfamiliar operation but which is imagined to be particular but unspecified. The generality is that in any problem (or proof) setting the diamond may denote *any* operation that satisfies the context, and thus any reasoning and results apply to *all* such operations. This generality leads to an abstraction, which is the creation of a new concept—binary operation—that is the set of all possible binary operations, familiar and unfamiliar, known and unknown, specified and unspecified. In this study, the students sometimes saw neither the generality nor the abstraction, but instead saw the diamond as another specific operation, distinct from both addition and multiplication. Furthermore, some of the students decided that the diamond was to be used when the elements were *a* and *b*, which we "don't know how to add."

Similarly, the instructor provided the students access to the set of all finite cyclic groups through the notation Z_n. The notation is general in the sense that it can stand for any such group, but it depends upon an abstraction: the set of all such groups. Once again, some of the students did not see the generality and abstraction, as evidenced by their statements that Z_3 and Z_6 were subgroups of Z_n. Unfortunately, the data are not sufficient regarding this phenomenon to provide a clear picture of the object that they were calling Z_n. The data do suggest, however, that these students saw Z_n not as a particular but unspecified group but rather as another specific group, and, furthermore, that "any Z group" (except perhaps Z itself) might be considered a subgroup of Z_n. I use the students' language "any Z group" here because the standard language and notation yields a sentence that is striking in its ambiguity: "For any n, [what is commonly called] Z_n might be considered a subgroup of [what these students called] Z_n." Again, it is not clear what object the latter Z_n denotes. Either the students were not aware of their ambiguous use of the symbol Z_n, or they were not particularly concerned by it.

The students' ambiguous use of notation proved to be an indicator of issues with both abstraction and precision. In the case of diamond and Z_n, the students were not making the intended abstractions, thereby introducing ambiguity into their use of the notations. In other cases, students used notation ambiguously when they did not make important distinctions, such as with the notation aH, which for some students represented both a particular coset and all such cosets. Similarly, some students used a symbol such as $4x$ to denote both a particular multiple of 4 and all multiples of 4. One student also used $4x$ to denote two (possibly distinct) multiples of 4.

Actually, it should not be surprising that students use notation ambiguously, for standard mathematical discourse uses notation that might appear to students to be ambiguous. Notational precision is not in the notation itself, but is implied by the surrounding text. Compare the following typical phrases that might come from an abstract algebra class:

- Suppose G is a group and $a, b \in G$.
- An operation * is commutative on a set G if $a*b = b*a$ for all $a, b \in G$.
- Suppose G is a group with 3 elements, e, a, and b, and suppose e is the identity.

This study suggests that some, perhaps many, students would call a and b *variables* in these statements, yet that label may obscure subtle but important distinctions revealed by the differences in syntax. In the first statement, a and b represent unspecified but particular elements of G. Unless stated otherwise, they cannot be assumed distinct. Generality may come later, in that whatever is argued about a and b will hold for any pair a, b in G. In the second statement, a and b are pattern generalizers. We imagine that a and b vary through all possible pairs in G, and in any such pair a and b are not necessarily distinct. The generality is in the statement itself. In the third statement, a and b are specific elements of the group G. They are names of the elements of G, and they are necessarily distinct. The generality comes later, and only as part of another abstraction: that there is only one group of order 3.

So what does this discussion imply about the relationship between generalization and abstraction? Sometimes it is hard to separate them. General reasoning seems to require abstractions. Commutativity, for example, requires imagining all possible pairs of elements from a set. General reasoning about a binary operation requires ability to imagine a generic operation. Is it necessary to construct the abstraction that is the set of

all binary operations? Is it possible to imagine $f : \Re \rightarrow \Re$ to be an arbitrary function without imagining the set of all of such functions? I say yes, although one needs to be able to imagine a wide range of diverse possibilities. In other words, general reasoning about a concept requires a significant psychological step toward the mathematical abstraction that is the set of all instances of that concept.

For some mathematical concepts, there is no simple and standard notation to distinguish between a particular thing and the set of all of them. Does $f(x) = x^2$ denote a generic function value or the set of all such values? What about $f(a) = a^2$ or $f(x_0) = x_0^2$? In most contexts the abbreviated form $f(x) = x^2$ is taken to denote the entire function, despite the misleading metonymy—in this case, substituting a generic pair for the set of all such pairs. The literature on the learning of functions suggests, however, that students have difficulty making the transition from x being a particular value to being a variable that takes all values in the domain.

By treating aK both as a specific coset and as representing all of them, some of the students demonstrated that they had applied a similar metonymy in a context where it is not often done. This ambiguity and flexible use of notation is not surprising when one considers that there is no common notation for the set of all cosets of a subgroup when the subgroup is not normal. Of course, in some contexts it is easy to make notational distinctions between a generic value and the set of all such values, such as in $4x + 1$ and $4Z + 1$.

From a psychological point of view, it is easy to see why students have trouble making distinctions between a generic value and the set of all such values. If a symbol can

represent any suitable value, it is a short conceptual leap to considering all of them. To better portray the relationship between the mathematical and psychological distinctions, the relevant distinctions are superimposed in Figure 26.

Figure 26. Mathematical and psychological distinctions

specific value → generic (particular) value → any value → all such values → the set of all such values
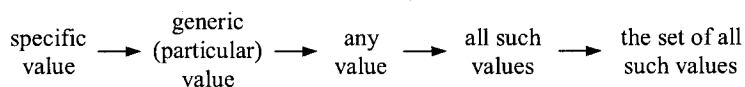
The important mathematical distinction is usually between the generic value and the set of all such values, and the distinction may be managed metaphorically by first imagining that a value is fixed and then imagining that it varies through all such values and the result is collected in a set. Mathematicians do not make an ontological distinction between a specific value, a generic value, and any such value, but the mathematical status of "all such values" is problematic unless those values are collected in a set. Thus, there are essentially two kinds of mathematical objects on this continuum, and mathematicians use notation and context as a way of managing the distinction between them.

The students in this study, in contrast, were often thinking simultaneously of any and all values, so that the important mathematical distinction was neither apparent nor relevant. When they talked about all values, they sometimes did not bother to collect the values in a set, suggesting that this distinction was also irrelevant. A specific value was psychologically distinct from the idea of any value, but the idea of a generic particular value seemed to be unavailable to them.

Thus, there are important mathematical distinctions between a generic value and the set of all such values, and standard notation conventions provide inconsistent and sometimes ambiguous support. But when and how might students learn such conceptual and

notational distinctions? Such distinctions seem to depend also on careful interpretation of the text that surrounds the introduction of the notation and on metaphors such as "imagine x is fixed." I suspect that many difficulties with quantification may be partially explained by this result.

Implications

The preceding sections have included many implications about the teaching and learning of abstract algebra in particular and mathematics more generally. In this section, I focus on what I see as this study's most important implications for teaching, teacher education, and research. Most of these implications grow out of the dialectic between making distinctions and managing abstraction.

Teaching

The results of this study make clear that students do not necessarily make standard conceptual, linguistic, and notational distinctions; in fact, the standard distinctions might not even be relevant or apparent to students. This finding has implications for teaching because of the potential for failure of communication when the teacher and the students are using conceptual grids that do not fit with one another. To overcome this obstacle, teachers must become aware of the distinctions that students are and are not making. This awareness is possible only by encouraging students to make their thinking explicit. Then, informed by knowledge of the students' thinking, teachers may help them make important conceptual distinctions and encourage them to make more of the standard distinctions in their language and notation.

The teaching of advanced mathematics must be sensitive to language, notation, and the important conceptual distinctions that the linguistic and notational distinctions are intended to convey. Consequently, language and notation should periodically be given explicit attention in instruction. This suggestion for teaching is consistent with and provides additional support for the communication and representation standards proposed by the NCTM (1989, 2000). Thus, one could suggest that mathematics teachers at all levels should ask open questions that allow students' linguistic, notational, and conceptual difficulties to become apparent. Students should be supported through instruction to translate early and often among the various signifiers of a mathematical object or property, including names, definitions, symbols, and other representations. Furthermore, students should be encouraged to articulate their thinking, for it is only through such articulation that nonstandard conceptions might be noticed.

The interview data collected for this study are noteworthy for their richness and for the fact that the students were often able to talk for long stretches with little intervention from me. The design of the study did not permit any attribution of the cause of this richness, but a plausible contributing factor was that, in this course, the students were often expected to articulate their thinking not only in their collaborative groups but also during office hours and on their written work. This interpretation implies that cooperative work can support learning by encouraging students to make their thinking explicit, even if they are not equipped to notice unusual conceptions and are not yet fluent in the standard vocabulary and syntax. Such work can also provide instructors with a window into students' thinking.

Asking students to make their thinking explicit can serve to also inculcate habits of mind that may be important to mathematics learning, such as asking the questions, "What does it mean to say that ... ?" or "What would we need to do to show that ... ?" Furthermore, by articulating their thinking, students who are not yet strong conceptually or procedurally can demonstrate important mathematical habits of mind that deserve to be supported by instruction.

Preparation of Teachers

Part of the rationale for conducting this study was that a course in abstract algebra is often required in the preparation of secondary school teachers. Thus, it makes sense to ask whether the study has implications for that preparation. A thorough investigation of the role and relevance of abstract algebra in the preparation of teachers would require a very different set of studies that would, of necessity, include significant work in secondary school mathematics classrooms to determine which ideas from abstract algebra are useful in teachers' instruction, planning, and reflection. Nonetheless, this study supports three observations.

First, abstract algebra could be a setting in which preservice teachers develop a deep sense of the nature and role of definitions and proof in mathematics. If secondary teachers are to develop in their students a sense of mathematical reasoning and proof, as is currently recommended (NCTM, 2000), then the teachers must themselves understand how definitions and proof support mathematical reasoning. These were secondary goals of the class that provided the setting for this study, the primary goal being that the students develop an intuitive and experiential sense of the concepts. Thus, it is perhaps not surprising that the students' proofs and definitions were more uneven in quality than,

say, their computations of cosets and quotient groups. It would be worth designing an abstract algebra course for preservice teachers in which intuition and rigor were twin goals and explicit foci of instruction.

Second, it is important that preservice teachers come to have a sufficiently abstract view of the concepts of inverse, identity, and binary operation, and to be able to see many examples as instantiations of the main ideas. Because the secondary mathematics curriculum includes the concept of inverse functions, for example, teachers need sufficient sense of the more abstract concept of inverse to be able to treat inverse functions in ways that are faithful to the abstract concept. Although cosets and quotient groups are certainly good candidate topics for teaching the roles of definitions and proof, this research has not convinced me that cosets and quotient groups are necessary background for high school mathematics teachers, because there are few obvious connections with high school mathematics. For preservice teachers who are to learn about quotient groups, the results of this study suggest paying particular attention to the structural relationships between Z and Z_n, for that is where the teachers can gain a firm and subtle sense of the relationships between addition of integers and addition modulo n.

Third, the results of this study suggest that sensitivity to conceptual, linguistic, and notational distinctions should be an explicit focus of the pedagogical preparation of teachers. As mentioned above, there is good reason to believe that secondary mathematics learning will be enhanced if students are often encouraged to make their thinking explicit. For that to happen in secondary school mathematics classrooms, preservice teachers should be encouraged not only to reflect on their own language and notation and to make their thinking explicit but also to reflect on such experiences as a

way to help them see the pedagogical benefits of mathematical communication and the importance of making careful distinctions.

Empirical and Theoretical Research

Many of the conclusions and implications in this chapter carry messages for research. In particular, the themes of making distinctions and managing abstraction can be seen as the beginning of a theoretical construct that takes seriously both logic and intuition. In this section, I make additional comments about some of the other theoretical constructs that informed this research and make suggestions for additional research.

First, the results of this study suggest that concept image may be a problematic construct if it is take to suggest a prescribed way of cutting up experience into concepts. The data and analysis demonstrate that students do not necessarily make standard conceptual, linguistic, and notational distinctions. To accommodate this emergent result, the analysis for this study depended upon a flexible notion of a concept image that allowed conceptual boundaries to migrate in order to explain the data. That flexibility was accomplished via a semiotic conceptual framework, which allowed consideration of the standard concept, the extent to which a student had mastered the concepts and processes, and the meanings that the student associated with the names and notations. In particular, by paying careful attention to the students' use of language and notation by separating names, symbols, and definitions in the analysis, I was able to characterize their concept images in ways that explained their nonstandard use of language and notation. Any viable notion of concept image must maintain similar flexibility that does not require that the name of the concept be the organizing determinant.

Second, as discussed in detail above, the results of this study suggest that the process/object distinction is more subtle and nuanced than is commonly portrayed in the literature. In particular, it appears that sometimes the developmental hierarchy can be circumvented when a representation supports an object conception without requiring a robust process conception. Many questions remain, however. For example, when is it advantageous to inject into instruction representations that support object conceptions before students have developed a robust process conception?

It might be profitable to acknowledge that not all process conceptions and not all object conceptions are the same. Even without comparing two different individual learners, it is clear that constructing an object conception is not the completion of a learning process but rather the beginning of a new phase of learning about the concept. In other words, an object conception creates new possibilities. Some of this might be captured by acknowledging the gradual development of conceptual proficiency, which might be described as increasing richness in one's process conception, increasing applicability of one's object conception, and increasing ability to move between the two.

This study suggests that much is to be gained by paying attention to the ways that students manage abstraction, particularly their use of metaphors and representations. Although it is clear that the students in this study did not necessarily make the standard distinctions between any object of a kind and all such objects, it appears that the idea of a particular unspecified object was not available to some of the students. I hypothesized above that metaphors such as "imagine it is fixed" may help, but many questions remain about the psychological requirements for imagining a generic object. Furthermore, there

may be many more as-yet-unidentified metaphors that both support and constrain thinking in abstract algebra in particular and advanced mathematics more generally.

