ABSTRACT

ALLAIN, ASHLEY. Development of an Instrument to Measure Proportional Reasoning Among Fast-Track Middle School Students. (Under the direction of North Carolina State University Graduate Faculty).

The purpose of the study was to develop a reliable and valid instrument for measuring proportional reasoning among fast-track middle school girls in Wake County, North Carolina. The study sample consisted of 70 girls who attended the summer 2000 Girls on Track program at Meredith College located in Raleigh, North Carolina. The grade level for each of the participants ranged from 6th grade through 8th grade for the 2000-2001 school year. The instrument used in this study is the Proportional Reasoning Assessment Instrument.

This instrument was developed by the researcher and is based upon problems discussed in relevant literature. The test items chosen include missing value, comparison, mixture, associated sets, part-part-whole, graphing and scale problems. The instrument is comprised of 10 open-ended items of varying difficulty levels.

Data were analyzed using Statistical Package for the Social Sciences Version 10.0 (SPSS) and EXCEL. A four-point grading rubric was used to score each test item. Two measure of internal consistency were calculated to determine reliability: Chrombach’s coefficient alpha and inter-rater reliability. A panel of experts examined the test instrument for the qualities of relevance, balance, and specificity to establish content validity. Criterion validity was established through determining the correlation between students’ scores on the Proportional Reasoning Assessment Instrument and the students’ scores on the North Carolina End-of-Grade exam. A detailed item analysis was
performed including item difficulty, item discrimination, item means, item variances, and inter-item correlations.

Results from the study reveal the Proportional Reasoning Assessment Instrument is a reliable and valid test instrument for measuring proportional reasoning among fast-track middle school girls. In addition, the instrument revealed common misconceptions among the students in the sample. The overall coefficient alpha is $\alpha = .7083$ and inter-rater agreement was 96%. The average difficulty is $p_{\text{mean}} = .4$ and the average discrimination is $D_{\text{mean}} = .405714$. Each test item contributed to the central purpose of the instrument due to the absence of negative discrimination indices.
DEVELOPMENT OF AN INSTRUMENT TO MEASURE PROPORTIONAL REASONING AMONG FAST-TRACK MIDDLE SCHOOL STUDENTS

by

ASHLEY ALLAIN

A thesis submitted to the Graduate Faculty of North Carolina State University in partial fulfillment for the Degree of Master of Science

MATHEMATICS EDUCATION

Raleigh

2000

APPROVED BY:

Chair of Advisory Committee

Karen R. Dawkins
I dedicate this thesis to my daughter, Abigail Ruth.

With faith in God, dedication and hard work,

you can accomplish all things
BIOGRAPHY

My name is Ashley Allain. I was born in Birmingham, Alabama and attended Vestavia Hills High School. I received a Bachelors of Science in Secondary Mathematics Education from the University of Alabama in May 1997. While attending the University of Alabama, I was treasurer of Kappa Delta Epsilon Education Honor Society and a member of Kappa Delta Pi and Pi Mu Epsilon honor societies. I was actively involved in various ministries at St. Francis of Assisi University parish where I taught Sunday school, served as a Eucharistic minister, wedding coordinator, and RCIA sponsor.

I was married to Rhett Allain on August 5, 1995 at St. Francis of Assisi University Parish. Upon graduation, we moved to Raleigh, North Carolina where Rhett entered the doctoral program in Physics at North Carolina State University. I began my career in education at Cary High School located in Cary, North Carolina. There I taught Pre-Algebra, Introductory College Mathematics, and Advanced Placement Statistics. After my first year of teaching, I entered the masters program at North Carolina State University.

My husband and I are blessed with a beautiful daughter, Abigail Ruth, who was born April 18, 2000. We attend the Catholic Community of St. Francis of Assisi where we are involved with youth ministry and social outreach programs.
ACKNOWLEDGEMENTS

The completion of this masters program could not have been possible without the support, patience and generosity of many people. I would like to begin by thanking God for giving me the wisdom to see me through this endeavor. It is through faith and trust in God I have come this far. The following verse from Philippians 4:13, “I can do all things through Christ who gives me strength,” has carried me over the past couple of years.

I would like to thank my husband, Rhett Allain. His love, support, and patience have been some of the greatest and most wonderful gifts. You have given me wonderful advice and counsel as well as been my best friend through all of this. I would like to thank my daughter, Abigail. Although you are just a baby, your laughter has brought much joy and happiness to my days. I would also like to thank my family whose prayers and support have helped me attain this goal.

I would like to thank my advisor, Dr. Sandra Berenson, for all of her hard work and support. She has been a blessing throughout this program. I would also like to thank Dr. Jacqueline Dietz for her encouragement and advice. I would like to thank Laurie Cavey for assisting with preliminary interviews and Matthew Clark for his contributions to this project as well. To everyone at the Center for Research in Science and Mathematics Education, I would like to express thanks for your support and encouragement. Finally, to the statistics, mathematics, and education faculty, thank you for challenging and expanding my mind through your instruction. Without your efforts, I could not have achieved this goal.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. LITERATURE REVIEW</td>
<td>6</td>
</tr>
<tr>
<td>Understanding Proportional Reasoning</td>
<td>6</td>
</tr>
<tr>
<td>Types of Proportional Reasoning Problems</td>
<td>7</td>
</tr>
<tr>
<td>Validity</td>
<td>12</td>
</tr>
<tr>
<td>Reliability</td>
<td>15</td>
</tr>
<tr>
<td>Item Analysis</td>
<td>18</td>
</tr>
<tr>
<td>Summary</td>
<td>19</td>
</tr>
<tr>
<td>3. METHODOLOGY</td>
<td>20</td>
</tr>
<tr>
<td>Research Design and Instrumentation</td>
<td>20</td>
</tr>
<tr>
<td>Data Collection</td>
<td>24</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>25</td>
</tr>
<tr>
<td>Summary</td>
<td>29</td>
</tr>
<tr>
<td>4. RESULTS</td>
<td>31</td>
</tr>
<tr>
<td>Preliminary Findings</td>
<td>31</td>
</tr>
<tr>
<td>Research Findings</td>
<td>32</td>
</tr>
<tr>
<td>Summary</td>
<td>37</td>
</tr>
<tr>
<td>5. SUMMARY, CONCLUSIONS, IMPLICATIONS, AND RECOMMENDITIONS</td>
<td>38</td>
</tr>
<tr>
<td>Summary of Procedures</td>
<td>38</td>
</tr>
<tr>
<td>Summary of Findings</td>
<td>39</td>
</tr>
<tr>
<td>Conclusions</td>
<td>45</td>
</tr>
<tr>
<td>Discussion</td>
<td>46</td>
</tr>
<tr>
<td>Recommendations</td>
<td>52</td>
</tr>
<tr>
<td>6. LIST OF REFERENCES</td>
<td>54</td>
</tr>
<tr>
<td>7. APPENDICES</td>
<td>56</td>
</tr>
<tr>
<td>Appendix A. Approval Letter for Research on Human Subjects.</td>
<td>57</td>
</tr>
</tbody>
</table>
Appendix B.  Proportional Reasoning Assessment Instrument……  59
Appendix C.  Key of Problem Types……………………………………  65
Appendix D.  Grading Rubric………………………………………………  67
Appendix E.  Comments From Preliminary Interviews………………  69
Appendix F.  Sample of Student Responses to Item 8………………  75
Appendix G.  Sample of Student Responses to Item 9………………  83
Appendix H.  Sample of Student Responses to Item 10………  88
LIST OF TABLES

Table 2.1  Missing value and comparison problems................................. 8
Table 2.2  Semantic problem types....................................................... 9
Table 2.3  Other types of proportional reasoning problems......................... 11
Table 2.4  Guidelines for interpreting coefficient alpha............................. 16
Table 3.1  Guidelines for interpreting difficulty indices............................ 28
Table 3.2  Guidelines for interpreting discrimination indices..................... 29
Table 4.1  Item means and variances for the Proportional Reasoning Assessment Instrument ................................................................. 32
Table 4.2  Item difficulty for the Proportional Reasoning Assessment Instrument ................................................................. 34
Table 4.3  Item discrimination for the Proportional Reasoning Assessment Instrument ................................................................. 36
Table 5.1  Summary of frequencies and percentages of student scores for each item ................................................................. 40
Table 5.2  Summary of statistical reliability analysis for the Proportional Reasoning Assessment Instrument ................................................................. 42
Table 5.3  Summary of item difficulty and item discrimination indices for each of the ten test items ................................................................. 43
LIST OF FIGURES

RESULTS

1. Bar graph of item difficulty for Proportional Reasoning Assessment Instrument .......................................................... 35

2. Bar graph of item discrimination for Proportional Reasoning Assessment Instrument ......................................................... 36

SUMMARY, CONCLUSIONS, DISCUSSION AND RECOMMENDATIONS

1. Item difficulty versus item discrimination ................................. 45

2. Item 8 of the Proportional Reasoning Assessment Instrument .......................................................... 47

3. Item 9 of the Proportional Reasoning Assessment Instrument .......................................................... 48

4. Item 10 of the Proportional Reasoning Assessment Instrument .......................................................... 50

5. Item 5 of the Proportional Reasoning Assessment Instrument .......................................................... 51
CHAPTER I
INTRODUCTION

This study focuses on the development of a valid and reliable instrument that measures proportional reasoning among fast-track middle school girls. This form of mathematical reasoning permeates many different areas of mathematics as well as everyday life. Behr, Lesh, and Post (1988) assert proportional reasoning is a pivotal concept that serves as the capstone of children’s elementary school arithmetic and the cornerstone of all that is to follow. This validates the importance for teachers to quickly assess the sophistication of their students’ thinking strategies, specifically the level at which their students are reasoning proportionally.