If we take Poincaré's dialectic seriously, it is fair to say that this study focused on intuition at the expense of logic. I studied mostly students' informal definitions, and proof was not an explicit area of investigation. This decision made sense given the context, the available literature, and my particular interests. My intent was to study the meanings that students construct for the various concepts in group theory and to characterize those understandings. It is now quite clear, however, that the logical side needs equal attention, for that is the way that students begin to make more careful distinctions that are necessary for clear thinking and for effective participation in mathematical discourse. The semiotic perspective makes clear, after all, that signs have meaning not in themselves, but within a system of signs. Thus, many questions remain about how students learn to use definitions, notation, and language precisely, and to use such rigor in the service of reasoning, proof, and communication and to support further development of intuition.

For the continued investigation of learning in abstract algebra in particular and advanced mathematical thinking more generally, the conceptual and analytic framework developed here can offer a good starting point. The framework was developed in order to characterize students' understandings, focusing on meaning and intuition. Because it supports identification of distinctions that students do and do not make, the framework seems equally applicable to investigating not only the logical side of mathematical thinking but also the relationship between logic and intuition.

Perhaps future research in advanced mathematical thinking could take advantage of the dialectic between intuition and logic. Beginning with logic, one might ask, What does it take to build stronger connections between symbolic proofs and the examples to which the proofs are intended to apply and the intuitions the proofs are intended to formalize? Beginning with intuition, one might ask, What does it takes to encourage students to make distinctions that are not apparent to them? An extended research program could ask, What does it take for students to come to sophisticated understandings of the nature and role of definitions and proof in mathematics?

Research in advanced mathematical thinking is still a young field with many open questions, many challenges, and many opportunities. This study has explored linguistic, notational, and conceptual issues for undergraduate students in a particular abstract algebra class and has provided detailed results about the students' learning of particular concepts. Some of the conceptual issues and theoretical explanations are of sufficient generality to suspect that they may be useful in describing advanced mathematical thinking more generally. Thus, the study has also provided insights about some of the theoretical constructs commonly used in research on advanced mathematical thinking. In addition, the study has put forward and elaborated the theoretical proposition that mathematical thinking and learning may be viewed as a balance between precision and abstraction and has suggested that balance is needed in both the research on advanced mathematical concepts and the practice of teaching advanced mathematics. I hope the empirical and theoretical results presented here will influence the research agenda and ultimately will serve to improve the learning of mathematical concepts that historically have been among the most challenging in the undergraduate curriculum.

REFERENCES

Ahlfors, L. V. et al. (1962). On the mathematics curriculum of the high school. *American Mathematical Monthly, 69*(3), 189-193.

Arcavi, A. (1994). Symbol sense: Informal sense-making in formal mathematics. *For the Learning of Mathematics, 14*(3), 24-35.

Asiala, M., Brown, A., DeVries, D. J., Dubinsky, E., Mathews, D., & Thomas, K. (1996). A framework for research and curriculum development in undergraduate mathematics education. In J. Kaput, A. H. Schoenfeld, & E. Dubinsky (Eds.), *Research in collegiate mathematics education, 2* (pp. 1-32). Providence, RI: American Mathematical Society.

Asiala, M., Brown, A., Kleiman, J., & Mathews, D. (1998). The development of students' understanding of permutations and symmetries. *International Journal of Computers for Mathematical Learning, 3*(1), 13-43.

Asiala, M., Dubinsky, E., Mathews, D. M., Morics, S., & Oktac, A. (1997). Development of students' understanding of cosets, normality, and quotient groups. *Journal of Mathematical Behavior, 16*(3), 241-309.

Balaguer, M. (1998). *Platonism and anti-Platonism in mathematics*. Oxford: Oxford University Press.

Ball, D. L. (2000). Working on the inside: Using one's own practice as a site for studying teaching and learning. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 365-402). Mahwah, NJ: Erlbaum.

Ball, D. L., & Bass, H. (2000). Making believe: The collective construction of public mathematical knowledge in the elementary classroom. In D. C. Phillips (Ed.), *Constructivism in education: Opinions and second opinions on controversial issues* (Ninety-ninth Yearbook of the National Society for the Study of Education, Part I, pp. 193-224). Chicago: University of Chicago Press.

Ball, D. L., & Bass, H. (in press). Making mathematics reasonable in school. In J. Kilpatrick, W. G. Martin, & D. E. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Benson, S., & Richey, M. (1994). Much ado about coset multiplication. *PRIMUS, 4*(1), 9-14.

Benson, S. (with Findell, B.) (in press). A modified discovery approach to teaching and learning abstract algebra. In A. Hibbard (Ed.), *Innovations in the teaching of abstract algebra*. Washington, DC: Mathematical Association of America.

Beth, E. W., & Piaget, J. (1966). *Mathematical epistemology and psychology* (W. Mays, Trans.). Dordrecht: Reidel.

Bhattacharya, P. B., Jain, S. K., & Nagpaul, S. R. (1986). *Basic abstract algebra*. Cambridge: Cambridge University Press.

Bittinger, M. L. (1982). *Logic, proof and sets* (2nd ed.). Reading, MA: Addison-Wesley.

372

Boero, P. (1993). About the transformation function of the algebraic code. In R. Sutherland (Ed.), *Algebraic processes and the role of symbolism* (pp. 55-62). London: University of London, Institute of Education.

Borasi, R. (1992). *Learning mathematics through inquiry*. Portsmouth, NH: Heinemann.

Bourbaki, N. (1950). The architecture of mathematics. *American Mathematical Monthly, 57*(4), 221-232.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Breidenbach, D., Dubinsky, E., Hawks, J., & Nichols, D. (1992). Development of the process concept of function. *Educational Studies in Mathematics, 23*(3), 247-285.

Brown, A., DeVries, D. J., Dubinsky, E., & Thomas, K. (1997). Learning binary operations, groups, and subgroups. *Journal of Mathematical Behavior, 16*(3), 187-239.

Bruner, J. S. (1977). *The process of education*. Cambridge, MA: Harvard University Press. (Original work published 1960)

Bruner, J. S. (1966). *Toward a theory of instruction*. Cambridge, MA: Belknap Press.

Burn, B. (1996). What are the fundamental concepts of group theory? *Educational Studies in Mathematics, 31*(4), 371-377.

Chandler, D. (in press). *Semiotics: The basics*. London: Routledge. Retrieved July 17, 2001, from http://www.aber.ac.uk/media/Documents/S4B/

Chazan, D., & Ball, D. (1999). Beyond being told not to tell. *For the Learning of Mathematics, 19*(2), 2-10.

Clark, J. M., Devries, D. J., Hemenway, C., St. John, D., Tolias, G., & Vakil, R. (1997). An investigation of students' understanding of abstract algebra (binary operations, groups and subgroups) and the use of abstract structures to build other structures (through cosets, normality and quotient groups). *Journal of Mathematical Behavior, 16*(3), 181-185.

Cobb, P., Gravemeijer, K., Yackel, E., McClain, K., & Whitenack, J. (1997). Mathematizing and symbolizing: The emergence of chains of signification in one first grade classroom. In D. Kirshner & J. A. Whitson (Eds.), *Situated cognition: Social, semiotic, and psychological perspectives* (pp. 151-233). Mahwah, NJ: Erlbaum.

Cobb, P., & Whitenack, J. W. (1996). A method for conducting longitudinal analyses of classroom videorecordings and transcripts. *Educational Studies in Mathematics, 30*(3), 213-228.

Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics Education, 23*(1), 2-33.

Committee on the Undergraduate Program in Mathematics, Panel on Teacher Training. (1971). Recommendations on course content for the training of teachers of mathematics. In *A compendium of CUPM recommendations: Studies discussions and recommendations by the Committee on the Undergraduate Program in Mathematics of the Mathematical Association of America* (pp. 158-202). Washington, DC: Mathematical Association of America.

Conference Board of the Mathematical Sciences. (2001). *The mathematical education of teachers*. Providence, RI: American Mathematical Society.

Conway, J. H. (1995, April 13). *Re: Non-Euclidean Geometry.* Retrieved July 15, 2001, from http://www.mathforum.org/epigone/geom.pre-college/150/Pine.3.07.9504130120.O300-c100000@broccoli.princeton.edu

Cuoco, A., Goldenberg, E. P., & Mark, J. (1996). Habits of mind: An organizing principle for mathematics curricula. *Journal of Mathematical Behavior, 15*(4), 375-402.

Davis, P. J., & Hersh, R. (1981). *The mathematical experience.* Boston: Birkhäuser.

Davis, R. B. (1990). Discovery learning and constructivism. In R. B. Davis, C. A. Maher, & N. Noddings (Eds.), *Constructivist views on the teaching and learning of mathematics* (Journal for Research in Mathematics Education Monograph 4, pp. 93-106). Reston, VA: National Council of Teachers of Mathematics.

Dean, E. E. (1996). Teaching the proof process: A model for discovery learning. *College Teaching, 44*(2), 52-55.

Devlin, K. (2000). *The math gene.* New York: Basic Books.

Dienes, Z. P. (1961). On abstraction and generalization. *Harvard Educational Review, 31*(3), 281-301.

Dorier, J.-L. (Ed.). (2000). *On the teaching of linear algebra.* Dordrecht: Kluwer.

Dossey, J. (Ed.). (1998). *Confronting the core curriculum: Considering change in the undergraduate mathematics major* (MAA Notes no. 45). Washington, DC: Mathematical Association of America.

Douglas, R. (Ed.). (1986). *Toward a lean and lively calculus: Report of the conference/workshop to develop curriculum and teaching methods for calculus at the college level* (MAA Notes no. 6). Washington, DC: Mathematical Association of America.

Dreyfus, T. (1991). Advanced mathematical thinking processes. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 25-41). Dordrecht: Kluwer.

Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 95-126). Dordrecht, The Netherlands: Kluwer.

Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1994). On learning fundamental concepts of group theory. *Educational Studies in Mathematics, 27*(3), 267-305.

Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1997). A reaction To Burn's "What are the Fundamental Concepts of Group Theory?" *Educational Studies in Mathematics, 34*(3), 249-253.

Dubinsky, E., & Leron, U. (1994). *Learning abstract algebra with ISETL.* New York: Springer-Verlag.

Dubinsky, E., Schoenfeld, A. H., & Kaput, J. (Eds.). (2000). *Research in collegiate mathematics education, 4.* Providence, RI: American Mathematical Society.

Durkin, K., & Shire, B. (1991). Lexical ambiguity in mathematical contexts. In K. Durkin & B. Shire (Eds.), *Language in mathematical education: Research and practice* (pp. 71-84). Milton Keynes, UK: Open University Press.

Duvall, R. (1999). Representation, vision, and visualization: Cognitive functions in mathematical thinking. Basic issues for learning. In F. Hitt & M. Santos (Eds.), *Proceedings of the twenty-first annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 3-26). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Encyclopaedia Britannica. (1999). *Encyclopaedia Britannica CD 99*. Chicago: Author.

Ernest, P. (1991). *The philosophy of mathematics education*. London: Falmer Press.

Ernest, P. (1994). Constructivism and the learning of mathematics. In P. Ernest (Ed.), *Constructing mathematical knowledge: Epistemology and mathematics education* (pp. 1-3). London: Falmer Press.

Ernest, P. (1996). Varieties of constructivism: A framework for comparison. In L. P. Steffe, P. Nesher, P. Cobb, G. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 335-350). Mahwah, NJ: Erlbaum.

Ferrini-Mundy, J., & Findell, B. (2001). The mathematical education of prospective teachers of secondary school mathematics: Old assumptions, new challenges. In T. Rishel (Ed.), *Mathematics and the mathematical sciences in 2010: What should students know?* (pp. 31-41). Washington, DC: Mathematical Association of America.

Ferrini-Mundy, J., & Graham, K. (1994). Research in calculus learning: Understanding of limits, derivatives, and integrals. In J. J. Kaput & E. Dubinsky (Eds.), *Research issues in undergraduate mathematics learning: Preliminary analyses and results* (MAA Notes no. 33, pp. 29-45). Washington, DC: Mathematical Association of America.

Foucault, M. (1971). *The order of things: An archaeology of the human sciences* (1st American ed.). New York: Pantheon Books.

Foucault, M. (1983). *This is not a pipe* (J. Harkness, Trans.). Berkeley: University of California Press.

Fraleigh, J. B. (1989). *A first course in abstract algebra* (4th ed.). Reading, MA: Addison-Wesley.

Freedman, H. (1983). A way of teaching abstract algebra. *American Mathematical Monthly, 90*(9), 641-644.

Frorer, P., Hazzan, O., & Manes, M. (1997). Revealing the many face of abstraction. *International Journal of Computers for Mathematical Learning, 2*(3), 217-228.

Gallian, J. A. (1994). *Contemporary abstract algebra* (3rd ed.). Lexington, MA: D. C. Heath.

Geissinger, L. D. (1989). Exploring small groups [Computer software]. San Diego, CA: Harcourt Brace Jovanovich.

Glaser, B. G. (1992). *Basics of grounded theory analysis: Emergence vs. forcing*. Mills Valley, CA: Sociology Press.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Gravemeijer, K., Cobb, P., Bowers, J., & Whitenack, J. (2000). Symbolizing, modeling, and instructional design. In P. Cobb, E. Yackel, & K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms: Perspectives on discourse, tools, and instructional design* (pp. 225-273). Mahwah, NJ: Erlbaum.

Gray, E. M., & Tall, D. O. (1994). Duality, ambiguity, and flexibility: A "proceptual" view of simple arithmetic. *Journal for Research in Mathematics Education, 25*(2), 116-140.

Greeno, J. G. (1991). Number sense as situated knowing in a conceptual domain. *Journal for Research in Mathematics Education, 22*(3), 170-218.

Hamming, R. W. (1980). The unreasonable effectiveness of mathematics. *American Mathematical Monthly, 87*(2), 81-90.

Hanna, G. (1991). Mathematical proof. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 54-61). Dordrecht: Kluwer.

Hanna, G. (1995). Challenges to the importance of proof. *For the Learning of Mathematics, 15*(3), 42-49.

Hannah, J. (2000). Visual confusion in permutation representations. In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *Research in collegiate mathematics education, 4* (pp. 188-209). Providence, RI: American Mathematical Society.

Harel, G. (1989). Learning and teaching linear algebra: Difficulties and an alternative approach to visualizing concepts and processes. *Focus on Learning Problems in Mathematics, 11*(1-2), 139-148.

Harel, G., & Kaput, J. (1991). The role of conceptual entities and their symbols in building advanced mathematical concepts. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 82-94). Dordrecht: Kluwer.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education, 3*. Providence, RI: American Mathematical Society.

Harkness, J. (1983). Translator's introduction. In M. Foucault, *This is not a pipe* (pp. 1-12). Berkeley: University of California Press.

Hart, E. W. (1986). *An exploratory study of the proof-writing performance of college students in elementary group theory*. Unpublished doctoral dissertation, The University of Iowa.

Hart, E. W. (1994). A conceptual analysis of the proof-writing performance of expert and novice students in elementary group theory. In J. J. Kaput & E. Dubinsky (Eds.), *Research issues in undergraduate mathematics learning: Preliminary analyses and results* (MAA Notes no. 33, pp. 47-62). Washington, DC: Mathematical Association of America.

Hazzan, O. (1994). A student's belief about the solutions of the equation $x = x^{-1}$ in a group. In J. P. da Ponte & J. F. Matos (Eds.), *Proceedings of the Eighteenth International Conference for the Psychology of Mathematics Education* (Vol. 3, pp. 49-56). Lisbon: PME.

Hazzan, O. (1999). Reducing abstraction level when learning abstract algebra concepts. *Educational Studies in Mathematics, 40*(1), 71-90.

Hazzan, O., & Leron, U. (1996). Students' use and misuse of mathematical theorems: The case of Lagrange's theorem. *For the Learning of Mathematics, 16*(1), 23-26.

Herstein, I. N. (1975). *Topics in algebra* (2nd ed.). New York: Wiley.

Hewitt, D. (1999). Arbitrary and necessary. Part 1: A way of viewing the mathematics curriculum. *For the Learning of Mathematics, 19*(3), 2-9.

Hewitt, D. (2001). Arbitrary and necessary. Part 2: Assisting memory. *For the Learning of Mathematics, 21*(1), 44-51.

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65-97). New York: Macmillan.

Hirsch, C. R. (1981). An exploratory study of the effectiveness of a "didactical shadow" seminar in abstract algebra. *School Science and Mathematics, 81*(6), 459-466.

Holton, D., & Wenzel, J. (1993). A student-centred approach to group theory. *International Journal of Mathematical Education in Science and Technology, 24*(6), 883-888.

Hotopf, W. H. N. (1980). Semantic similarity as a factor in whole-word slips of the tongue. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 97-109). New York: Academic Press.

Hungerford, T. W. (1974). *Algebra*. New York: Springer-Verlag.

Johnson, W. (1983). A note on Lagrange's theorem. *American Mathematical Monthly, 90*(2), 132-133.

Kaplan, R. (2000). *The nothing that is: A natural history of zero*. Oxford: Oxford University Press.

Kaput, J., Schoenfeld, A. H., & Dubinsky, E. (Eds.). (1996). *Research in collegiate mathematics education, 2*. Providence, RI: American Mathematical Society.

Kaput, J. J., & Dubinsky, E. (Eds.). (1994). *Research issues in undergraduate mathematics learning: Preliminary analyses and results* (MAA Notes no. 33). Washington, DC: Mathematical Association of America.

Kemerling, G. (2001). *Definition and meaning*. Retrieved July 15, 2001, from http://www.philosophypages.com/lg/e05.htm

Kilpatrick, J. (1987). What constructivism might be in mathematics education. In J. C. Bergeron, N. Herscovics, & C. Kieran (Eds.), *Proceedings of the Eleventh International Conference for the Psychology of Mathematics Education* (pp. 3-27). Montreal: University of Montreal.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Kleiner, I. (1986). The evolution of group theory: A brief survey. *Mathematics Magazine, 59*(4), 195-215.

Kleiner, I. (1987). A sketch of the evolution of (noncommutative) ring theory. *L'Enseignement Mathématique, 33*(3-4), 227-267.

Kleiner, I. (1995). The teaching of abstract algebra: An historical perspective. In F. Swetz, J. Fauvel, O. Bekken, B. Johansson, & V. Katz (Eds.), *Learn from the masters* (pp. 225-239). Washington, DC: Mathematical Association of America.

Kline, M. (1973). *Why Johnny can't add: The failure of the new math*. New York: St. Martin's Press.

Knuth, D. E. (1984). *The TeXbook*. Reading, MA: Addison-Wesley.

Kosslyn, S. M., & Hatfield, G. (1984). Representation without symbol systems. *Social Research, 51*(4), 1019-1045.

Lacan, J. (1977). *Ecrits: A selection*. London: Tavistock.

Lajoie, C., & Mura, R. (2000). What's in a name? A learning difficulty in connection with cyclic groups. *For the Learning of Mathematics, 20*(3), 29-33.

Lakatos, I. M. (1976). *Proofs and refutations*. Cambridge, UK: Cambridge University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh : The embodied mind and its challenge to Western thought.* New York: Basic Books.

Lakoff, G., & Núñez, R. E. (1997). The metaphorical structure of mathematics: Sketching out cognitive foundations for a mind-based mathematics. In L. D. English (Ed.), *Mathematical reasoning: Analogies, metaphors, and images* (pp. 21-89). Mahwah, NJ: Erlbaum.

Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being.* New York: Basic Books.

Lauten, A. D., Graham, K., & Ferrini-Mundy, J. (1994). Student understanding of basic calculus concepts: Interaction with the graphics calculator. *Journal of Mathematical Behavior, 13*(2), 225-237.

Leitzel, J. R. C. (Ed.). (1991). *A call for change: Recommendations for the mathematical preparation of teachers of mathematics* (MAA Reports no. 3). Washington, DC: Mathematical Association of America.

Leron, U., & Dubinsky, E. (1995). An abstract algebra story. *American Mathematical Monthly, 102*(3), 227-242.

Leron, U., & Hazzan, O. (1997). The world according to Johnny: A coping perspective in mathematics education. *Educational Studies in Mathematics, 32*(3), 265-292.

Leron, U., Hazzan, O., & Zazkis, R. (1995). Learning group isomorphism: A crossroads of many concepts. *Educational Studies in Mathematics, 29*(2), 153-174.

Magee, B. (1979). *Men of ideas.* New York: Viking Press.

Mahavier, W. T. (1997). A gentle discovery method (The Modified Texas Method). *College Teaching, 45*(4), 132-135.

Mariotti, M. A., & Fischbein, E. (1997). Defining in classroom activities. *Educational Studies in Mathematics, 34*(3), 219-248.

Markovits, Z., & Sowder, J. (1994). Developing number sense: An intervention study in grade 7. *Journal for Research in Mathematics Education, 25*(1), 4-29.

Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics, 15*(3), 277-289.

Matthews, M. R. (2000). Appraising constructivism in science and mathematics education. In D. C. Phillips (Ed.), *Constructivism in education: Opinions and second opinions on controversial issues* (Ninety-ninth yearbook of the National Society for the Study of Education, Part I, pp. 161-192). Chicago: University of Chicago Press.

Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York: W. H. Freeman.

Michener, E. R. (1978). *Understanding understanding mathematics: Artificial intelligence memo no. 488* (No. LOGO-50). Cambridge, MA: Massachusetts Institute of Technology.

Moore, R. C. (1990). *College students' difficulties in learning to do mathematical proofs.* Unpublished doctoral dissertation, University of Georgia, Athens.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics, 27*(3), 249-266.

Morriss, P. (1998). Discovering a geometric volume relationship in calculus. *Mathematics Teacher, 91*(4), 334-336.

Nardi, E. (2000). Zeros and ones in advanced mathematics: Transcending the intimacy of number. *For the Learning of Mathematics, 20*(3), 44-50.

National Commission on Mathematics and Science Teaching for the 21st Century. (2000). *Before it's too late: A report to the nation from the National Commission on Mathematics and Science Teaching for the 21st Century.* Washington, DC: U.S. Department of Education.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Research Council. (1992). *Moving beyond myths: Revitalizing undergraduate mathematics.* Washington, DC: National Academy Press.

National Research Council. (2001a). *Educating teachers of science, mathematics, and technology: New practices for the new millennium.* Washington, DC: National Academy Press.

National Research Council. (2001b). *Knowing and learning mathematics for teaching: Proceedings of a workshop.* Washington, DC: National Academy Press.

Nemirovsky, R., & Monk, S. (2000). "If you look at it the other way ...": An exploration into the nature of symbolizing. In P. Cobb, E. Yackel, & K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms: Perspectives on discourse, tools, and instructional design* (pp. 177-221). Mahwah, NJ: Erlbaum.

Nicholson, J. (1993). The development and understanding of the concept of quotient group. *Historia Mathematica, 20*(1), 68-88.

Peirce, C. S. (1955). Logic as semiotic: The theory of signs. In J. Buchler (Ed.), *Philosophical writings of Peirce* (pp. 98-119). New York: Dover.

Piaget, J. (1970a). *Genetic epistemology* (E. Duckworth, Trans.). New York: Columbia University Press.

Piaget, J. (1970b). *Structuralism* (C. Maschler, Trans.). New York: Basic Books.

Piaget, J., & Garcia, R. (1989). *Psychogenesis and the history of science* (H. Feider, Trans.). New York: Columbia University Press.

Pimm, D. (1987). *Speaking mathematically.* London: Routledge.

Pimm, D. (1995). *Symbols and meanings in school mathematics.* London: Routledge.

Plato. (1998). *Phaedrus* (J. H. Nichols, Jr., Trans.). Ithaca, NY: Cornell University Press.

Poincaré, H. (1946). *Foundations of science: Science and hypothesis, the value of science, science and method* (G. B. Halsted, Trans.). Lancaster, PA: Science Press.

Polya, G. (1957). *How to solve it* (2nd ed.). Princeton, NJ: Princeton University Press.

QSR International. (2000). N5 [Computer software]. Melbourne, Australia: Author.

Raman, M. (1998, April 13-17). *Epistemological messages conveyed by high school and college mathematics textbooks.* Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Reid, C. (1986). *Hilbert-Courant*. New York: Springer-Verlag.

Rickart, C. E. (1995). *Structuralism and structures: A mathematical perspective*. River Edge, NJ: World Scientific.

Rin, H. (1982). *Linguistic barriers to students' understanding of definitions in a college mathematics course*. Unpublished doctoral dissertation, University of California, Berkeley.

Rotman, B. (1988). Toward a semiotics of mathematics. *Semiotica, 72*(1/2), 1-35.

Russell, B. (1938). *Principles of mathematics*. New York: Norton.

Sapir, E. (1949). The status of linguistics as a science. In D. G. Mandelbaum (Ed.), *Selected writings of Edward Sapir in language, culture, and personality* (pp. 160-166). Berkeley: University of California Press. (Originally published in *Language, 5*, (1929): 207-214.)

Saussure, F. d. (1959). *Course in general linguistics* (W. Baskin, Trans.). New York: Philosophical Library.

Scher, D., & Findell, B. R. (1996, September). *Research in undergraduate mathematics education: A map of the territory*. Paper presented at the Research Conference in Collegiate Mathematics Education, Central Michigan University.

Schoenfeld, A. H., Kaput, J., & Dubinsky, E. (Eds.). (1998). *Research in collegiate mathematics education, 3*. Providence, RI: American Mathematical Society.

Seife, C. (2000). *Zero: The biography of a dangerous idea*. New York: Viking.

Selden, A. A., & Selden, J. (1978). Errors students make in mathematical reasoning. *Bogazici Üniversitesi Dergisi, 6*, 67-87.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*(1), 1-36.

Sfard, A. (1994). Reification as the birth of metaphor. *For the Learning of Mathematics, 14*(1), 44-55.

Sfard, A. (1995). The development of algebra: Confronting historical and psychological perspectives. *Journal of Mathematical Behavior, 14*(1), 15-39.

Sfard, A. (2000). Symbolizing mathematical reality into being; or how mathematical discourse and mathematical objects create each other. In P. Cobb, E. Yackel, & K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms: Perspectives on discourse, tools, and instructional design* (pp. 37-98). Mahwah, NJ: Erlbaum.

Sfard, A., & Linchevski, L. (1994). The gains and the pitfalls of reification: The case of algebra. *Educational Studies in Mathematics, 26*(2-3), 191-228.

Skemp, R. R. (1987). *The psychology of learning mathematics* (Expanded American ed.). Hillsdale, NJ: Erlbaum.

Steen, L. A. (Ed.). (1992). *Heeding the call for change: Suggestions for curricular action* (MAA Notes no. 22). Washington, DC: Mathematical Association of America.

Steffe, L. P. (1990). On the knowledge of mathematics teachers. In R. B. Davis, C. A. Maher, & N. Noddings (Eds.), *Constructivist views on the teaching and learning of mathematics* (pp. 167-184). Reston, VA: National Council of Teachers of Mathematics.

Steffe, L. P., & Wiegel, H. G. (1996). On the nature of a model of mathematical learning. In L. P. Steffe, P. Nesher, P. Cobb, G. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 477-498). Mahwah, NJ: Erlbaum.

Stillings, N. A., Weisler, S. E., Chase, C. H., Feinstein, M. H., Garfield, J. L., & Rissland, E. L. (1995). *Cognitive science: An introduction.* Cambridge, MA: MIT Press.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Newbury Park, CA: Sage.