In Piaget’s stages of development, proportional reasoning is considered to usher in the beginning of the formal operations stage (Inhelder and Piaget, 1958). As a result, the focus of many research studies revolves around adolescent students. Through the course of time, however, the true definition of proportional reasoning has become clouded. Many researchers have attempted to define proportional reasoning in their own terms. Inhelder and Piaget (1975) suggest proportional reasoning is a relationship between two relationships rather than a relationship between two concrete objects. In this section, I explore how different researchers define proportional reasoning as well as look at the types of proportional reasoning problems that have emerged from their research. Based upon this research, I have developed a test instrument that will measure proportional reasoning among talented middle school girls.
The Need for the Study

Proportional reasoning has been widely studied over the years and researchers agree that it is a pivotal concept for students. In elementary school, students are introduced to additive and multiplicative strategies that lay the groundwork for formal proportional reasoning. This is typically referred to as pre-proportional reasoning and involves tasks of the form \( A + B = C + D \) and \( A \times B = C \times D \). Other examples are tasks that involve relationships of the form \( A = n + B \) and \( A = n \times B \). Behr et al. (1988) suggest unless there is evidence that reveals students recognize the structural similarity represented by both sides of the equation then they are not engaged in proportional reasoning. This is important for educators and teachers when attempting to unravel the misconceptions that often arise in mathematics classes.

Many see proportional reasoning as the foundation upon which many advanced mathematical concepts build. Behr, Lesh, and Post (1988) state, “proportional reasoning involves some of the most important understandings having to do with equivalence, variables, and transformations” (p. 97). Many of the concepts studied in Algebra, Geometry, and Calculus require students to reason proportionally. Furthermore, Karplus, Pulos and Stage (1983) emphasize the importance of effective proportional reasoning in science education. Rate of change, speed, density, revolutions of gears, and gas mileage are among some of the concepts that involve the use of ratios and proportional reasoning.

Finally, the National Council of Teachers of Mathematics (2000) assert, “the need to understand and be able to use mathematics in everyday life and in the workplace has never been greater and will continue to increase” (p.3). Proportional reasoning is a type of reasoning that students will most likely use in their profession and in everyday
situations. Some examples include: working with scale models and unitary rates, comparison-shopping, adjusting a recipe, diluting solutions in a biochemistry lab, or adjusting gears for a bike race. Proportional reasoning plays an important role in many real world scenarios, as well as in many advanced mathematical concepts.

Purpose of the Study

The purpose of the study is to measure proportional reasoning among fast-track middle school girls through the development of the Proportional Reasoning Assessment Instrument. Using the attendance roster from the summer 2000 Girls on Track program at Meredith College located in Raleigh, North Carolina, the researcher selected all 70 girls registered as subjects to which the Proportional Reasoning Assessment Instrument would be administered.

Research Questions

During the course of this research, the researcher investigated the following research questions:

1. Can a reliable instrument be constructed to measure proportional reasoning among fast-track middle school girls based upon the examination of research on proportional reasoning?
2. Can a valid instrument be constructed to measure proportional reasoning among fast-track middle school girls based upon the examination of research on proportional reasoning?
3. Through the administration of this instrument, can student misconceptions be identified?
Definition of Terms

The following terms used in this study are defined to help ensure clarity:


**Fast-Track Students** – “Students selected to take Algebra I in middle school” (S. Berenson, personal communication, March 13, 2001).

**Intensive Variable or ‘Rate’** – “The speed, shape, or other characteristic whose specification leads to a constant ratio relationship” (Karplus, et al. 1983, p.219).

**Objectives** – “Detailed descriptions of the specific outcomes that a test is designed to assess” (Gronlund, 1988).

**Proportional Reasoning** – “A term that denotes reasoning in a system of two variables between which there exists a linear functional relationship” (Karplus, et al., 1983, p.219).

**Test Reliability** – “The consistency, stability, and precision of test scores” (Gall, M., Borg & Gall, J., 1996, p.197).

**Test Validity** – “The appropriateness, meaningfulness, and usefulness of specific inferences made from test scores” (Gall, M., Borg & Gall, J., 1996, p.197).

Assumptions of the Study

The researcher made the following assumptions regarding the study:

1. The sample in this study is an adequate representation of the population of fast-track middle school girls in Wake County, North Carolina.
2. The students who were enrolled in the Girls on Track program completed the Proportional Reasoning Assessment Instrument.

3. The students responded to each question to the best of their ability.

Limitations of the Study

Several limitations existed in the study:

1. The study was limited to the population of fast-track middle school girls in Wake County, North Carolina.

2. The study was limited to students attending the Girls on Track summer program at Meredith College.

3. The study was limited to the problems on the Proportional Reasoning Assessment Instrument due to time constraints of the Girls on Track program. The conclusions from this study are most applicable for fast-track middle school girls similar to those in Wake County, North Carolina.

Summary

Proportional Reasoning is a form of mathematical reasoning that serves as the foundation for many mathematical concepts; therefore, it is essential that teachers be able to quickly assess the level at which their students reason proportionally. The focus of this study is to develop a valid and reliable instrument, based on the existing research, that will measure proportional reasoning among fast-track middle school girls.

Chapter Two describes the framework of the problems chosen for the Proportional Reasoning Assessment Instrument and outlines the basis for determining the reliability and validity of this instrument.
CHAPTER II
REVIEW OF LITERATURE

The goal of this study is to develop a reliable and valid instrument for measuring proportional reasoning among fast-track middle school girls in Wake County, North Carolina. This chapter presents the framework for the problems chosen for the Proportional Reasoning Assessment Instrument. It also explores the definition of proportional reasoning and introduces literature discussing the different types of proportional reasoning problems. Other topics discussed include how to determine the validity and reliability of a test instrument, along with how this information will be implemented in measuring the reliability and validity of this instrument. Finally, the significance of an item analysis will be presented.

Framework

Understanding Proportional Reasoning

The topic of proportional reasoning has been widely studied over the years and a detailed discussion of this form of mathematical reasoning is not the intention of this research. However, a few preliminary remarks must be considered. There have been many definitions of proportional reasoning published. I would like to present three definitions of proportional reasoning because of their relative nature in the later discussion of the types of proportional reasoning problems.

Karplus, Pulos, and Stage (1983) refer to proportional reasoning as “a term that denotes reasoning in a system of two variables between which there exists a linear functional relationship” (p.219). This linear functional relationship allows various
situations to be described in terms of a constant ratio. The use of a constant ratio is prevalent in Chemistry and Physics, for example, in problems associated with density or Ohm’s law. Lamon (1993) asserts, “proportional reasoning consists of being able to construct and algebraically solve proportions” (p.41). Finally, Behr et al. (1988) view proportional reasoning as “a form of mathematical reasoning that involves a sense of co-variation and of multiple comparisons, and the ability to store and process several pieces of information” (p.92). These definitions lay the groundwork for the discussion of the types of proportional reasoning problems chosen for the Proportional Reasoning Assessment Instrument.

Types of Proportional Reasoning Problems

According to available literature, there are a variety of proportional reasoning problems. In this section, I define the different types of proportional reasoning problems as they relate to the definitions discussed in the previous section.

As previously mentioned, Karplus, et al. (1983) report proportional reasoning centers around a linear functional relationship between two variables. According to these researchers, the process of proportional reasoning can be outlined in the following steps: “identification of two extensive variables that are applicable, recognition of the rate of intensive variables whose constancy determines the linear function, and application of the given data and relationships” (Karplus, et al., p.219). Karplus, et al. (1983) suggest that this process leads to two distinct types of problems: missing value problems and comparison problems. Missing value problems involve using the given data and relationships to determine the value for an extensive variable. Comparison problems
focus on comparing two values of the intensive variables that were computed from the data. Table 1 shows examples of problems that are indicative of each type of problem.