Suzuka, K. (2001). *Cultivating unfamiliar terrains.* Unpublished doctoral dissertation, Michigan State University, East Lansing.

Tall, D. (1991). The psychology of advanced mathematical thinking. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 3-21). Dordrecht: Kluwer.

Tall, D. (1992). The transition to advanced mathematical thinking: Functions, limits, infinity and proof. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 495-511). New York: Macmillan.

Tall, D., Thomas, M., Davis, G., Gray, E., & Simpson, A. (2000). What is the object of the encapsulation of a process? *Journal of Mathematical Behavior, 18*(2), 223-241.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics, with particular reference to limits and continuity. *Educational Studies in Mathematics, 12*(2), 151-169.

Thurston, W. P. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society, 30*(2), 161-177.

Touval, A. (1997). Investigating a definite integral: From graphing calculator to rigorous proof. *Mathematics Teacher, 90*(3), 230-232.

Tucker, A. C., & Leitzel, J. R. C. (Eds.). (1994). *Assessing calculus reform efforts: A report to the community* (MAA Reports no. 6). Washington, DC: Mathematical Association of America.

Vinner, S. (1976). The naive concept of definition in mathematics. *Educational Studies in Mathematics, 7*(4), 413-429.

Vinner, S. (1991). The role of definitions in the teaching and learning of mathematics. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 65-81). Dordrecht: Kluwer.

Vinner, S. (1992). The function concept as a prototype for problems in mathematics learning. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy* (MAA Notes no. 25, pp. 195-213). Washington, DC: Mathematical Association of America.

Von Glasersfeld, E. (1989). Constructivism in education. In T. Husen & T. N. Postlethwaite (Eds.), *The international encyclopedia of education: Research and studies* (Supplementary Vol. 1, pp. 162-163). Oxford: Pergamon Press.

Von Glasersfeld, E. (1990). An exposition of constructivism: Why some like it radical. In R. B. Davis, C. A. Maher, & N. Noddings (Eds.), *Constructivist views on the teaching and learning of mathematics* (Journal for Research in Mathematics Education Monograph 4, pp. 19-29). Reston, VA: National Council of Teachers of Mathematics.

Von Glasersfeld, E. (1996). Aspects of educational constructivism and its educational recommendations. In L. P. Steffe, P. Nesher, P. Cobb, G. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 307-314). Mahwah, NJ: Erlbaum.

Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Trans. newly rev. and edited ed.). Cambridge, MA: MIT Press. (Original work published 1934)

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Trans.). Cambridge, MA: Harvard University Press.

Wertsch, J. V. (1981). *The concept of activity in Soviet psychology*. Armonk, NY: M. E. Sharpe.

Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Cambridge, MA: Harvard University Press.

Whorf, B. L. (1956). *Language, thought, and reality: Selected writings*. Cambridge, MA: MIT Press.

Wilensky, U. (1991). Abstract meditations on the concrete and concrete implications for mathematics education. In I. Harel & S. Papert (Eds.), *Constructionism* (pp. 193-203). Norwood, NJ: Ablex.

Wittgenstein, L. (1973). *Philosophical investigations: The English text of the third edition* (G. E. M. Anscombe, Trans.). New York: Macmillan.

Zandieh, M. J. (2000). A theoretical framework for analyzing student understanding of the concept of derivative. In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *Research in collegiate mathematics education, 4* (pp. 103-127). Providence, RI: American Mathematical Society.

Zaslavsky, O., & Peled, I. (1996). Inhibiting factors in generating examples by mathematics teachers and student teachers: The case of binary operation. *Journal for Research in Mathematics Education, 27*(1), 67-78.

Zazkis, R. (1998). Divisors and quotients: Acknowledging polysemy. *For the Learning of Mathematics, 18*(3), 27-30.

Zazkis, R., & Dubinsky, E. (1996). Dihedral groups: A tale of two interpretations. In J. Kaput, A. H. Schoenfeld, & E. Dubinsky (Eds.), *Research in collegiate mathematics education, 2* (pp. 61-82). Providence, RI: American Mathematical Society.

Zazkis, R., Dubinsky, E., & Dautermann, J. (1996). Coordinating visual and analytic strategies: A study of students' understanding of the group D_4. *Journal for Research in Mathematics Education, 27*(4), 435-457.

Zemach, E. M. (1992). *The reality of meaning and the meaning of "reality"*. Hanover, NH: University Press of New England.

APPENDIX A

SYLLABUS AND INFORMATIONAL HANDOUTS

- Syllabus

- Questionnaire

- Arithmetic of sets discussion and proofs

- Selected answers to the homework and other problems

- Notes on cosets, April 29

383

Math 761, Spring 1996
Abstract Algebra
MWF Kingsbury 308 and Th Kingsbury M309
Textbook: *Contemporary abstract algebra* by Joe Gallian
Instructor: *Steve Benson*   **Office:** *KN M331*   **Phone:** *x-2684*
Office hours: *MWF 1:30-2:30; TuTh 11-12 (and by appointment)*
TA: *Brad Findell*   **Office:** *KN M333*   **Phone:** *x-2674*
Office hours: *MF 11-12; WTh 3-4*

This course will most likely be *very* different from your previous mathematics classes, both in content and in instructional style. However, the choices we'll be making are based on a significant amount of research on teaching and learning, as well as on our own experience as students and teachers of mathematics (specifically abstract algebra). It might be a little hard to get used to, at first, but I *really* believe that this is the best way to get us to our goal, developing an understanding of, and facility with, abstract algebra concepts and their applications. We will be taking a *discovery approach* in this class, which means that, during class time, *you* will be <u>doing mathematics</u>, rather than watching *me* do mathematics. Of course, there may very well be occasions, during class discussions for instance, that I might clarify some points or explain some details or define some terminology. As part of the discovery approach, we will often be working in KN M327 (the computer lab in Kingsbury) or the *Morse Hall* computer lab (a.k.a. *Spicerack*) using a variety of software packages designed for abstract algebra This software will be useful both for learning and for applying the concepts we'll be learning. Although we won't be following the book in "lock-step" fashion, we will be using it as a supplement to what we do in class. I find Gallian's book to be very readable with lots of nice examples, problems, and references.

Grading policy:		
	Classwork	**30%**
	Homework	**30%**
	Midterm exams	**20%**
	Final exam	**20%**

Classwork will involve class participation and work (done individually or in groups). Most of this work will be collected, some of it will be graded, but **all** collected work will be looked at and commented on (a preposition is a terrible thing to end a sentence *with*). **Homework** is self-explanatory, I suppose. Like classwork, most, if not all, of it will be looked at and some will be assigned a grade. Sometimes, the grade will depend only on whether you made a reasonable attempt to solve the problem, while other times the homework will be graded in a more "traditional" way. In addition to regular homework, there will be occasional "projects" which will be designed to pull together some of the key concepts we've dealt with in class. These projects might involve both group and individual work, depending on the particular project, and you'll typically be given several days to complete them. Since you have a little extra time, it will also be important for you to be sure that you write up your work carefully and completely. *Being able to solve a mathematical problem* is *just part of the mathematical process. It's also important that*

you be able *to explain your work and its justifications to others.* Of course, you should also write up your homework carefully, but I realize you won't have as much time to work on it! There will be **two midterm exams** during the semester (times to be announced during the first two weeks of class). There will be both in-class and take-home components to these exams, as well as on the **final exam**.

Brad and Steve will divide up the task of looking at, and making comments on, the homework, but **Steve will do the actual grading**. In addition, Steve will be the grader of the projects and exams (lucky him!).

In order to learn mathematics, one must do mathematics, not sit idly by while someone else does it. An analogy might be helpful: Learning mathematics is like learning to drive a car. While you may very well pick up some pointers while watching someone else drive, the things you learn are mainly *procedural* (where to put your hands, use one foot (unless there's a clutch), etc.). However, when you first sit behind the wheel and begin to accelerate, you realize that there's a lot more to it than in that class video simulation. Even once you're an experienced driver, you'll often make a wrong turn, forcing you to do a U- turn or even pull off the road and check the map. Sometimes, you even have to go back home, make a phone call, and start over, but you still eventually make it to your destination. Doing mathematics is like that, too. We often make false starts and mistakes, but if we keep at it, we *can* solve the problem, or learn the concept, or prove the Theorem. At the risk of stretching this analogy to the breaking point, isn't it usually very difficult to remember directions to a location you've only been <u>driven to</u> (pardon the split infinitive), and never driven yourself? Until you've actually driven somewhere yourself, it's hard to find your way back, especially if you have to take a detour which takes you off of the usual path. If you just sit and watch the teacher, even if you understand why everything he, or she, does is correct, you may very well still have a lot of difficulty remembering how to do the exact same thing at a later time. And it might seem impossible to use these ideas from class to solve new problems in the homework or test situations (sound familiar?).

Mathematics is like a video game; if you just sit and watch, you're wasting your quarter (or semester).

I'm really looking forward to this class. I really like the material you'll be learning and I hope that you will also have a good time. It will be a lot of work, but if we work together , I know that everyone can be successful.

As a final suggestion, I recommend that you keep your work in a notebook to which you can often refer. In particular, keep your "scratch work", examples an notes from class, definitions, theorems, and your own questions and conjectures in the notebook. If you like, think of the notebook as a class journal. Occasionally, I *will* be asking you to provide examples, questions, and conjectures concerning class material, so this will be good practice. I also think it will be a useful study device as you prepare for classwork, homework, and exams. At the end of the semester, I would like to see your notebooks, if you're willing to share them with me. However, the notebooks **will not be graded**. In future class discussions, I will talk more about these notebooks and why I think they're important.

Class questionnaire for Math 761

Please provide the following information. Also, be sure to check the box if you're willing to share your phone number and email address with the class.

1. Name:

2. Major/concentration:

3. phone number   May I share your number with the class? | yes | | no |

4. email address   May I share this with the class? | yes | | no |

5. past classes:

6. future plans:

7. questions, goals, hopes, concerns you have about this class, or mathematics in general:

Math 761–(arithmetic of sets discussion and proofs)

Definition. *If A and B are subsets of the integers, then $A + B$ is the set $\{a + b \mid a \in A, b \in B\}$. (When necessary to avoid confusion, we will specify that the sum is in Z.) Furthermore, if n is a positive integer, then $nA = \{na \mid a \in A\}$ and $A + n = \{a + n \mid a \in A\}$.*

For example, the sum of $\{1, 3, 4\}$ and $\{2, 6\}$ in Z is $\{3, 5, 6, 7, 9, 10\}$, $2A = \{2, 6, 8\}$ and $A + 2 = \{3, 5, 6\}$.

As another example, let's consider some infinite subsets. Recall that the Division Algorithm guarantees that every integer can be expressed (uniquely) in the form $3q$, $3q + 1$, or $3q + 2$, for some integer q. In other words, every integer is in exactly one of the sets $3Z$, $3Z + 1$, or $3Z + 2$. I'm curious, then: What's $(3Z + 1) + (3Z + 2)$? Let's choose several elements from the two sets to see if we can make a guess at what their sum is. The integers $1, -2, 4$, and 10 are all elements of $3Z + 1$, and 2, -1, -4, an 5 are all elements of $3Z + 2$. Therefore, the sum of the two sets contains $1 + 2$, $1 + (-4)$, $-2 + 5$, $4 + (-1)$, $10 + 5$, and $1 + (-4)$. That is, 3, -3, an 15 are all in $(3Z + 1) + 3Z + 2$. What do you notice about these elements? They are all multiples of 3, but does that necessarily mean that the sum contains <u>just</u> multiples of 3? And, if this is the case, are <u>all</u> of the multiples of 3 in the sum? Investigating further, we see that the following integers are also in the sum in question: $10 + (-4) = 6$, $4 + 5 = 9$, and $4 + (-4) = 0$. I'm *almost* convinced, but not quite. How can we convince ourselves, and others, that $(3Z + 1) + (3Z + 2) = 3Z$?

Is it OK to use the familiar rules involving addition to prove this? In particular, is it "legal" to just say $(3Z + 1) + (3Z + 2) = 3Z + 3Z + 1 + 2 = 3Z + 3Z + 3$? Then, since "multiples of 3" + "multiples of 3" give you multiple of 3, we're done. There are a few things "fishy" about these statements. First of all, how do I know the familiar rules apply? We just *defined* how to add sets and haven't checked that all of the properties still hold. I fact, we might be a little dubious, since we weren't terribly certain this was the "right" way to define the sum of sets. Also, the statement that we alluded to above $(3Z + 3Z = 3Z)$ isn't so obvious, either.

Let's try a slightly different approach. It does seem sort of obvious that if you add an integer which is 1 more than a multiple of 3 to an integer which is 2 more than a multiple of 3, you'll get an integer which is 3 more than a multiple of 3. But if a number is 3 more than a multiple of 3, then it must be a multiple of 3, itself. This is a fairly convincing argument that $(3Z + 1) + (3Z + 2)$ is a subset of $3Z$, but it doesn't convince me that <u>every</u> multiple of 3 is in the sum. In addition, it's entirely possible that someone might not be convinced by the intuitive argument, above. Let's see if we can make it a little more "rigorous" (that is, let's leave no room for doubt).

We want to eventually show that $(3Z + 1) + (3Z + 2) = 3Z$. For the sake of brevity, let's let $A = 3Z + 1$ and $B = 3Z + 2$. We want to show, then, that $A + B = 3Z$. A technique that often comes in handy when attempting to show that one set equals another is to show that they are both subsets of one another. That

is, we'll show that $A + B \subseteq 3Z$ and $3Z \subseteq A + B$.