Table 2.1
Missing Value and Comparison Problems

<table>
<thead>
<tr>
<th>Type of Problem</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing Value</td>
<td>A car is driven 175 km in 3 hours. How far will it travel in 12 hours at the same speed?</td>
</tr>
<tr>
<td>Comparison</td>
<td>Car A is driven 180 km in 3 hours. Car B is driven 400 km in 7 hours. Which car was driven faster?</td>
</tr>
</tbody>
</table>


Gerald Noelting (1980) takes a slightly different approach with comparison problems. He works with mixture problems in which students compare two ratios of various concentrations of orange juice and water. With these experiments, orange juice flavor represents the intensive variable while glasses of water and orange juice concentrate are the extensive variables. According to Karplus, et al. (1983) his problems are unique in that they “require subjects to compare two ratios rather than requiring them to compute an answer that would produce a desired ratio” (p.48).

Lamon (1993) outlines a framework of semantic problem types breaking down the different ways students may encounter proportional reasoning problems. The first category of problems is called well-chunked measures. This type of problem “involves the comparison of two extensive measures, resulting in an intensive measure, or rate” (Lamon, 1993, p.42). The second category of problems involves the cardinality of a single subset being expressed in terms of the cardinalities of two or more subsets of which it is composed (Lamon, 1993). These types of problems are called part-part-
whole. Associated sets are proportion problems in which “the relationship between two elements is unknown or unclear unless their relationship is defined within the problem situation” (Lamon, 1993, p.42). Finally, stretchers and shrinkers involve the scaling up or down of one or more characteristics of a fixed ratio. More specifically, there is a one-to-one continuous ratio-preserving mapping that exists between two values that are representative of given characteristics (Lamon, 1993). Table 2 shows examples of problems associated with each semantic problem type.

Table 1.2

<table>
<thead>
<tr>
<th>Semantic Problem Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-Chunked Measures</td>
<td>The student is shown a page from a driver’s log book, recording total mileage at various intervals on a long trip. After 2, 5, 7, and 8 hours of driving, distance was recorded as 130 miles, 325 miles, 445 miles, and 510 miles, respectively. Was this truck driver traveling at a constant speed throughout the trip?</td>
</tr>
<tr>
<td>Part-Part-Whole</td>
<td>The student is shown a picture of two egg cartons, one containing a dozen eggs (8 white eggs to 4 brown eggs) and the other containing 1-½ dozen eggs (10 white eggs to 8 brown eggs). Which carton contains more brown eggs relative to white eggs?</td>
</tr>
<tr>
<td>Associated Sets</td>
<td>The student is shown a picture of 7 girls with 3 pizzas and 3 boys with 1 pizza. Who gets more pizza, the girls or the boys?</td>
</tr>
<tr>
<td>Stretcher/Shrinkers</td>
<td>The student is shown a picture of two trees. Tree A is 8 feet high and tree B is 10 feet high. This picture was taken 5 years ago. Today, tree A is 14 feet high and tree B is 16 feet high. Over the last five years, which tree’s height has increased most?</td>
</tr>
</tbody>
</table>

Note. From Lamon (1993), p.44.

Behr et al. (1988) assert, “the essential characteristics of proportional reasoning involves reasoning about the holistic relationship between two rational expressions such
as rates, ratios, quotients, and fractions” (p.93). Based upon this relationship, Behr et al. (1988) outline several key steps that must occur if a student is successfully engaging in proportional reasoning. First, the student must assimilate and synthesize the multifarious counterparts of pairs or series of rational expressions. Second, the student must conclude if disparity exists among expressions or if the expressions are equivalent. Third, the student must be able to determine missing values despite the numerical aspects of the problem scenario.

Based upon this detailed definition of proportional reasoning, Behr et al. (1988) define seven types of problems, some of which have been mentioned by previous researchers. These include: missing value problems, comparison problems, transformation problems, mean value problems, conversion problems, problems involving numerals with unit labels, and between-mode translation problems. Missing value and comparison problems hold the same definition as cited by Karplus, et al. Transformation problems can be categorized either as direction of change or transformations to produce equality. Mean value problems refer to problems involving geometric means as well as harmonic means. Conversion problems involve converting between ratios, rates, and fractions. Finally, between-mode translation problems involve having the student depict a given relationship in a representation system different from the original. Examples of each type of problem are included in Table 3.
Table 2.2  

Other types of proportional reasoning problems

<table>
<thead>
<tr>
<th>Type of Problem</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing Value</td>
<td>The student is told that the variables A, B, C, and D have the following relationship, ( \frac{A}{B} = \frac{C}{D} ). Given three of the four values, the student must determine the missing value.</td>
</tr>
<tr>
<td>Comparison</td>
<td>The student is shown ( \frac{A}{B} ), in which all four values of A, B, C, and D are given. The student must determine if ( \frac{A}{B} ) is &lt;, &gt;, or = to ( \frac{C}{D} ).</td>
</tr>
<tr>
<td>Transformation</td>
<td>Direction of Change: An equivalence of the form ( \frac{A}{B} = \frac{C}{D} ) is given. Then, one or two of the values of the A, B, C, or D is changed. The student must determine if the transformed value is &lt;, &gt;, or = to the original relationship.</td>
</tr>
<tr>
<td></td>
<td>Transformations to Produce Equality: An inequality is given in the form ( \frac{A}{B} &lt; \frac{C}{D} ). Then, A is increased by x amount. The student must determine x so that ( \frac{A + x}{B} = \frac{C}{D} ).</td>
</tr>
<tr>
<td>Mean Value</td>
<td>Geometric Mean: The student is shown the following relationship between A and B, ( \frac{A}{x} = \frac{x}{B} ). Then, they must determine the value for x so that the relationship is true.</td>
</tr>
<tr>
<td></td>
<td>Harmonic Mean: The student is shown the following relationship between A and B, ( \frac{A}{B} = \frac{(A - x)}{(x - B)} ). Then, they must determine the value of x so that the relationship is true.</td>
</tr>
<tr>
<td>Conversion</td>
<td>The ratio of students with blue eyes and those with green eyes is 17:21. What fraction of the students has green eyes?</td>
</tr>
<tr>
<td>Unit Labels</td>
<td>( \frac{3 \text{ feet}}{2 \text{ seconds}} = x \text{ miles per hour or } 5 \text{ feet/second} = x \text{ miles/hour} )</td>
</tr>
<tr>
<td>Between-Mode</td>
<td>Jane can ride her bike eight kilometers in two hours. Her ratio of kilometers per hour is: ( (a) \ 2: 8 \ (b) \ 4: 1 \ (c) \ 8: 2 \ (d) \not \text{ given} )</td>
</tr>
</tbody>
</table>

Note. From Behr et al. (1988), pp.95-96.
In this section, I have presented a variety of proportional reasoning problems that have emerged from relevant research. These problems provide an extensive base from which to develop an instrument that measures the proportional reasoning ability of middle school students.

Validity

When developing a test instrument, the researcher determines if the actual test meets the goals for which it was designed. In other words, the researcher must determine if the test is valid. Ravid (1994) states, “The validity of a test refers to the appropriateness of specific inferences and interpretations made using the test scores” (p.255). According to the Standards for Educational Psychological Testing, there are three types of validity: content validity, criterion-related ability, and construct validity” (AERA-APA_NCME, 1985, p. 9-18). Hopkins (1998) outlines the relationship between the specific inferences to be made and the type of validity. First, inferences pertaining to performance on a universe of items are associated with content validity. Second, inferences based on the performance on some criterion are associated with criterion-related validity. Third, inferences that relate to the degree to which certain psychological traits or constructs are actually represented by test performances correspond to construct validity.

Establishing validity is not a straightforward process and depends upon the purpose of the measure. Hopkins (1998) suggests, “the relevant type of validity in the measurement of academic achievement is content validity” (p.72). According to Ravid (1994), “content validity refers to the adequacy with which an instrument measures a representative sample of behaviors and content domain about which inferences are to be
Content validity involves examining the test items to ensure they represent the content to be tested, therefore, a numerical calculation of validity is not required. Hopkins (1998) indicates, “For a test to have high content validity, it should be a representative sample of both the content/topics and the cognitive process/abilities objectives of a given course or unit----the test should contain a representative sample of the content and uses to which the content is to be applied” (p.73).

In order to establish content validity, qualified professionals examine the test, its table of specifications and method development (Ravid, 1994). During the process of this investigation, three qualities of content validity must be determined: relevance, balance, and specificity (Doran, 1980). The quality of relevance involves looking at the behavior required to respond correctly to a test item and the purpose or objective in writing the item. Balance is found by determining the degree to which the proportion of items that are testing specific outcomes correspond to an “ideal” test. The quality of specificity is slightly more ambiguous. According to Doran (1980), “if subject matter experts should receive perfect scores (objectivity) then test-wise but course –naïve students should receive near chance scores, thus indicating course-specific learnings are being measured” (p.101).