First, suppose that $x \in A + B$. Then $x = a + b$, where $a \in 3Z + 1$ and $b \in 3Z + 2$, so $a = 3q + 1$ for some integer q and $b = 3k + 2$ for some integer k. Therefore, $x = (3q+1) + (3k+2) = 3q + 3k + 3 = 3(q+k+1)$, which is an element of $3Z$, since q, k, and 1, and therefore their sum, are all integers. We have succeeded in showing that any element of $A + B$ is an element of $3Z$, so we have proved $A + B \subseteq 3Z$.

We'll now attempt to prove that $3Z \subseteq A + B$. To that end, let y be an element of $3Z$ (there's no reason not to call this element x, but we don't want to cause any undue confusion and there are a lot of letters in the alphabet!). Then $y = 3n$ for some integer n. In order to show that y is also an element of $A + B$, we need to show that y can be expressed as the sum of one element from A and one element from B. That is, we need to show that $y = (3q + 1) + (3k + 2)$, for some integers q and k. Well, we can always rewrite $3n$ as $3(n - 1) + 3 = 3(n - 1) + 1 + 2$, which is *almost* what we want. But now, notice that $y = 3n = 3(n-1)+1+2 = 3n+3(-1)+1+2 = (3n + 1) + (3(-1) + 2)$, and we see that y is an element of $A + B$. Thus, we have succeeded in showing that every element of $3Z$ is also an element of $A + B$. That is, $3Z \subseteq A + B$.

We may therefore conclude that $A + B = 3Z$, and our conjecture is proved.

How would that proof have appeared in a textbook? Here's one possibility.

Proposition. $(3Z + 1) + (3Z + 2) = 3Z$.

Proof. Let $3q + 1$ and $3k + 2$ be elements of $3Z + 1$ and $3Z + 2$, respectively. Then $(3q+1) + (3k+2) = 3(k+q+1) \in 3Z$, so $(3Z+1) + (3Z+2) \subseteq 3Z$. Conversely, if $3n \in 3Z$, then $3n = [3(n - 1) + 1] + [3(-1) + 2] \in (3Z + 1) + (3Z + 2)$, showing that $3Z \subseteq (3Z + 1) + (3Z + 2)$, and the proposition is proved.

What are the main differences between these two proofs? Which is the "better" proof? Is one more correct than the other? In my mind, these proofs are *identical*, logically. Their main difference is that the first proof goes further in <u>explaining</u> each of the steps, sometimes even "talking you through" the thought processes involved in solving the problem. In this respect, the first proof is "better" if you want to know <u>how and why</u>, while the second proof is best if you just want to be convinced <u>that</u> the proposition is true. In general, the purpose of a proof is to convince you, the reader, that a statement is true. Unfortunately, in my opinion, the *raison d'être* of a proof is to convince, not necessarily to enlighten. I'd like to suggest that we write our proofs with both goals in mind. After all, if we really want to know what's going on, we want to know more that just <u>what</u> is true; we also want to know <u>why</u> things are true (knowing <u>why</u> things are true often helps us remember <u>that</u> they are true).

Exercises:

1. State and prove conjectured values of the sums $3Z + (3Z + 1)$ and $(3Z + 2) + 3Z$.

2. Complete the addition table for $3Z$, $3Z + 1$, and $3Z + 2$.

Math 761 (selected answers to the homework and other problems)

Proposition. *The set* 3**Z** *is a group under integer addition.*

Proof. We need to show that the four group axioms hold for 3**Z** (recall that 3**Z** = $\{3z \mid z \in \mathbf{Z}\}$, so the statement "$x$ is an element of 3**Z**" means that $x = 3a$ for some integer a.). First, of all, let's confirm that the associative law holds. If $3a$, $3b$, and $3c$ are in 3**Z** (notice that this means a, b, and c must be integers), then $3a + (3b + 3c) = (3a + 3b) + 3c$, since $3a$, $3b$, and $3c$ are all integers and addition is associative in **Z**. Therefore, addition is associative in 3**Z**. (That is, the associativity of addition in 3**Z** is *inherited* from the associativity of addition in **Z**). To see that 3**Z** is closed under addition, let $3a$ and $3b$ be elements of 3**Z** and note that $3a + 3b = 3(a + b)$, by the distributive law (in **Z**). But a and b are integers, so $3(a + b)$ is an element of 3**Z**. Thus, 3**Z** is closed under addition. Now, we need to determine whether 3**Z** contains an identity element. It's true that $0 + 3a = 3a + 0 = 3a$ for each $3a$ in 3**Z**, since 0 is the identity element of **Z** and 3**Z** is a subset of **Z**, but we also need to show that 0 is an element of 3**Z**. But $0 = 3(0)$, so 0 <u>is</u> an element of 3**Z** (since 0 is an integer), and we may conclude that 0 is the identity element of 3**Z**. The last property we need to confirm is the "inverse property." It's true that if $3a$ is an element of 3**Z**, then $3a + (-3a) = 0 = (-3a) + 3a$, so $(-3a)$ is the inverse of $3a$ <u>in the integers</u> but how do we know that $-3a$ is an element of 3**Z**? Well, $-3a = 3(-a)$, and if a is an integer, then so is $-a$, so $-3a$ is, indeed, an element of 3**Z**, and we've been successful in proving that every element of 3**Z** has an inverse in 3**Z**. We've now shown that all of the group axioms hold for 3**Z** under addition, so <u>3**Z** is a group under addition</u>.

Notice that several of the properties for 3**Z** follow directly from the fact that **Z** is a group under addition. Specifically, the fact that addition was associative in 3**Z** was a direct consequence of the fact that the associative law held for <u>all</u> integers, and therefore must hold for our specific subset of **Z**. Also, the identity of 3**Z** was the identity of **Z** and the inverse of each element was just its inverse from **Z**. However, we still had to check that the identity, the inverse of each element, and the result of every addition were <u>in 3**Z**</u> (these are the local, or locational, properties).

Many of you noticed that the set $\{a + bi \mid a, b \in \mathbf{Z}\}$ is <u>not</u> a group under complex number multiplication, since the element $0 = 0+0i$ does not have an inverse (if such an inverse $a + bi$ existed, then we'd have $0 = 0(a + bi) = 1$, which is impossible, since $0 \neq 1$). However, many of you also stated that the set $\{a + bi \mid a, b \in \mathbf{Z}$ and $a^2 + b^2 \neq 0\}$ <u>is</u> a group under multiplication. First, notice that this set is the original set with 0 removed. However, the set is still <u>**not closed under taking inverses**</u>, since $1 + 2i$ is in the set, but $(1 + 2i)^{-1} = \frac{1}{5} + \frac{-2}{5}i$, which is clearly not in our set, since neither $\frac{1}{5}$ nor $\frac{-2}{5}$ are integers.

Notice that we have to omit <u>a lot</u> of the elements of the set $S = \{a + bi \mid a, b \in \mathbf{Z}\}$ in order to end up with a group. In particular, the only complex numbers in the set S that have multiplicative inverses in S are $1, -1, i,$ and $-i$. The proof of this fact is left to you.

We finish this handout with a discussion of the importance of understanding set notation. Many of the group properties involve determining whether the set (which we're trying to determine is, or isn't, a group) contains certain elements (e.g. an identity, inverses for each element of the set, the "product" of set elements). For example, suppose that we

know that the set G is a group under its operation $\Diamond$. Given an element a of G, we define the subset $S_a = \{a \Diamond g \Diamond a^{-1} \mid g \in G\}$. That is, S_a consists of elements that can be expressed as $a \Diamond g \Diamond a^{-1}$ for some g in G, where the element a is always the same. That is, if x, y, and z are elements of G, then $a \Diamond x \Diamond a^{-1}$, $a \Diamond y \Diamond a^{-1}$, and $a \Diamond z \Diamond a^{-1}$ are all elements of S_a. I'm getting really tired of writing the symbol $\Diamond$, so I'm now going to write xy to denote $x \Diamond y$. Notice that we know that S_a is a subset of G, since aga^{-1} will always be in G whenever a and g (and thus g^{-1}, too) are in the group G, which is defined to be closed under its operation and under taking inverses.

As we decided in class, in order to determine whether S_a is a group under G's operation, we need only check that the following properties hold: **(1)** The identity of G is in S_a. **(2)** For each x in S_a, x^{-1} is in S_a, too (that is, S_a is closed under taking inverses). **(3)** For each x and y in S_a, xy is in S_a , too. Since each of these properties involve elements being in S_a , it's important for us to know how to determine whether a given element is in S_a. By the definition of S_a, we know that an element $\mathbf{X}$ is in S_a if and only if $\mathbf{X} = aga^{-1}$ for some g in G. With this in mind, let's proceed. First, let's show that the identity of G is in S_a. That is, if e is the identity of G, we need to <u>find</u> a g in G so that $e = aga^{-1}$. How can we find such a g? Let's work backwards. Suppose we <u>had</u> found g. Then we'd know that $e = aga^{-1}$, so we can solve for g by "multiplying" (actually, we're $\Diamond$-ing) by a^{-1} and a on the left and right sides of each equation. Therefore, $a^{-1}ea = a^{-1}(aga^{-1})a = (a^{-1}a)g(a^{-1}a) = ege = g$, so we've found that $g = a^{-1}ea = e$. But by solving for g, we had to *suppose* that g existed, and that's part of what we're trying to prove, so we need to make sure that this g really works. Clearly, though, if $g = e$, then $aga^{-1} = aea^{-1} = e$, so the identity e <u>is</u> an element of S_a.

Now, let's show that S_a is closed under taking inverses. To that end, suppose that x is an element of S_a. Then $x = aga^{-1}$ for some g in G, so $x^{-1} = (aga^{-1})^{-1}$. But is x^{-1} in S_a? It might help to figure out exactly what $(aga^{-1})^{-1}$ is. We know that $(aga^{-1})^{-1}$ is the element b of G so that $(aga^{-1})b = b(aga^{-1}) = e$. But if $b(aga^{-1}) = e$, then $b = ag^{-1}a^{-1}$ (to see this, *carefully* solve the equation for b), which <u>is</u> an element of S_a , since g^{-1} is in G. Therefore, whenever x is in S_a , we've shown that x^{-1} is in S_a , too.

Finally, we need to show that S_a is closed under the operation of G. If x and y are in S_a, then $x = aga^{-1}$ and $y = aha^{-1}$ for some g and h in G, so

$$xy = (aga^{-1})(aha^{-1}) = ag(a^{-1}a)ha^{-1} = a(gh)a^{-1},$$

which is an element of S_a , since g and h (and therefore gh) are in the group G. We've succeeded in showing, then, that S_a is closed under the operation of G.

Since we've shown that the set S_a satisfies the required properties, we may now conclude that S_a is a group under the operation of G.

Now, how would a "textbook" proof look?

Theorem. *If G is a group, $a \in G$, and $S = \{aga^{-1} \mid g \in G\}$, then S is a group under the operation of G.*

Proof. Notice that if e is the identity of G, then $e = aea^{-1} \in S$. Similarly, if aga^{-1} is in S, then $(aga^{-1})^{-1} = ag^{-1}a^{-1}$ is in S, since g^{-1} is an element of the group G (as the inverse of an

element of the group G). Finally, if aga^{-1} and aha^{-1} are in S, then $(aga^{-1})(aha^{-1}) = agha^{-1}$ is in S, since gh is in G (as the product of elements of the group G). Therefore, S is a **subgroup** of G.

A definition is in order here, don't you think? A **subgroup** of a group G is a subset of G which is itself a group under G's operation. In order to prove the above Theorem, we used the fact that the associativity law automatically holds in S, since it holds in all of G. Similarly, we didn't need to check whether e satisfied the identity property, nor did we need to check whether the elements of S <u>had</u> inverses or satisfied the inverse property, as these properties were already known to be true in the whole group G. To summarize, we used the following Theorem, the proof of which depends on the arguments of this paragraph.

Theorem. *A subset S of a group G is a subgroup of G if and only if the identity of G is in S, S is closed under the operation of G, and if the inverse of each element of S is in S.*

Math 761 Notes on Cosets
April 29, 1996

We began by stating that $f : Z_6 \to D_3$ was a homomorphism with $f(0) = (1)$ and $f(1) = (123)$. From the fact that f was a homomorphism, we calculated $f(2) = (132), f(3) = (1), f(4) = (123)$, and $f(5) = (132)$.

Below is the operation table for the range of f. If we consider f^{-1} of each element of the range we get another operation table.

$\cdot$	(1)	(123)	(132)
(1)	(1)	(123)	(132)
(123)	(123)	(132)	(1)
(132)	(132)	(1)	(123)

$\begin{array}{c} f^{-1} \\ \longrightarrow \end{array}$

$+_6$	{0,3}	{1,4}	{2,5}
{0,3}	{0,3}	{1,4}	{2,5}
{1,4}	{1,4}	{2,5}	{0,3}
{2,5}	{2,5}	{0,3}	{1,4}

The operation table for Z_6 can be rearranged with elements reordered according to their images in D_3 under f. Consider coloring the reordered table so that two elements have the same color if they have the same image in D_3.

$+_6$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

$\begin{array}{c} reorder \\ \longrightarrow \end{array}$

$+_6$	0	3	1	4	2	5
0	0	3	1	4	2	5
3	3	0	4	1	5	2
1	1	4	2	5	3	0
4	4	1	5	2	0	3
2	2	5	3	0	4	1
5	5	2	0	3	1	4

The operation table for the cosets of K where $K = \ker f$.

$+_6$	K	$1+K$	$2+K$
K	K	$1+K$	$2+K$
$1+K$	$1+K$	$2+K$	K
$2+K$	$2+K$	K	$1+K$

Math 761 Notes (cont.)
April 29, 1996

If $g : D_4 \to U_8$ is a homomorphism with $g(H) = 3$ and $g(R_{90}) = 5$, we can calculate the other values of g as follows:

$$
\begin{aligned}
g(R_0) &= 1 \\
g(R_{180}) &= g(R_{90})g(R_{90}) = 5 \cdot 5 = 1 \\
g(R_{270}) &= g(R_{90})^{-1} = 5^{-1} = 5 \\
g(V) &= g(H)g(R_{180}) = 3 \cdot 1 = 3 \\
g(D) &= g(H)g(R_{90}) = 3 \cdot 5 = 7 \\
g(D') &= g(R_{90})g(H) = 5 \cdot 3 = 7
\end{aligned}
$$

Below is the operation table for the range of g, which in this case is all of U_8. If we consider g^{-1} of each element of the range we get another operation table.