While content validity is typically the relevant type of validity with regards to determining academic achievement, there are a few difficulties involved in its determination. First, content validity reflects the opinions of experts who may or may not be able to measure the extent to which test items reflect any likely attitude. Next, content validity is based upon the adequacy and knowledge of the initial selection of judges. Finally, when determining content validity, if you are only looking at the reactions of
respondents as reactions, and are working with reliable measurements, then they are measuring something consistently. The subjectivity comes into play when this something is being evaluated by a group of judges (Sax, 1968).

In the course of developing the Proportional Reasoning Assessment Instrument, a panel of experts consisting of mathematics education professors and graduate students evaluated the test items as a means of establishing content validity. Through the process of this evaluation, three issues were addressed. First, the panel reviewed the test items to ensure they reflect a representative sample of the various types of proportional reasoning problems found in relevant literature. Second, the panel evaluated the test items to ensure they reflect varying degrees of difficulty. Finally, the panel reviewed each item to ensure each question was worded appropriately to avoid misinterpretations. In addressing these issues, the panel of experts analyzed the qualities of relevance, balance, and specificity that are inherent in determining content validity of a test instrument.

A second form of validity was established for the Proportional Reasoning Assessment Instrument: criterion validity. Criterion validity refers to the relationship between scores on the test instrument and a defined external variable, or criterion (Kizzier, 1997). According to Popham (2000), criterion validity can be categorized into predictive or concurrent validity. The difference between the two involves the element of time. Predictive validation studies require a substantial time interval between the administration of the test instrument and the collection of criterion data, whereas no time interval is required for concurrent validation studies (Popham, 2000). More specifically, with concurrent validation, the criteria and the predictor are evaluated at the same point in time. The main question is to see if one measure can be substituted for the other.
(Hartmann, 2000). With both predictive and concurrent validation studies, it is important to select a legitimate criterion to ensure quality research.

One of the main objectives of the Girls on Track program is to explore the relationship between proportional reasoning and success in Algebra 1. Behr et al. (1988) assert proportional reasoning is vital to the success in middle-school algebra courses as well as more advanced high school mathematics. A study conducted by Clark and Berenson (2000) explores the connection between the scores on the Proportional Reasoning Assessment Instrument and the scores each of the 70 girls received on the State of North Carolina’s End-of-Grade (EOG) tests. The End-of-Grade tests was administered to the girls the year before they took Algebra I. For this reason, the validity study will follow a concurrent validation strategy using the scores on the End-of-Grade tests as the criterion of interest.

Reliability

When developing a test instrument, there are two areas of significant interest: validity and reliability. Doran (1980) asserts reliability relates to the “precision” of the instrument while validity relates to the “accuracy” of the instrument. Ravid (1994) explains, “The reliability of an assessment method is the extent to which the methods provide consistent, dependable, and stable information about the characteristics of students” (p.285).

There are three ways to approach determining the reliability, or consistency of an instrument. First, when looking at consistency across time, the reliability is measured in terms of stability. This method typically involves a test-retest administration of the instrument. Second, when consistency is described in terms of form, the reliability is
measured through equivalence. This is usually accomplished by administering two parallel forms of the instrument. Finally, internal consistency is a measure of reliability that is determined from only one administration of a test (Doran, 1980). Ravid (1994) states, “each item on a test can be viewed as a series of repeated measures” (p. 246).

Unlike the determination of content validity, a numerical value for reliability is given in terms of a reliability coefficient. The following chart outlines guidelines that have been accepted by most researchers in the educational arena.

Table 2.4

Guidelines for Interpreting Coefficient Alpha

<table>
<thead>
<tr>
<th>Reliability Coefficient</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>.95 - .99</td>
<td>Very High, Rarely Found</td>
</tr>
<tr>
<td>.90 - .95</td>
<td>High, Sufficient for Measurement of Individuals</td>
</tr>
<tr>
<td>.80 -.90</td>
<td>Fairly High, Possible for Measurement of Individuals</td>
</tr>
<tr>
<td>.70 -.80</td>
<td>Okay, Sufficient for Group Measurement, Not Individuals</td>
</tr>
<tr>
<td>Below .70</td>
<td>Low, Useful Only for Group Averages or Surveys</td>
</tr>
</tbody>
</table>

Note 1. From (Doran, 1980, p.104).

Ravid (1994) reports the major disadvantage with measuring reliability in terms of stability or equivalence is that the test has to be administered twice, whereas, the internal consistency approach allows the use of scores from only one test administration. For this reason, this measure of reliability is commonly reported in research. The internal consistency approach is based upon “the assumption that when a test measures a single basic concept, items correlate with each other and people who answer one item correctly are likely to correctly answer similar items” (Ravid, 1994, p.245).
There are several methods for determining reliability when taking the internal consistency approach. Among the most commonly used are: the split-half method, the Kuder-Richardson methods (KR-20 and KR-21), coefficient alpha and inter-rater reliability. For tests with various item formats and especially with instruments that involve a rating scale, like the Likert scale, Ravid (1994) suggests calculating coefficient alpha as the reliability measure. In addition, Ravid (1994) states, “Coefficient alpha is considered by researchers to provide good reliability estimates in most situations” (p.247).

Ravid (1994) suggests determining inter-rater reliability when the scoring of a test instrument involves the graders having to make subjective decisions. In his book *Practical Statistics for Educators*, Ravid (1994) discusses an example related to scoring essay tests. In this example, two or more readers read essays and then they assign a score on each criterion using a rating scale. “The scores from two or more essay readers can be used in two ways: (a) to compute a correlation coefficient, or (b) to compute the percentage of agreement. The correlation coefficient and the percentage of agreement indicate the reliability and the consistency of the measure as used by the judges. A high correlation coefficient shows consistency between the readers” (p.247-248).

I determined the reliability of the Proportional Reasoning Assessment Instrument by calculating two measures of internal consistency: coefficient-alpha and inter-rater reliability. I chose the coefficient alpha method due to the various formats of the test items and because I scored the items using a rating scale. SPSS statistical software package calculated this reliability coefficient. Due to the subjective nature in scoring the test items, I determined a measure of inter-rater reliability. I calculated the percent of
agreement between two graders based upon a sample of ten exams. Finally, Ravid (1994) suggests the standard error of measurement (SEM) provides additional information about the variability of the students’ scores. Therefore, I reported the SEM to further determine the reliability and accuracy of this test instrument.

Item Analysis

An item analysis provides further insight into the cohesiveness of the test instrument. Doran (1980) suggests, “Item and test analysis are the chores that teachers must tend to in order to ensure that their part of the testing and grading process is as valid and fair as possible” (p.95). A test instrument may appear satisfactory to the researcher or instructor, but may not be measuring the desired skills and objectives. According to Hopkins (1998), an item analysis reveals important insights into the student’s thinking and understanding of the material being assessed. As a result, possible student misconceptions may be identified. In addition, unsuspected defects of specific test items may be revealed and overall student performance can be used as a tool to provide feedback that may impact subsequent instruction.

Hopkins (1998) notes the “chief purpose of an item analysis is to determine the difficulty and the discrimination of each item” (p. 254). Computer programs easily compute values for these indices or there are several methods for determining these values through hand calculations. Doran (1980) defines the item difficulty index, $p$, as the proportion of a given sample choosing the correct response. The item discrimination index, $D$, indicates how frequently students performing well on the total test correctly answer an item and how frequently students performing poorly on the test incorrectly
answer an item (Popham, 2000). These two indices identify items that are ambiguous as well as provide insight into the reliability of the overall instrument (Hopkins, 1998).

Summary

In this chapter, I presented three definitions of proportional reasoning and introduced various types of proportional reasoning problems. In addition, I presented the methods of establishing reliability and validity of this test instrument as well as a brief introduction to item analysis.
CHAPTER III

METHODOLOGY

The purpose of the study was to measure proportional reasoning among fast-track middle school girls through the development of the Proportional Reasoning Assessment Instrument. In this chapter, the study’s research methodology is presented by discussing (a) the research design, (b) the sample, (c) the instrumentation, (d) data collection procedures, and (e) data analysis procedures.

Research Design

This instrument will be used in a longitudinal study conducted by the Center for Research in Mathematics and Science Education at North Carolina State University. There was a need to be able to analyze the cognitive strategies employed by the students. For this reason, I chose an open-ended test as the research design for this study. The Proportional Reasoning Assessment Instrument consisted of ten questions in which the students were asked to show all of their work.

The Sample

The sample for this study consisted of 70 girls who attended the summer 2000 Girls on Track program at Meredith College located in Raleigh, North Carolina. The grade level for each of the girls ranged from 6th grade through 8th grade for the 2000-2001 school year. The cohort leaders administered the exam to the girls on the third day of camp, July 5, 2000. The girls were instructed to read each question carefully and to show all of their work. Participating teachers collected the exams with a 100% completion rate.
Instrumentation

The instrument used in this study was the Proportional Reasoning Assessment Instrument (see Appendix B). I developed this instrument from problems discussed in relevant literature. Karplus, et al. (1983) reported two distinct types of proportional reasoning problems: missing value and comparison. Noelting (1980) also studied comparison problems; however, he focused on problems involving mixtures. Lamon (1993) outlines a framework of four semantic problem types: well-chunked measures, part-part-whole, associated sets, and stretchers and shrinkers. Finally, Behr, et al. (1988) outline seven types of proportional reasoning problems: missing value, comparison, transformation, mean-value, conversion, unit labels, and between-mode translation problems.

The first step in developing the Proportional Reasoning Assessment Instrument was to choose a representative sample of these various proportional reasoning problems. Professors and graduate students reviewed the test to determine the level of difficulty of the questions as well as the clarity. A panel of experts reviewed the test items to ensure they were an adequate representation of the various types of proportional reasoning problems. In addition, I administered the test to two girls who attended the summer 1999 Girls on Track program. The two girls commented on the wording of each test item and the difficulty of each item (see Appendix E).

There were two factors affecting the format of the test instrument. First, the test instrument needed to give some insight into the cognitive strategies of students. For this reason, an open-ended test format was chosen. Second, there were time constraints. Test
administration could not take longer than thirty minutes. For these reasons, the test instrument consisted of ten items of varying difficulty levels.

Item 1 of the Proportional Reasoning Assessment Instrument is a comparison problem taken from research by Karplus, et al. (1983). The context of the problem involves two girls purchasing pieces of gum. Based upon given information, students determine which girl bought the cheaper gum. This item is similar to the lemonade puzzles previously used in research by Karplus, et al (1983) and involves the use of familiar settings, small values, discrete quantities (pieces) and a quasi-continuous quantity (money). Behr, et al. (1988) also identified comparison problems.

Item 2 is similar to the missing value problems presented by Karplus (1983). The context of this problem involves a boy making coffee. Based upon the fact that it takes 8 cups of water to make 14 small cups of coffee, the boy must determine how many small cups of coffee can be made with 12 cups of water. Once again, this item revolves around a setting familiar to most students and small, discrete quantities.

Item 3 is an associated sets problem taken from research by Lamon (1993). The context of the problem involves students determining who gets the most pizza between a group of boys and a group of girls. This problem requires students to use written representations rather than symbolic representations to arrive at the solution. Once again, the problem is placed in a setting familiar to most students.

Item 4 is a part-part-whole problem that is also taken from research by Lamon (1993). The context of this problem involves two cartons of eggs. One carton contains a dozen eggs and the other contains a dozen and a half eggs. In addition, each carton contains both brown and white eggs. The students must decide which carton contains
more brown eggs relative to white eggs. Based upon comments from the two student reviewers and from colleagues, I modified the item from its original version. Unlike the original problem; however, the students are given both a written description of the scenario and a picture of the two egg cartons.

Items 5, 6, and 7 are mixtures problems taken from research by Noelting (1980). The context of the problems involves making orange juice for a party. The item shows two trays containing various amounts of orange juice and water. I present the information both pictorially and as an ordered pair. The student must determine which tray will create a drink that has a stronger orange taste or if the two drinks will taste the same. Item 5 has the amount of orange juice the same with the amount of water different. Item 6 has different amounts of orange juice and water for each tray. The ratios of orange juice and water are not of equivalent classes. Item 7 uses amounts of orange juice and water that represent equivalent classes.

Item 8 involves both comparisons and graphical interpretation. I developed this item through collaboration with colleagues. The context of the problem centers on a bike ride. The students are shown a graph of a bike journey divided into three intervals. The x-axis represents the variable time and the y-axis represents the variable distance. The students must deduce from the graph how fast the girl is traveling in each time interval.

Item 9 can be classified as a “stretcher” and is similar to a problem presented by Lamon (1993). The context of the problem involves two trees. Given information about the height of the two trees, students must determine which tree has increased the most relative to its initial height. For this problem, students are given only a written description of the scenario.
Item 10 can also be classified as a stretcher and is similar to the problems posed by Lamon (1993) in her research. This problem is slightly different than question 9 in that the scaling up of the given quantities is non-linear. The context of this problem involves constructing flags. The students are given information about the dimensions of flag 1 and are told that flag 2 needs to be 3 feet longer. They must determine how much cloth is needed to construct flag 2 while maintaining the same ratio of length to height as flag 1. I present the information both as a written description and a table. This item resembles the comparison problems as discussed by Karplus, et al. (1983) and Behr, et al. (1988).

Data Collection

There were 70 girls enrolled in the summer 2000 Girls On Track program at Meredith College in Raleigh, North Carolina. Each of these girls participated in the research study. The cohort leaders administered the *Proportional Reasoning Assessment Instrument* to the students on July 5, 2000. It was important for future research that the exam be administered during the first few days of the summer program before the girls received further instruction on the topic of proportional reasoning. Program leaders divided the girls into their cohorts where half took the exam in one room and the other group took the exam in a different room. One cohort leader in each group administered the exam. Therefore, there were only two teachers involved in administering the test. The students were instructed to read each question carefully and to show all of their work. Upon completion, the cohort leader collected the exams. I collected the exams from the cohort leaders later that afternoon. Prior to implementing these procedures, approval was granted from Gary A. Mirka, Ph.D., Chairman for the NCSU Institutional
Review Board on Research Involving Human Subjects at North Carolina State University, Raleigh, North Carolina (see Appendix A).

Analysis of Data

I analyzed information received from the Proportional Reasoning Assessment Instrument using Statistical Package for the Social Sciences Version 10.0 (SPSS) and EXCEL. I used a four-point grading rubric to score each test item (see Appendix D). In this section, I will discuss the methods for determining reliability, validity, difficulty, and discrimination analysis.

Reliability

I determined the reliability of the Proportional Reasoning Assessment Instrument by calculating two measures of internal consistency: Chrombach’s coefficient-alpha and inter-rater reliability. I used SPSS to calculate the coefficient alpha reliability coefficient. I determined inter-rater reliability by calculating the percent of agreement between two graders based upon a random sample of ten exams.

Content Validity

A panel of experts consisting of professors and graduate students evaluated the test items in an effort to establish content validity. First, the panel reviewed the test item to ensure the items reflect a representative sample of the various types of proportional reasoning problems identified in relevant literature. Second, the panel evaluated the test items to ensure they reflect varying degrees of difficulty. Finally, the panel reviewed the individual test items to ensure each question was worded appropriately to avoid misinterpretations. In addressing these issues, the panel analyzed the three qualities of
relevance, balance, and specificity that are inherent in determining content validity of a test instrument.

**Criterion Validity**

I used a study conducted by Clark and Berenson (2000) to establish criterion validity. For this study, the researchers explore the connection between the girls’ scores on the State of North Carolina End-of-Grade test and their scores on the **Proportional Reasoning Assessment Instrument**. The girls took the End-of-Grade test during the spring prior to attending the Girls on Track summer program. The scores on this test represent the criterion used in establishing validity. The cohort leaders administered the **Proportional Reasoning Assessment Instrument** on July 5, 2000. I calculated the correlation between the girls’ scores on the End-of-Grade test and the **Proportional Reasoning Assessment Instrument**.

**Item Analysis- Difficulty**

Difficulty is directly related to discrimination. If the difficulty is either too high or too low, it becomes almost impossible to differentiate among students. According to Doran (1980), the difficulty of an item is often expressed as \( P = \frac{R}{N} \) where \( R \) represents the number of students who answer the item correctly and \( N \) represents the total number of students (p.96). Because of the four-point grading rubric used for scoring this test, \( R \) represents the number of students scoring a 4 on the test item. I used the following procedure outlined by Hopkins (1998) to determine item difficulty for the **Proportional Reasoning Assessment Instrument**.

1. Determine each student’s raw score.
2. Order the N papers in descending order.
3. Take the highest third of the tests and label the “high group”.

4. Take the lowest third of the tests and label the “low group”.

5. Determine the proportion, \( p_H \), of students from the “high group” scoring a 4 on the test item.

6. Determine the proportion, \( p_L \), of students from the “low group” scoring a 4 on the test item.

7. Item difficulty is found by averaging \( p_H \) and \( p_L \). The formula is
\[
p = \frac{p_H + p_L}{2}.
\]

The Proportional Reasoning Assessment Instrument was administered to seventy girls; therefore, the scores cannot be divided evenly into thirds. For the purpose of this analysis, I chose to have the highest twenty-three scores represent the “high group” and the lowest twenty-three scores represent the “low group.” This leaves twenty-four scores for the middle group. I chose this method so both the high and low groups would contain the same number of scores.

Under ideal circumstances, the difficulty of each test item should be halfway between the random guessing score and 100% (Doran, 1980). When the difficulty is at this optimal level, discrimination between students is more straightforward. There are several criteria used for interpreting item difficulty. Hopkins (1998) suggests, “the maximum measurement of individual differences by an item is at a maximum when the item difficulty is .5” (p.257). Others suggest the range of .4 to .6 for item difficulty (Doran, 1980). For the purposes of this evaluation, I chose to use a table that presents a range of values for item difficulty that combines moderately difficulty items with those possessing
difficulty at the extremes. Table 3.1 presents the guidelines used in interpreting difficulty indices found for each test item.