$\cdot_8$	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

$\downarrow g^{-1}$

	$\{R_0, R_{180}\}$	$\{H, V\}$	$\{R_{90}, R_{270}\}$	$\{D, D'\}$
$\{R_0, R_{180}\}$	$\{R_0, R_{180}\}$	$\{H, V\}$	$\{R_{90}, R_{270}\}$	$\{D, D'\}$
$\{H, V\}$	$\{H, V\}$	$\{R_0, R_{180}\}$	$\{D, D'\}$	$\{R_{90}, R_{270}\}$
$\{R_{90}, R_{270}\}$	$\{R_{90}, R_{270}\}$	$\{D, D'\}$	$\{R_0, R_{180}\}$	$\{H, V\}$
$\{D, D'\}$	$\{D, D'\}$	$\{R_{90}, R_{270}\}$	$\{H, V\}$	$\{R_0, R_{180}\}$

APPENDIX B

EXAMS

- Midterm exam 1

- Midterm exam 1, take-home portion

- Take-home exam 2

- Final exam

Math 761
Midterm exam #1

Please record your answers *and their explanations* on a separate sheet of paper. This sheet is yours to keep.

You may, if necessary, use the fact(s) that the sets *Z*, *Q*, *R*, and *C* (the integers, rational numbers, real numbers and complex numbers) are groups <u>under addition</u>, and the *nonzero* elements of *Q*, *R*, and *C*, respectively, form <u>multiplicative</u> groups.

1. a) Explain how we know that there are no integer solutions to $3x \equiv 8$ (mod 90)

    b) Find all integer s *x* so that $0 \le x < 91$ **and** $3x \equiv 8$ (mod 92).
       (*you may find it helpful to notice that* $3 \cdot 31 \equiv 1$ (mod 92).)

2. Write complete definitions of the following phrases. That is, carefully describe exactly what each phrase means.

    a) *e* is an identity of the set *S* under the operation $\Diamond$.

    b) The operation $\Diamond$ is associative on the set *S*.

    c) The operation $\Diamond$ is commutative on the set *S*.

    d) In the group $G, x^{-1} = y$.

    e) The set *S* is a group under the operation $*$.

3. As you might expect, $\frac{1}{2}Z$ is defined to be the set $\{\frac{1}{2}z \mid z \in Z\}$.

    a) Confirm that $\frac{1}{2}Z$ is a group under addition.

    b) Is $\frac{1}{2}Z$ a group under multiplication? Explain.

4. a) Construct the multiplication table for $\{4, 8, 12, 16\}$ in Z_{20}.

    b) Determine whether $\{4, 8, 12, 16\}$ is a group under multiplication (mod 20).

5. The following operation table is <u>not</u> a group table. Which properties fail? Explain. You **may** use the fact that you <u>know</u> that this isn't a group table in your explanation.

$*$	a	b	c	d	e	f
a	a	b	c	d	e	f
b	b	c	d	e	f	a
c	c	d	a	f	b	e
d	d	e	f	a	c	b
e	e	f	b	c	a	d
f	f	a	e	b	d	c

Math 761
Midterm #1 - take home portion, Due Friday, March 8

Be sure to justify your responses to the following problems. You may talk to Brad and Steve about these problems, and you may use your book and class notes, but please work alone.

1. The set $\{e, a, b, c, d\}$ is a group under the operation which is given in the table below. Unfortunately, I only had time to copy down part of the table. Fill out the rest of the table and check to see that we get a group. Explain how you __know__ that your choice for each entry is the only possibility. Then use *Exploring Small Groups* to confirm that you have, indeed, created a group table.

*	e	a	b	c	d
e	e				
a		b			e
b		c	d	e	
c		d		a	b
d					

2. Fill out all possible operation tables which make the set $\{e, a, b, c\}$ a group. You may assume that e is the identity of each group. In the process, be sure to explain your decisions and __prove__ that the set is a group under each of the specified operations. You may use the computer to check the associative property, but you must justify all of the other properties "by hand". You don't have to provide a formula for the operation, but explain __each__ decision you make while filling out the table. For example, explain how we know that the first row and column is the same for each possible table. What are the possibilities for a^2? Are there any that __don't__ work (that is, that won't allow the set to be a group)? In a particular table, once you've "chosen" a^2, does this "force" some (or all) of the other values in the table? Why can't we have $bc = b$? Is it possible to fill out the table so that the set is a nonabelian group?

3. Recall that the funny addition operation $*$ on the set **R** of real numbers is defined by $a*b = a+b+ab$. Show that $*$ satisfies the associative property .

4. Prove of disprove: If $x, y,$ and z are elements of a group G, then $(xyz^{-1})^{-1} = x^{-1}y^{-1}z$.

5. Prove or disprove: If $x, y,$ and z are elements of a group G, then $(xyz^{-1})^{-1} = zy^{-1}x^{-1}$.

Math 761
Take-home exam #2   **due Monday, April 29**

1. Are the following statements true, or false? Prove or find a counterexample.

   a) If G is an abelian group, then the set $\{g \in G \mid g^2 = e\}$ is a subgroup of G.

   b) If G is a group, then the set $\{g \in G \mid g^2 = e\}$ is a subgroup of G.

2. Find the subgroup generated by the given element(s) in the specified group G.

   a) The subgroup of S_4 generated by (134).

   b) The subgroup of S_5 generated by (124)(35).

   c) The subgroup of S_4 generated by (14) *and* (124). **Additional question**: Is there an element α in S_4 so that the subgroup you just found is generated by α? Explain.

3. Let the function $f: Z \to Z_4$ be defined by $f(x) = x \pmod 4$.

   a) Show that f is a homomorphism. Is f one-to-one? Is f onto?

   b) Find the kernel of f. Recall that the kernel of a homomorphism $f: G \to G'$ is the set $\ker(f) = \{g \in G \mid g^2 = e'\}$, where e' is the identity of G'.

4. Suppose $f: G \to G'$ is a group homomorphism.

   a) Prove that $\ker(f)$ is a subgroup of G.

   b) Prove that if g is an element of G, then $(f(g))^{-1} = f(g^{-1})$.

   c) Prove that $f(G)$ is a subgroup of G'. Recall that $f(G) = \{f(g) \mid g \in G\}$.

   d) Prove that if G is an abelian group, then $f(G)$ is an abelian group.

5. Let K be the kernel of the group homomorphism $f: G \to G'$ and suppose a and b are elements of G.

   a) Show that if b is in the set $aK = \{ak \mid k \in K\}$, then $f(b) = a$.

   b) Show that if $f(a) = f(b)$, then there exists an element k in K so that $b = ak$ by following the steps below.

      i) First, explain how we know that there is a k <u>in G</u> so that $b = ak$.

      ii) Now show that the k you found in part i) is <u>in K</u>.

   c) Explain why we may now conclude that $aK = \{x \in G \mid f(x) = f(a)\}$, **and therefore** $f(a) = f(b)$ if and only if $aK = bK$.

6. Prove that if f is a group homomorphism, then f is one-to-one **if and only if** $\ker(f) = \{e\}$.

MATH 761   NAME_____
Final Exam
Friday, May 10

1. Provide complete definitions for the following terms and phrases. Be sure to explain any otherwise undefined terms.

   a) associative operation

   b) commutative operation

   c) group

   d) subgroup

   e) homomorphism

   f) isomorphism

   g) the order of the element a in the group G

2. Carefully define the following groups, being sure to include the operation under which these sets are groups.

   a) $\mathbf{Z}_n$

   b) $\mathbf{U}_n$

3. Describe the following sets in words AND using set notation.

   a) the kernel of the group homomorphism $f : G \to G'$.

   b) $A + B$ (where A and B are subsets of the integers)

   c) $A * B$ (where A and B are subsets of the group G, under the operation $*$)

4. Let α be the permutation of $\{1, 2, 3, 4\}$ defined by $1 \to 2$, $2 \to 3$, $3 \to 1$, and $4 \to 4$, and let β be the permutation defined by $1 \to 4$, $2 \to 3$, $3 \to 2$, and $4 \to 1$. Find the order of $\alpha\beta^2\alpha^{-1}$.

5. Let $f : \mathbf{Z} \to 2\mathbf{Z}$ be defined by $f(z) = 2z$.

   a) Determine whether f is a homomorphism.

   b) Determine whether f is one-to-one.

   c) Determine whether f is onto.

   d) Determine whether $\mathbf{Z}$ is isomorphic to $2\mathbf{Z}$.

6. Recall that if g is an element of G, then the cyclic subgroup of G generated by g is defined to be the set $\langle g \rangle = \{g^n \mid n \in \mathbf{Z}\}$.

   a) How does the order of $\langle g \rangle$ compare to the order of g? Provide a brief explanation.

   b) If G has finite order and g is an element of G, can g have infinite order? Explain.

   c) Recall that Lagrange's Theorem states that if H is a subgroup of a <u>finite</u> group G, then the order of H is a divisor of the order of G. With this in mind, finish the following statement:

   Lagrange's (other) Theorem: If g is an element of a finite group G, then the order of g is

   _____ .

   Provide a brief explanation of why the Theorem is true.

7. Let G be a group and define the set S as follows: $S = \{x \in G \mid x^4 = x\}$

   a) Carefully state what it means for a to be in S.

   b) Carefully state what it means for b <u>not</u> to be in S

   c) Suppose $G = D_3$ and S is defined as above. Note that the table for D_3 is at the bottom of the page.

       (i) Is (123) in S?

       (ii) Is (13) in S?

   d) Now let G be an arbitrary group. Is S a subgroup of G? If so, prove it. If not, which of the properties fail? Under what circumstances will S be a subgroup of G? Prove that, under these circumstances, S is a subgroup of G.

8. Find the cyclic subgroup of $\mathbf{Z}_{15}$ generated by 5.

   a) List the left cosets of H in $\mathbf{Z}_{15}$, where H is the subgroup you found above.

   b) Is H a <u>normal</u> subgroup of $\mathbf{Z}_{15}$? Explain.

   c) Provide the table for the quotient group $\mathbf{Z}_{15}/H$.

9. Let $H = \{(1), (25)\}$ and $G = \{(1), (245), (254), (24), (25), (45)\}$. The table for G is provided below for your viewing pleasure.

   a) Confirm that H is a subgroup of G

   b) Find the left cosets of H in G.

   b) Find the right cosets of $\{(1), (25)\}$ in G.

   c) Carefully compute $(24)H(254)H$.

   d) Is H a normal subgroup of G? Explain.

10. **Extra credit:**

   a) Is $\mathbf{Z}_4$ a subgroup of $\mathbf{Z}$?

   b) Is $\mathbf{Z}_4$ a subgroup of $\mathbf{Z}_8$?

   c) When working on the problem "do the elements a and b commute?", your friend says "I think that a does, but b doesn't." Can your friend be right? How do you respond to their comment?

	(1)	(245)	(254)	(24)	(25)	(45)
(1)	(1)	(245)	(254)	(24)	(25)	(45)
(245)	(245)	(254)	(1)	(25)	(45)	(24)
(254)	(254)	(1)	(245)	(45)	(24)	(25)
(24)	(24)	(45)	(25)	(1)	(254)	(245)
(25)	(25)	(24)	(45)	(245)	(1)	(254)
(45)	(45)	(25)	(24)	(254)	(245)	(1)

	(1)	(123)	(132)	(12)	(13)	(23)
(1)	(1)	(123)	(132)	(12)	(13)	(23)
(123)	(123)	(132)	(1)	(13)	(23)	(12)
(132)	(132)	(1)	(123)	(23)	(12)	(13)
(12)	(12)	(23)	(13)	(1)	(132)	(123)
(13)	(13)	(12)	(23)	(123)	(1)	(132)
(23)	(23)	(13)	(12)	(132)	(123)	(1)

APPENDIX C

SAMPLE PROBLEM SETS AND HOMEWORK ASSIGNMENTS

- Problem sheet 1, January 17

- Assignment 1

- Problems, February 12: Definition of group

- Introduction to *Exploring Small Groups*, March 8

- Guiding questions, April 11

- Problems, April 22

- Final assignment

- Review problems, May 6

400

Math 761. Abstract Algebra
Problem Sheet 1
January 17, 1996

1 Number Theory

Abstract algebra has roots in number theory, in geometry, and in methods of solving equations. The first few assignments are designed to explore these roots. Perhaps the most famous number theory problem is Fermat's last theorem[1] (or, more accurately, Fermat's conjecture) which states that there are no non-trivial[2] integer solutions to the equation $x^n + y^n = z^n$ when n is an integer greater than 2. Investigation of this question alone has led to incredible achievements in the development of the fields of algebra, number theory, and algebraic geometry. We cannot consider the question here, but instead consider related questions.

1. Euler proved in 1770 that $x^3 + y^3 = z^3$ has no non-trivial integral solutions. In order to gain an understanding of this problem, we might first ask whether $x^3 + y^3$ is ever divisible by 3? If so, what can you conclude about x and y. If not, why not?

   Hint: A useful way to explore this equation is through modular arithmetic, sometimes called arithmetic of remainders. We say $a \equiv b$ mod n if a and b have the same remainder when divided by n. For example, $3 \equiv 24$ mod 7 because both have a remainder of 3 when divided by 7.

2. Investigate solving equations of the form $ax \equiv b$ mod n for $n = 5$ and $n = 6$, where a and b are constants. For example, does $3x \equiv 5$ mod 6 have a solution? Try several different values of a and b. Be sure that you have found ALL solutions. Summarize your results.