Table 3.1

Guidelines for interpreting difficulty indices

<table>
<thead>
<tr>
<th>Index of Difficulty</th>
<th>Item Difficulty Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>.85 to 1.00</td>
<td>Very Easy</td>
</tr>
<tr>
<td>.60 to .85</td>
<td>Moderately Easy</td>
</tr>
<tr>
<td>.35 to .60</td>
<td>Moderately Difficult</td>
</tr>
<tr>
<td>.00 to .35</td>
<td>Very Difficult</td>
</tr>
</tbody>
</table>

Note. From Doran (1980), p.97

Item Analysis - Discrimination

The discrimination index is sometimes called the validity index and represents a measure of how well an item distinguishes between the high students and the low students (Doran, 1980). Beuchart and Mendoza (1979) assert, “A standard index of item discrimination is the coefficient of correlation of the examinees’ scores on an item with their total scores on the rest of the test” (p.116). However, I used the High and Low Third System (Doran, 1980, p.98) outlined in the following steps to calculate discrimination for the Proportional Reasoning Assessment Instrument:

1. Order the N scores in descending order.

2. Take the highest third of the tests and label the “high group”.

3. Take the lowest third of the tests and label the “low group”.

4. Item discrimination is found by the following formula: 
   \[ D = \frac{(H - L)}{N/2}, \]
   where H is the number of students in the high group scoring a 4 on the test item, L is the number of students in the low group scoring a 4 on the test item, and N is the total number of students in the sample.
I divided the students into high and low groups using the same procedure outlined in the previous section. Beuchart & Mendoza (1979) found the differences in results from using D versus the coefficient of correlation are virtually nonexistent. According to Doran (1980), one of the main attributes of this method is that the discrimination index falls between the limits of +1.00 and –1.00. Test items that have negative discrimination values should be closely analyzed because these items do not contribute to the overall reliability of the test. Table 3.2 outlines a guide for interpreting discrimination when the sample size is at least 30.

Table 3.2

**Guidelines for interpreting discrimination indices**

<table>
<thead>
<tr>
<th>Index of Discrimination</th>
<th>Item Discrimination Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>.40 and up</td>
<td>Excellent Discrimination</td>
</tr>
<tr>
<td>.30 to .39</td>
<td>Good Discrimination</td>
</tr>
<tr>
<td>.10 to .29</td>
<td>Fair Discrimination</td>
</tr>
<tr>
<td>.01 to .10</td>
<td>Poor Discrimination</td>
</tr>
<tr>
<td>Negative</td>
<td>Item may be miskeyed or intrinsically ambiguous</td>
</tr>
</tbody>
</table>


**Summary**

In this chapter, I presented the study’s research methodology. More specifically, I discussed the research design, the sample, the instrumentation, data collection, procedures, and data analysis. I developed the Proportional Reasoning Assessment Instrument for the purpose of measuring proportional reasoning among fast-track middle school girls. The sample for this study comprised 70 girls who attended the summer 2000 Girls On Track program at Meredith College located in Raleigh, North Carolina. Experts determined content validity through a review of the test instrument. I calculated
two measures of internal consistency: Chrombach’s alpha reliability coefficient and inter-rater reliability. I used SPSS to calculate coefficient alpha and to perform other data analysis procedures. Finally, I discussed procedures for determining item difficulty and item discrimination.
CHAPTER IV

RESULTS

The purpose of the study was to measure proportional reasoning among fast-track middle school girls through the development of the Proportional Reasoning Assessment Instrument. The sample for this study comprised 70 girls who attended the summer 2000 Girls on Track program at Meredith College located in Raleigh, North Carolina. Cohort leaders administered the test to each girl on the third day of camp, July 5, 2000. In this section, I will present the results of reliability, validity, item difficulty, and item discrimination analyses.

Preliminary Findings

I used SPSS and EXCEL to perform data analysis of the Proportional Reasoning Assessment Instrument. Table 4.1 presents the preliminary findings. The overall mean score on the 70 exams is 27.66 with a standard deviation of 5.8. The item means range in value from 1.83 to 3.79 with the overall average being 2.77. This gives a range in the scores of 1.96 with a standard deviation of 0.50. The average inter-item correlation is .18.
Table 4.1

Item means and variances for the Proportional Reasoning Assessment Instrument

<table>
<thead>
<tr>
<th>Test Item</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.79</td>
<td>0.74</td>
<td>0.55</td>
</tr>
<tr>
<td>2</td>
<td>3.00</td>
<td>1.31</td>
<td>1.71</td>
</tr>
<tr>
<td>3</td>
<td>2.87</td>
<td>1.20</td>
<td>1.45</td>
</tr>
<tr>
<td>4</td>
<td>2.66</td>
<td>1.14</td>
<td>1.30</td>
</tr>
<tr>
<td>5</td>
<td>2.66</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>6</td>
<td>2.89</td>
<td>1.20</td>
<td>1.44</td>
</tr>
<tr>
<td>7</td>
<td>3.01</td>
<td>1.15</td>
<td>1.32</td>
</tr>
<tr>
<td>8</td>
<td>2.59</td>
<td>1.20</td>
<td>1.43</td>
</tr>
<tr>
<td>9</td>
<td>1.83</td>
<td>0.70</td>
<td>0.49</td>
</tr>
<tr>
<td>10</td>
<td>2.37</td>
<td>1.24</td>
<td>1.54</td>
</tr>
</tbody>
</table>

Research Findings

Reliability

According to Ravid (1994), “the reliability of an assessment method is the extent to which the methods provides consistent, dependable, and stable information about the characteristics of students” (p.285). Under ideal situations, a test instrument should produce the same results each time it is administered to the same population. I determined the reliability of the Proportional Reasoning Assessment Instrument by calculating two measures of internal consistency: coefficient-alpha and inter-rater reliability.
Coefficient alpha ranges in value from 0.0 to 1.0. A reliability coefficient of 1.0 would indicate the test produces consistent results with no measurement error. SPSS reported a coefficient alpha for the Proportional Reasoning Assessment Instrument as \( \alpha = .71 \) using a sample size of 70 students.

I calculated inter-rater reliability due to the subjective nature in the scoring of the test items. Two graders each scored a random sample of ten exams for a total of 100 test items. I determined the percent of agreement between the two graders to be 96%, meaning 96 out of the 100 test items received the same score.

**Content Validity**

A panel of experts conducted an extensive review of the items on the Proportional Reasoning Assessment Instrument in an effort to establish content validity. The panel found the instrument to contain a representative sample of proportional reasoning problems. Also, the panel found the test items to represent a variety of difficulty levels. I confirmed this assertion through conducting preliminary interviews of girls attending the 1999 Girls on Track summer program (see Appendix E). Finally, the panel reviewed each test item for clarity. Suggestions were made to improve the overall clarity especially for items 8 and 9.

**Criterion Validity**

I used a study conducted by Clark and Berenson (2000) to establish criterion validity. They calculated the correlation between the scores on the Proportional Reliability Assessment Instrument and the scores on the State of North Carolina End-of-Grade tests. The correlation is \( r = .69 \).
Item Analysis – Difficulty

I performed an analysis of item difficulty using all available data. Table 4.2 and Figure 4.1 present the results.

Table 4.2

<table>
<thead>
<tr>
<th>Test Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_H$</td>
<td>23</td>
<td>22</td>
<td>19</td>
<td>15</td>
<td>13</td>
<td>21</td>
<td>22</td>
<td>11</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>$p_L$</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>$p$</td>
<td>.57</td>
<td>.61</td>
<td>.46</td>
<td>.37</td>
<td>.28</td>
<td>.50</td>
<td>.52</td>
<td>.28</td>
<td>.07</td>
<td>.35</td>
</tr>
</tbody>
</table>

Figure 4.1- Bar graph of item difficulty for Proportional Reasoning Assessment Instrument

The average difficulty is $p_{mean} = .40$ suggesting the overall test is moderately difficult. Analysis revealed items 5, 8, and 9 to be very difficult with p-values ranging from .07 to .28. These p-values indicate that students may possess strong
misconceptions. I analyze these items in greater detail in the next chapter. Items 1, 3, 4, 6, 7, and 10 are moderately difficult with p-values ranging from .35 to .57. Item 2 with a p-value of .61 can be labeled as moderately easy. As you can see, none of the test items were found to be very easy. This analysis reveals 7 of the 10 items have difficulty indices ranging from .35 to .60. This suggests the test items will be good discriminators.

Item Analysis – Discrimination

The discrimination index is a measure of how well the test instrument differentiates between students who did well on the test and students whose performance was poor. I used all available data to determine the discrimination indices for the Proportional Reasoning Assessment Instrument. Table 4.3 and Figure 4.2 present the results.

Table 4.3

<table>
<thead>
<tr>
<th>Test Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_H )</td>
<td>23</td>
<td>22</td>
<td>19</td>
<td>15</td>
<td>13</td>
<td>21</td>
<td>22</td>
<td>11</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>( n_L )</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>( D )</td>
<td>.86</td>
<td>.69</td>
<td>.73</td>
<td>.56</td>
<td>.56</td>
<td>.81</td>
<td>.86</td>
<td>.39</td>
<td>.13</td>
<td>.51</td>
</tr>
</tbody>
</table>
Figure 4.2 – Bar graph of item discrimination for Proportional Reasoning Assessment Instrument

The average discrimination is $D_{mean} = .61$. Based on the guidelines outlined in the previous chapter, the test instrument has excellent discrimination. Item 8 has good discrimination and item 9 is considered to have fair discrimination. All of the other items are considered to be excellent discriminators. It is noted items 8 and 9 have lower discrimination indices when compared to the rest of the items. These test items were also considered to be more difficult and will be looked at more closely in the next chapter.

Summary

In this chapter, I presented the results of the reliability, validity and item analyses. The overall mean score on the 70 exams was 27.66 with a standard deviation of 5.8. The average item mean was 2.77 and the average inter-item correlation was .18. Coefficient alpha was $\alpha = .71$ which is sufficient for group measurement. Inter-rater reliability was determined to be 96%. I established content and criterion validity. The average item
difficulty was $p_{\text{mean}} = .40$ and the average item discrimination was $D_{\text{mean}} = .61$. These values indicate the test instrument is moderately difficult and has good discrimination.
CHAPTER V
SUMMARY, CONCLUSIONS, DISCUSSION, AND RECOMMENDATIONS

The purpose of this study was to measure proportional reasoning among fast-track middle school girls through the development of the Proportional Reasoning Assessment Instrument. The following research questions provided focus for the study:

1. Can a reliable instrument be constructed to measure proportional reasoning among fast-track middle school girls based upon the examination of research on proportional reasoning?

2. Can a valid instrument be constructed to measure proportional reasoning among fast-track middle school girls based upon the examination of research on proportional reasoning?

3. Through the administration of this instrument, can student misconceptions be identified?

In this chapter, I discuss the results from the reliability, validity, and item analyses. In addition, I present recommendations for practice and future research.

Summary of Procedures

The sample for this study consisted of 70 girls who attended the summer 2000 Girls on Track program at Meredith College located in Raleigh, North Carolina. The grade level for each of the girls ranged from 6th grade through 8th grade for the 2000-2001 school year. Cohort leaders administered the Proportional Reasoning Assessment Instrument on July 5, 2000. Program leaders divided the seventy girls into cohorts and half took the exam in one room and the other half in a different room. Cohort leaders administered the exam instructing students to read each question carefully and to show all
of their work. The two cohort leaders collected the exams with a 100% completion rate.

I used a four-point rating scale to score each test item.

I analyzed the data from the Proportional Reasoning Assessment Instrument using Statistical Package for the Social Sciences Version 10 (SPSS) and EXCEL. I determined reliability calculating two measures of internal consistency: Chronbach’s coefficient alpha and inter-rater reliability. I established content validity by having a panel of experts evaluate the test items based on relevance, balance, and specificity. I used a study conducted by Clark and Berenson (2000) to establish criterion validity. In this study, the researchers calculated the correlation between the test scores on the Proportional Reasoning Assessment Instrument and the students’ scores on the State of North Carolina End-of-Grade tests. I calculated item difficulty by arranging the raw scores in descending order and then dividing the scores into thirds. The average of the proportion of students in the high group scoring a four on the item with the proportion of students in the low group scoring a four on the item represents the difficulty index. I determined item discrimination using the High and Low Third System (Doran, 1980). I calculated item means, item variances and inter-item correlation using SPSS.

Summary of Findings

I obtained the data from the 70 students enrolled in the summer 2000 Girls on Track program at Meredith College located in Raleigh, North Carolina. Table 5.1 summarizes the frequencies and percentages of the students’ scores for each of the ten items on the Proportional Reasoning Assessment Instrument. Results of data analysis from the individual exams revealed that (a) the Proportional Reasoning Assessment Instrument is a reliable instrument for measuring proportional reasoning among fast-track
middle school girls; (b) the Proportional Reasoning Assessment Instrument is a valid instrument for measuring proportional reasoning among fast-track middle school girls; and (c) the instrument revealed common misconceptions among the students.

Table 5.1

Summary of frequencies and percentages of student scores for each item

<table>
<thead>
<tr>
<th>Test Item</th>
<th>Scale 1</th>
<th>Scale 2</th>
<th>Scale 3</th>
<th>Scale 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f</td>
<td>p</td>
<td>f</td>
<td>p</td>
</tr>
<tr>
<td>1 (fs/c)</td>
<td>4</td>
<td>5.71</td>
<td>1</td>
<td>1.43</td>
</tr>
<tr>
<td>2 (mv)</td>
<td>19</td>
<td>27.14</td>
<td>2</td>
<td>2.86</td>
</tr>
<tr>
<td>3 (as)</td>
<td>11</td>
<td>15.71</td>
<td>22</td>
<td>31.43</td>
</tr>
<tr>
<td>4 (ppw)</td>
<td>10</td>
<td>14.29</td>
<td>31</td>
<td>44.29</td>
</tr>
<tr>
<td>5 (m-1)</td>
<td>2</td>
<td>2.86</td>
<td>44</td>
<td>62.86</td>
</tr>
<tr>
<td>6 (m-2)</td>
<td>10</td>
<td>14.29</td>
<td>24</td>
<td>34.29</td>
</tr>
<tr>
<td>7 (m-3)</td>
<td>7</td>
<td>10</td>
<td>24</td>
<td>34.29</td>
</tr>
<tr>
<td>8 (c/g)</td>
<td>21</td>
<td>30</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>9 (s)</td>
<td>20</td>
<td>28.57</td>
<td>46</td>
<td>65.71</td>
</tr>
<tr>
<td>10 (c/s)</td>
<td>22</td>
<td>31.43</td>
<td>23</td>
<td>32.86</td>
</tr>
</tbody>
</table>

Note. See Appendix C for a key of problem types.
There are a few test items that stand out. Ninety-one percent (64) of the students scored a 4 on item 1, while only 5\% (4) scored a 4 on item nine. It is interesting to note that item 1 is a comparison problem involving familiar settings with discrete, quasi-continuous quantities. On the three mixture problems, items 5, 6, 7, there were no students who scored a 3. It appears from looking at the table, the students either answered these questions completely or possessed a misconception or applied an incorrect strategy in solving the problem. Over 27\% (19) of the students scored a 1 on item 2, which is a missing value problem involving discrete quantities. Thirty percent (21) of the students scored a 1 on item 8, which is a graphical interpretation problem involving comparisons. Items 9 and 10 were labeled “stretchers” with over 28\% (20) of the students scoring a 1 on item 9 and 31\% (22) scoring a 1 on item 10. With such large percentages of students scoring a 1 on items 2, 8, 9, and 10, this suggests these students possess misconceptions related to these items or applied an incorrect strategy to solve the problems.

There were two measures of internal consistency computed for the Proportional Reasoning Assessment Instrument: Chrombach’s coefficient alpha and inter-rater reliability. The overall coefficient alpha for this instrument is $\alpha = .71$. According to Doran (1980), an alpha between .70 and .80 is sufficient for group measurement, but not for individuals. Therefore, this instrument is appropriate for drawing inferences about a group as a whole, not an individual student. Currently, this instrument has only been tested on fast-track middle school girls. I found inter-rater agreement to be 96\%; that is, the two raters scored the exams exactly the same 96 percent of the time. This suggests
the grading rubric (see Appendix D) provides enough detail to yield consistent results or scores.

Table 5.2 shows a summary of the reliability analysis for the Proportional Reasoning Assessment Instrument. The last three columns are the most integral in drawing conclusions about the instrument’s reliability. The Corrected-Item-Total Correlation shows the correlation between the respective item and the total sum score without the respective item. The Squared Multiple Correlation shows the correlation between the respective item and all of the others. The Alpha if Item Deleted shows the internal consistency of the scale (coefficient alpha) if the respective item would be deleted.