3. Investigate $a + x \equiv b$ mod n for $n = 5$ and $n = 6$.

4. Investigate $x^2 + 3x + 2 \equiv 0$ mod n for $n = 5$ and $n = 6$.

2 Arithmetic with sets

1. If $A = \{1, 3, 4\}$ and $B = \{2, 6\}$, can $A + 1$ make sense in a way that pays attention to arithemetic of integers? What about $A + B$, AB, $2A$, and $A + A$? What about $2Z$ and $2Z + 1$, where Z is the set of integers? Are there any choices to be made? If so, what are the advantage and disadvantages of each alternative. If not, why not.

2. Now suppose A is the set of even integers and B is the set of odd integers. What can you say about $A + B$? $A + A$? Try all possibilities. Describe, as completely as you can, arithmetic with these sets.

3. Compare the sets $3Z$, $3Z + 1$, $3Z + 2$, $3Z + 3$, $3Z + 4$, $3Z + 5$, $3Z + 6$, and $3Z + 7$. Describe, as completely as you can, arithmetic with these sets.

---

[1] At the joint meeting of the AMS and the MAA in Orlando last week, Andrew Wiles of Princeton University gave a series of lectures on his proof of this theorem.

[2] There are obvious solutions if x, y, or z are zero, but as these solutions are not very interesting, they are called "trivial."

MATH 761
Assignment #1
Due Monday, January 22

1. Read pages 3-4, omitting the last paragraph of page 4, and pages 7-9 (on modular arithmetic). In particular, carefully read through the proof of the Division Algorithm (on page 4) until you feel that you'd be able to explain it to someone else.

   (a) Rewrite the proof of the division algorithm, being sure to fill in all of the details that you feel are missing.

   (b) Explain the Well-Ordering Principle, in a way that a high school student would understand.

2. Work through problem #2 from the first handout. Compare the cases $n = 5$ and $n = 6$, using at least 4 different choices for a and b. State a conjecture based on what you find. Are there generalizations that seem to be true for all n, a, and b?

3. Work through problem #3 from the handout. Compare the cases $n = 5$ and $n = 6$, using at least 4 different choices for a and b. State a conjecture based on what you find. Are there generalizations that seem to be true for all n, a, and b?

4. Work through problem #4 from the handout. Compare the cases $n = 5$ and $n = 6$, State a conjecture based on what you find. Are there generalizations that seem to be true for all n?

5. Investigate the values of n^3 (mod 6). Make and prove a conjecture based on your "data".

6. Do #27 on page 19 of the textbook. Explain how you know you've found all n that satisfy the specified condition.

7. **Work on, but don't turn in.** Prove that $x^3 + y^3$ is divisible by 3 if and only if $x + y$ is divisible by 3.

Read the rest of pp. 3-13 (up to, but not including, equivalence relations).

Definition of group (and its consequences)

As we defined in class, a set S, along with an operation $\Diamond$, is a **group** if the following properties are satisfied:

1. For all a and b in S, $a \Diamond b$ is in S.
   We then say that S is **closed under the operation $\Diamond$**.

2. For all a, b, and c in S, $(a \Diamond b) \Diamond c = a \Diamond (b \Diamond c)$.
   We then say that $\Diamond$ **is associative on S**.

3. There exists an element e in S so that $e \Diamond a = a \Diamond e = a$ for each a in S.
   We then say that e **is the identity element of S**.

4. For each a in S, there is an element a^{-1} in S so that $a \Diamond a^{-1} = a^{-1} \Diamond a = e$ (where e is the identity element of S). We say that a^{-1} **is the inverse of a in S**.

Problems for group discussion:

1. Determine whether the subset $\{2,4,6,8\}$ of $\mathbf{Z}_{10}$ is a group under multiplication (in $\mathbf{Z}_{10}$).

In earlier problems with operation tables, we noticed that in some rows and columns of some tables, not all elements of the set occurred. For example, consider the multiplication table for $\mathbf{Z}_{10}$. Several of the rows and columns have repeated elements, so some elements have to be left out; for example, 1 is not in the "2 row".

2. In an earlier problem, you were asked to determine whether $ab = ac$ implies $b = c$ for a variety of sets and associated operations. Is this **cancellation law** true for groups? That is, if S is a group under the operation $\Diamond$ and $a \Diamond b = a \Diamond c$, then is it necessarily true that $b = c$? In other words, is it possible to cancel the a from both sides of the equation?

3. The above version of the cancellation law is often called the **left cancellation law**, since the a is canceled on the left. Determine whether the **right cancellation law** is true for groups.

4. What do the right and left cancellation laws imply about whether the rows and columns in a group's operation table have any repeated elements?

5. Show that each row and column in a group's operation table contains every group element exactly once. For example, to show that the element c occurs in the "a row", we need to show that the equation $a \Diamond x = c$ has a solution in the group. Show that this is true, explain why it shows that c is in the "a row", and show that c also occurs in the "a column".

6. Suppose that $(S, \Diamond)$ is a group with identity e. If $a \Diamond x = a$ for some a and x in S, what can you conclude about x? What if $x \Diamond a = a$? That is, what kinds of elements can act like an identity in a group?

7. How many inverses can an element in a group have? That is, if a is an element of a group S with operation $\Diamond$ and identity e, and if $x \Diamond a = e$ or $a \Diamond x = e$, what can you conclude about x?

8. If S is a group with operation $\Diamond$ and identity e, and a and b are elements of S so that $a^{-1} = b^{-1}$, what can you conclude about a and b?

Additional terminology:

Let S be a group under the operation $\Diamond$. If $a \Diamond b = b \Diamond a$ for all a and b in S, then we say that $\Diamond$ **is commutative on S** and that S **is an abelian group**. If S is not abelian, we say that S is **nonabelian**.

9. Show that the "mixed" cancellation law is true for abelian groups. That is, prove that if the groups, with operation $\Diamond$, is abelian and if $a \Diamond b = c \Diamond a$, then $a = c$.

Math 761
Introduction to the *Exploring Small Groups* software package

1. Sit down at one of the IBM clones (one of the planets) in the M227 computer lab. If the machine isn't on, turn it on. If the screen saver is on, "jiggle" the mouse.
   (Note: if you're working on **Jupiter**, you might need to press the **spacebar** in order to "wake it up".)

2. Start up *Exploring Small Groups*. If you're in **Windows**, double-click the "Groups" icon in the **Mathematics Software** folder. If you're in **DOS**, type `cd ..\math\groups\esg` followed by <RETURN> (or <Enter>), then type `start` followed by <RETURN>. If this doesn't work, type `esg` then <RETURN>.

3. At some point, it will be good to read the information provided when you start up the program, but today, we'll hit <RETURN> until we're at the **Table Generation Menu**, which I'll usually refer to as the **Main Menu**. From the Main Menu, we can choose to create our own operation table, or use one of the "canned" tables which the software "knows". **If you want to exit the program, you need to get back to this menu. In order to do that, hit the *F10* button until the computer asks you if you want to return to the table generation menu. Type *y* and you will be back at this menu. You can then type 5 to exit the program.**

4. We'll first try some of the tables the software already knows. Type **34** then <RETURN> to see the "commutative loop".

The operation table now appears before you, along with a variety of commands that you can use. Your screen should look *something* like:

1. Check Commutative Property	**5. Check Subset Closure**
2. Check Associative Property	**6. Check Group Axioms**
3. Check Identities and Inverses	**7. Table Alterations**
4. Check Cancellation Property	

*	A	B	C	D	E	F
A	A	B	C	D	E	F
B	B	C	D	E	F	A
C	C	D	A	F	B	E
D	D	E	F	A	C	B
E	E	F	B	C	A	D
F	F	A	E	B	D	C

It's fairly easy to use the software, but you've got to be sure to read the instructions on the screen. Sometimes, you've got to hit the spacebar (or some other button) to continue. In addition, you often have the choice of stepping through various calculations. For example, let's check to see if the associative law holds for the set $\{A, B, C, D, E, F\}$ (notice that it's not too hard to see that the set is closed under the operation, that A is the identity of the set, and that $A^{-1} = A$, $B^{-1} = F$, $C^{-1} = C$, $D^{-1} = D$, $E^{-1} = E$, and $F^{-1} = B$).

5. Choose to check the associative property by typing **2**. You can choose to step through the calculations, or just check to see whether the property holds, or not.

6. Of course, if you're just interested in whether or not the set is a group under the given operation, we can just check the group axioms by typing **6**. Not surprisingly, the associative law fails to hold.

7. One of the nice features of *ESG* is that we can change the table (the order of elements and the names of elements). To see this, type **7**. When the table alterations menu appears, type **2** so that we can rename the elements. Replace the elements *A, B, C, D, E,* and *F* with 1, 2, 3, 4, 5, and 6 (in that order). Be sure to hit <Return> after each new element is entered, then type **N** if the new elements are entered as you intended.

8. Now, we want to return to the main menu. Type *F10* a few times, then **Y**, when asked whether you want to go back to the table generation menu. Then, type **3** so that we can define an operation table ourselves. I want to look at the operation table for U_9 under multiplication (mod 9), so we need to create a table for {1, 2, 4, 5, 7, 8}. Therefore, type **6**, since our set has 6 elements. We then want to change the name of the elements, so we can tell that U_9 is, in fact, the set we're talking about. Do this, then fill out the rest of the table. Have the computer check that this is, indeed, a group. Having done this, you will be able to use the *Group properties menu*.

9. By choosing the Powers and orders option (#1), you can compute the powers of elements much more quickly than is possible by hand. Use this option to compute the powers and orders of each element of the group. What are the different **orders** of the elements of the group?

10. Get back to the main menu and choose the Sample Library (#1), then choose to view the operation table for U_{20}. Again, find the **order** of each element, after checking that the group axioms hold, so that we may gain access to the Group properties menu. While you're doing this, find the sets $\{g^2 \mid g \in U_{20}\}$, $\{g^3 \mid g \in U_{20}\}$, and $\{g^4 \mid g \in U_{20}\}$, the collections of squares, cubes, and fourth powers in U_{20}.

11. Use the table alterations menu (option #7) to focus on the set $\{g^2 \mid g \in U_{20}\}$. Accept the "restrict to a closed subset" option to see whether the set is closed under the operation. After this is done, hit the *F10* key to get back to the group properties menu. Check to see whether the group axioms hold. That is, determine whether $\{g^2 \mid g \in G\}$ is a group.

12. In order to check whether $\{g^3 \mid g \in U_{20}\}$ and $\{g^4 \mid g \in U_{20}\}$ are groups, you'll have to repeat steps 10 and 11. In particular, you'll need to get back to the main menu, then choose the operation table for U_{20}, then restrict the table to the specified set, etc.

There is no better way to learn <u>how</u> to use *Exploring Small Groups* than to just start using it. The following exercises have been designed with this in mind.

Problems

1. Determine whether the subset {AD, BC, ~A, ~B, ~C, ~D} (of set #30 in the Sample Library) is a group. By choosing the table alterations menu (option 7),

then the "restrict to a closed subset option (4)" we can choose to focus only on the above listed elements.

2. List any of the following elements that exist: an identity for the above set and each element which has an inverse (with its inverse).

3. Create an operation table for the set $\{e, a, b\}$ so that the set, along with the operation, forms a group (I'll bet you've already guessed which element we should choose to be the identity). While it's possible to create the table without the use of the software, I'd recommend using the computer to check the associative law, at least. In order to do this, choose the **user defined table** command from the main menu, then ask for a table for a 3 element set. Accept the option to change the names of the elements so that your table starts off looking like the one below.

$$
\begin{array}{c|ccc}
* & e & a & b \\
\hline
e & & & \\
a & & & \\
b & & &
\end{array}
$$

In determining how to fill out the table, recall that a number of group properties come directly from the operation table. For instance, we *know* what to put in each entry of the e row and column, right? Be sure to explain each of your choices for table entries. When you're done, check to see whether you've created a group, or not.

4. Once you've created a group, determine whether it is possible to make any *different* choices and still end up with a group? Explain.

5. For each of the following group operation tables, determine whether the set $\{g^2 \mid g \in G\}$ is a group (follow 11 and 12 above). Does there seem to be a way to predict when the set will be a group? Check tables 0602, 0802, 0804, 0901, 0902, and 1204 from the Group Library and 12 and 18 from the Sample Library.

Math 761
Guiding Questions
April 11, 1996

You now have many examples of groups: permutation groups, dihedral groups, the groups Z_n and U_n, groups involving matrices, and also groups involving the integers, rational, real, and complex numbers. In order to understand these groups, it is useful to ask:

What are the easiest groups to describe and understand?

For these groups, and especially for the more complicated groups, it is useful to break the groups down into smaller pieces. We can ask:

What are all the subgroups of a given group?

Some of you have noticed that the subgroup of S_4 generated by (1432) looked a lot like Z_4. We might ask whether it is essentially the same as Z_4. Or we might ask the following more general question:

Given a group (or a subgroup of a group), is it "essentially the same" as another more familiar group?

In order to show that two groups are essentially the same, we must set up a correspondence between them. Such a correspondence can be given by a function which maps one group to the other.

Can we specify a function which shows that two groups are essentially the same?

But there must be more than just a correspondence between the elements of the two groups. The function must also show a relationship between the operations of the two groups.

What properties must such functions have in order to show a relationship between the group operations?

When groups are not essentially the same, they might still have important similarities. So given two groups, we might ask:

Given a function which maps one group to another, what kind of relationship does the function establish between the arithmetic of the two groups?

This last question sounds quite abstract, but in fact it is just a generalization of the question, "How is addition in Z related to addition in Z_n?"

Math 761
Problems to work on the week of April 22

Suppose that G and G' are groups under the operations * and *', respectively. Recall that a function $f: G \rightarrow G'$ satisfying the property that for all x and y in G, $f(x*y) = f(x) *' f(y)$, is called a (group) **homomorphism**. If f is also one-to-one *and* onto (or, as the French are wont to say, *injective* and *surjective*), then we say that f is an **isomorphism** and that G is **isomorphic** to G'. We further define $f(G) = \{f(g) \mid g \in G\}$ to be the **image of G under f**. (That is, $f(G)$ is the set of all possible outputs from f arising from inputs in G).

The kernel of f is defined to be the set $\{g \in G \mid f(g) = e'\}$, where e' is the identity of G' and is often abbreviated as ker(f). Note also that, even though the kernel typically refers to a homomorphism, we'll use the word to denote the elements of G which f sends to e' even if f is not a homomorphism.

1. Determine whether the following functions are group homomorphisms. Which, if any, are isomorphisms?

   a) $f: \mathbf{Z}_5 \rightarrow \mathbf{Z}_{10}$ defined by $f(x) = x \pmod{10}$.

   b) $f: \mathbf{Z}_5 \rightarrow \mathbf{Z}_5$ defined by $f(x) = x^2$.

   c) $f: \mathbf{D}_3 \rightarrow \mathbf{D}_3$ defined by $f(x) = x^2$.

   d) $f: \mathbf{Z}_6 \rightarrow \mathbf{Z}_6$ defined by $f(x) = 2x$.

   e) $f: \mathbf{Z} \rightarrow \mathbf{Z}_5$ defined by $f(x) = x \pmod 5$.

   f) $f: \mathbf{Z} \rightarrow 2\mathbf{Z}$ defined by $f(x) = 2x$.

2. For which of the above f is ker(f) a subgroup of the domain of f?

3. Partially define $f: \mathbf{Z}_6 \rightarrow D3$ by $f(0) = (1)$ and $f(1) = (123)$. In order that f be a homomorphism, how should we define $f(2) = f(1 + 1)$? $f(3)$? $f(4)$? $f(5)$? $f(6)$?

4. Suppose $f: G \rightarrow G'$ is a group homomorphism and g is an element of G.

   a) Prove that if n is a positive integer, then $f(g^n) = f(g)^n$.

   b) Prove that if n is an integer (not necessarily positive), then $f(g^n) = f(g)^n$.

   c) If g is an element of G and the order of g is 6, must the order of $f(g)$ be 6, as well? Hint: see 1. d) above. What can you say about the order of $f(g)$?

5. For each of the following groups G, determine whether the function $f: G \rightarrow G$, defined by $f(x) = x^2$ is a homomorphism. For which, if any, G is f an isomorphism?

   a) $G = \mathbf{Z}_4$

   b) $G = \mathbf{Z}_5$

   c) $G = \mathbf{Z}_6$

   d) $G = \mathbf{D}_3$

Math 761
Problems to work on the week of April 22

6. Use *Exploring Small Groups* to find all of the subgroups of D_3 and D_4. Note that while *many* of the subgroups of a given group are generated by a single element, some groups have subgroups which *are* generated by more than one element. For example, {(1), (12), (34), (12)(34)} is a subgroup of S_4 but it is not generated by a single element of S_4. It is, however, generated by (12) *and* (34). Make a conjecture about the order of any subgroup of a finite group.

7. Suppose G is a finite group, H is a subgroup of G, and a is an element of G. Recall that aH is defined to be the set $\{ah \mid h \in H\}$. In this context, we call aH the left coset of H containing a.

   a) Explain why it makes sense to say that "aH contains a." Therefore, every element of G is in <u>at least one</u> left coset of H.

   b) Show that $|aH| = |H|$ by showing that if $x \neq y$, then $ax \neq ay$. Therefore, every left coset of H has $|H|$ elements.

   c) Show that c is in aH if and only if $c^{-1}a$ is in H.

   d) Suppose that c is in $aH \cap bH$. Show that $aH = bH$ by showing that $a^{-1}b$ is in H. Therefore, if left cosets overlap, they are identical. Therefore, every element of G is in <u>at most one</u> left coset of H.

   e) We've thus shown that every element of G is in exactly one left coset of H and that each left coset has $|H|$ elements. That is, the left cosets of H in G "partition" G into disjoint, equal size pieces. Explain how we may now conclude that $|aH|$ is a divisor $|G|$.

Math 761.
Final assignment.

One of the central tenets of this class has been that you, the students, should come to some personal conviction and some group consensus about definitions, questions and answers. One consequence of this approach has been that sometimes consensuses have come slowly and sometimes they haven't come at all. Still, in order to participate in discussions about the ideas in this course, it is important to reach consensus on many of the basic facts.

The following activities are designed to bring the class toward greater consensus. They will help you prepare for the final exam and will also help Steve and Brad make up the final exam. In the activities below, you are asked to write a list of questions. Some of the questions you write will appear on the final.

Responses to each of the following activities are **due at the last class meeting**. During that class we will discuss and consolidate your responses.

1. Write a list of questions about "basic facts" that you believe everyone who completes this course should know. These questions might take one of the following forms:

What is _____ ?
What is meant by _____ ?
How many ways _____ ?
How do you _____ ?

Include answers for each of these questions.

Example:

Q: What is Z_n?
A: *You decide.* [Did you think we might slip and give an answer?]

2. Come up with a list of questions about the big ideas of the course. These questions should necessarily be more general (and abstract) than the questions above.

Example:

Q: Why do the rules for exponents make sense? Are the any differences between positive, negative, and zero exponents? What assumptions must be made in order to begin? Explain.

A: Again, you decide.

Math 761. Abstract Algebra
Review Problems
May 6, 1996

1. From your take-home exams, it seems that some students are unclear about proper use of set notation. Determine which of the following items are the same as which others. Assume G is a group and e is the identity.

   (a) $\{g \in G \mid g^2 = e\}$

   (b) $\{g^2 \mid g \in G\}$

   (c) $\{x \mid x \in G, x^2 = e\}$

   (d) $h \in G$ where $h^2 = e$

   (e) g^2 where $g \in G$

   (f) $\{k^2 \mid k \in G, k^2 = e\}$

   (g) $\{x^2 = e \mid x \in G\}$

   (h) $\{g^2 \in G \mid g^2 = e\}$

   (i) The squares of the elements in G.

   (j) The elements in G whose squares are the identity.

   (k) The squares of the elements in G whose squares are the identity.

   (l) An element in G whose square is the identity

   (m) The square of an element in G.

2. You should find that you know how to do each of the following problems, once you understand what the question is asking.

   (a) Let H be the subgroup of S_4 generated by (1432). Find the left cosets of H in S_4. How many should there be? Is H a normal subgroup?

   (b) Make an operation table for the quotient group $Z_{12}/\langle 9 \rangle$.

   (c) Compute the left cosets of $\langle 4 \rangle$ in U_{15}. Is the subgroup normal? Why or why not? If so, make an operation table for the quotient group.

   (d) In a group of order 18, what are the possible orders of elements in the group?

APPENDIX D

IRB APPROVAL AND CONSENT FORMS

- Institutional Review Board approval

- Consent form

- Consent form for use of video data

412

UNIVERSITY OF NEW HAMPSHIRE

Office of Sponsored Research
111 Service Building
51 College Road
Durham, New Hampshire 03824-3585

(603) 862-2000 PROPOSALS & AWARDS
(603) 862-3716 ACCOUNTING
(603) 862-3750 DIRECTOR
(603) 862-3564 FAX

March 20, 1996

Mr. Brad Findell
Mathematics
Kingsbury Hall
Campus Mail

IRB Protocol #1694 - Learning In Abstract Algebra

Dear Mr. Findell:

The Institutional Review Board (IRB) for the Protection of Human Subjects in Research has reviewed the protocol for your project as Exempt as described in Federal Regulations 45 CFR 46, Subsection 46.101(b)(2). Approval is granted to conduct the project as described in your protocol. If you decide to make any changes in your protocol, you must submit the requested changes to the IRB for review and approval prior to any data collection from human subjects.

The protection of human subjects is an ongoing process for which you hold primary responsibility. In receiving IRB approval for your protocol, you agree to conduct the project in accordance with the ethical principles and guidelines for the protection of human subjects in research as described in "The Belmont Report." Additional information about other pertinent Federal and university policies, guidelines, and procedures is available in the UNH Office of Sponsored Research.

There is no obligation for you to provide a report to the IRB upon project completion unless you experience any unusual or unanticipated results with regard to the participation of human subjects. Please report these promptly to this office.

If you have any questions or concerns, please feel free to contact Kara Eddy, Regulatory Compliance Officer (for the IRB), at 862-2003. Please refer to the IRB # above in all future correspondence related to this project. We wish you success with the research.

Sincerely,

Kathryn B. Cataneo, Executive Director
Research Administration
(for the IRB)

KBC: ke

Enclosure

cc: Karen Graham (advisor), Mathematics

LEARNING IN ABSTRACT ALGEBRA

INFORMED CONSENT FORM

Learning in Abstract Algebra is a dissertation in mathematics education. The aim of the dissertation is to describe how students think about the concepts in abstract algebra. It is hoped that this dissertation will lead to better teaching and learning in undergraduate mathematics.

You may participate in this study in any, all, or none of the following ways:

- by allowing copies of your written work to be included as data;
- by allowing your discussions to be audiotaped during regular classtime;
- by allowing your discussions to be videotaped during regular classtime; or
- by participating in videotaped interviews with the researcher.

Because the interviews will require time outside of class, you will be paid $6/hour for that time. Approximately four interviews of about one hour each will be scheduled during the semester.

Many students who participate in research of this type typically find the process to be helpful in their own learning. They benefit because in order to communicate with the researcher and with other students, they reflect upon and deepen their understandings of the mathematical concepts involved.

PLEASE READ THE FOLLOWING STATEMENTS AND RESPOND AS TO WHETHER OR NOT YOU ARE WILLING TO PARTICIPATE.

1. I understand that the use of human subjects in this project has been approved by the UNH Institutional Review Board (IRB) for the Protection of Human Subjects in Research.

2. I understand the scope, aims, and purposes of this research project and the procedures to be followed and the expected duration of my participation.

3. I have received a description of any potential benefits that may be accrued from this research and understand how they may affect me or others.

4. I understand that my consent to participate in this research is entirely voluntary, and that my refusal to participate will have no effect on my grade in Math 761.

5. I further understand that if I consent to participate, I may discontinue or modify my participation at any time with no effect on my grade in Math 761.

6. I confirm that no coercion of any kind was used in seeking my participation in this research project.

7. I understand that if I have any questions pertaining to the research or my rights as a research subject, I have the right to call Dr. Van Osdol (862-2690) or the UNH Office of Sponsored Research (862-2000) and be given the opportunity to discuss such questions in confidence.

8. I understand that I will be paid $6/hour for participation in interviews to be conducted outside of classtime. I further understand that there will be no financial compensation for other participation.

9. I understand that anonymity and confidentiality of all data records associated with my participation in this research, including my identity, will be fully maintained to the best of the researcher's ability.

10. I understand that data from this study may be used in presentations for audiences of researchers and teachers.

11. I agree to respect the confidentiality and anonymity of the other participants to the best of my ability

12. I certify that I have read and fully understand the purpose of this research project and its risks and benefits for me as stated above.

I, _____, CONSENT to participate in this research project in the following ways. (Initial all that apply.)

    ____ by allowing copies of my written work to be included as data;
    ____ by allowing my discussions to be audiotaped during regular classtime;
    ____ by allowing my discussions to be videotaped during regular classtime;
    ____ by participating in a videotaped interview with the researcher.

I, _____, DECLINE to participate in this research project.

_____   _____
Signature of Student   Date

LEARNING IN ABSTRACT ALGEBRA

INFORMED CONSENT FORM FOR USE OF VIDEO DATA

Learning in Abstract Algebra is a dissertation in mathematics education. The aim of the dissertation is to describe how students think about the concepts in abstract algebra. It is hoped that this dissertation will lead to better teaching and learning in undergraduate mathematics.

Data collected as part of this research project will be held strictly confidential. I will use videotapes and audiotapes primarily to develop written transcripts. When using excerpts from these transcripts in research papers and presentations, I will use pseudonyms to protect your anonymity. When using actual videotape rather than transcripts in a research presentation, however, it is not always possible to maintain anonymity. Thus it is important that I request specific permission for such use of video.

PLEASE READ THE FOLLOWING STATEMENTS AND RESPOND AS TO WHETHER OR NOT YOU ARE WILLING TO ALLOW USE OF VIDEO DATA.

1. I understand that the use of human subjects in this project has been approved by the UNH Institutional Review Board (IRB) for the Protection of Human Subjects in Research.

2. I understand the scope, aims, and purposes of this research project and the procedures to be followed and the expected duration of my participation.

3. I have received a description of any potential benefits that may be accrued from this research and understand how they may affect me or others.

4. I understand that my consent to the use of video in presentations is entirely voluntary, and that my refusal to participate will have no effect on my grade in Math 761.

5. I further understand that if I consent to the use of video in presentations, I may withdraw my consent at any time with no effect on my grade in Math 761.

6. I confirm that no coercion of any kind was used in seeking my consent to the use of video in presentations.

7. I understand that if I have any questions pertaining to the research or my rights as a research subject, I have the right to call Dr. Van Osdol (862-2690) or the UNH Office of Sponsored Research (862-2000) and be given the opportunity to discuss such questions in confidence.

8. I understand that there will be no financial compensation for my consent to the use of video in presentations.

9. I understand that any video data from this study may be used in presentations for audiences of researchers and teachers.

10. I certify that I have read and fully understand the purpose of this research project and its risks and benefits for me as stated above.

I, _____, CONSENT to the use in presentations of video of me gathered as part of this research project.

I, _____, DO NOT GIVE CONSENT to the use in presentations of video of me gathered as part of this research project.

_____   _____
Signature of Student   Date