Table 5.2

Summary of statistical reliability analysis for the Proportional Reasoning Assessment Instrument

<table>
<thead>
<tr>
<th>Variable</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Corrected Item-Total Correlation</th>
<th>Squared Multiple Correlation</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>23.87</td>
<td>31.04</td>
<td>0.25</td>
<td>0.26</td>
<td>0.70</td>
</tr>
<tr>
<td>Item 2</td>
<td>24.66</td>
<td>24.87</td>
<td>0.54</td>
<td>0.38</td>
<td>0.65</td>
</tr>
<tr>
<td>Item 3</td>
<td>24.79</td>
<td>28.81</td>
<td>0.27</td>
<td>0.21</td>
<td>0.70</td>
</tr>
<tr>
<td>Item 4</td>
<td>25.00</td>
<td>28.87</td>
<td>0.29</td>
<td>0.31</td>
<td>0.70</td>
</tr>
<tr>
<td>Item 5</td>
<td>25.00</td>
<td>28.90</td>
<td>0.36</td>
<td>0.45</td>
<td>0.69</td>
</tr>
<tr>
<td>Item 6</td>
<td>24.77</td>
<td>24.44</td>
<td>0.66</td>
<td>0.68</td>
<td>0.63</td>
</tr>
<tr>
<td>Item 7</td>
<td>24.64</td>
<td>24.23</td>
<td>0.72</td>
<td>0.68</td>
<td>0.62</td>
</tr>
<tr>
<td>Item 8</td>
<td>25.07</td>
<td>28.15</td>
<td>0.32</td>
<td>0.19</td>
<td>0.69</td>
</tr>
<tr>
<td>Item 9</td>
<td>25.83</td>
<td>33.45</td>
<td>-0.03</td>
<td>0.16</td>
<td>0.73</td>
</tr>
<tr>
<td>Item 10</td>
<td>25.29</td>
<td>28.87</td>
<td>0.24</td>
<td>0.13</td>
<td>0.71</td>
</tr>
</tbody>
</table>
Item 9 stands out in terms of not being consistent with the rest of the scale. Its correlation with the sum scale is -.03, while all of the others correlate at .24 or better.

Looking at the right-most column, the reliability of the test instrument would increase to .73 if this item were deleted. While this would increase the reliability of the test, it would only be a slight increase. Therefore, you may not want to delete it from the instrument.

I computed item difficulty and item discrimination indices for each test item. The average difficulty is $p_{\text{mean}} = .40$ and the average discrimination is $D_{\text{mean}} = .61$. Table 5.3 shows a summary of the difficulty and discrimination indices for each test item.

Table 5.3

<table>
<thead>
<tr>
<th>Item Number</th>
<th>Group</th>
<th>Number of Correct Responses (n =23)</th>
<th>Item Discrimination</th>
<th>Item Difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H</td>
<td>23</td>
<td>0.86</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>H</td>
<td>22</td>
<td>0.69</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>H</td>
<td>19</td>
<td>0.73</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>H</td>
<td>15</td>
<td>0.56</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>H</td>
<td>13</td>
<td>0.56</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>H</td>
<td>21</td>
<td>0.81</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>H</td>
<td>22</td>
<td>0.86</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>H</td>
<td>11</td>
<td>0.39</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>H</td>
<td>3</td>
<td>0.13</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>H</td>
<td>14</td>
<td>0.51</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
There are 9 items possessing a discrimination index above .35 suggesting the overall test discriminates nicely between the high and low groups. In fact, 8 of the 10 questions have excellent discrimination, possessing indices greater than .51. There is one item labeled a good discriminators and only 1 item considered to be a fair discriminator. There are no poor discriminators on the instrument. It is noted that item 9 is a fair discriminator among students. Cross-referencing this item with its corresponding difficulty index reveals item 9 to be a very difficult item.

Figure 5.1 shows the relationship between item difficulty and item discrimination for the Proportional Reasoning Assessment Instrument. The median D-values for the 4 very difficult items ($p \leq .35$) and the 5 moderately difficult items (.35 < $p$ ≤ .60) are .45 and .50, respectively. There was only one item labeled moderately easy (.60 < $p$ ≤ .85) and its corresponding D-value is .61. All of the items contribute to the reliability of the test instrument due to the absence of negative D-values. The average D-value for this test is $D = .61$. This suggests that the mean total score of the high group is about 61% greater than the mean total score of the low group.
Figure 5.1 – Item difficulty versus item discrimination

Conclusions

From the findings of this study, I made the following conclusions:

1. The Proportional Reasoning Assessment Instrument is a sufficiently reliable instrument for group measurement among fast-track middle school girls.

2. The Proportional Reasoning Assessment Instrument is a valid instrument for measuring proportional reasoning among fast-track middle school girls.

3. Item 9 is a very difficult item and would increase alpha from $\alpha = .71$ to $\alpha = .73$ if deleted from the instrument.

4. Overall, the Proportional Reasoning Assessment Instrument possesses an appropriate level of difficulty.

5. Overall, the Proportional Reasoning Assessment Instrument discriminates among students.
6. The items on the Proportional Reasoning Assessment Instrument contribute to the central purpose of the instrument due to the fact there are no negative discriminators.

Discussion

The purpose of this study was to measure proportional reasoning among fast-track middle school girls through the development of the Proportional Reasoning Assessment Instrument. Results of the study show this instrument is both valid and reliable for group measurement among talented girls and it possesses an appropriate level of difficulty and discrimination. In addition, careful analysis of the instrument reveals common misconceptions among the students, particularly for items 8, 9, and 10 as well as an interesting result for item 5. Finally, I will discuss results from a study conducted by Clark and Berenson (2000) as they relate to the establishment of criterion validity.

Item 8 is a problem developed by the researcher involving comparisons and graphical interpretation. The students see a graph of Sarah’s bike journey that is divided into three intervals. The x-axis represents the variable time and the y-axis represents the variable distance. The students are asked to deduce from the graph how fast the girl is traveling in each time interval. Figure 5.2 shows the graph of the bicycle journey as seen on the test instrument.
Thirty percent of students scored a 1 on item 8. According to the grading rubric (see Appendix D), these students may have possessed a misconception, applied an incorrect strategy, failed to show their work, or obtained a completely incorrect answer. This item is labeled very difficult with a difficulty index of $p = .28$. The discrimination index for this item is $D = .39$ making it a good discriminator. Although this value is relatively high, it is still lower than the other discrimination indices. This indicates the possible presence of student misconceptions. Careful review of the individual exams confirmed this assertion.

There were only 20 students (28.6%) who scored a 4 on this item indicating they possess complete understanding of the concept and applied a correct strategy. Among the remaining 50 students, the most common misconception occurred in the interpretation of interval B. Most of the students said Sarah traveled at a constant speed during this interval instead of realizing Sarah was actually stationary. Only a few students realized Sarah traveled faster in interval C than in interval A. Also, several students assigned
numerical values to the graph in order to determine their solution. Appendix F shows a sample of student responses.

Item 9 is adapted from Lamon (1993) and is classified as a stretcher. The students are given information about the height of two trees. They must determine which tree has increased the most relative to its initial height. For this problem, students are only given a written description of the scenario. Figure 5.3 gives the details of this item.

| Two trees were measured five years ago. Tree A was 8 feet high and tree B was 10 feet high. Today, tree A is 14 feet high and tree B is 16 feet high. Over the last five years, which tree’s height increased the most relative to its initial height? Show any calculations that lead you to your answer. |

Figure 5.3 – Item 9 of the Proportional Reasoning Assessment Instrument

Around 28% (20) of students scored a 1 on item 9. As previously stated, this indicates these students may have applied an incorrect strategy, possessed a misconception, showed no work, or obtained a completely incorrect answer. It is interesting to note 46 students (65.7%) scored a 2 on this item and there were no students scoring a 3. According to the grading rubric (see Appendix D), a score of 2 implies the student may have some understanding of the topic or a misconception, applied an incorrect strategy, showed no work, or obtained a correct answer possibly through guessing. A score of 3 means the student demonstrated understanding of the topic and applied a correct strategy, however, obtained an incorrect answer possibly due to a math error.

Item 9 has a difficulty index of $p = .07$ and a discrimination index of $D = .13$. According to these values, this item is very difficult and a fair discriminator. The low discrimination index signals either the presence of misconceptions among the students or
lack of clarity in the test item. Review of student responses to this item reveal the presence of two main misconceptions among the students.

There were only 4 students (5.7%) who scored a 4 on this item indicating they understand the concept and applied a correct strategy to obtain the answer. There were two common incorrect responses to this problem (see Appendix G). The first was that Tree B grew the most relative to its initial height because it was taller in the end. Basically, the students found out each tree grew a total of 6 feet. Because both trees increased by the same amount, students assumed tree B grew by a larger percentage because 16 is larger than 14. The other common answer was both trees grew equally. This is partially correct. Students answering in this manner received a score of 2 provided they showed appropriate strategies for obtaining the solution. Both trees grew by 6 feet; however, that does not completely answer the question. There were only a few students who figured out that tree A increased by 75% and tree B increased by 60%.

Item 10 is similar to problems posed by Lamon (1993) and can also be classified as a stretcher. This problem is different from item 9 in that the scaling up of given quantities is non-linear. This item revolves around the context of constructing flags. Students are given the dimensions of flag 1 and are told that flag 2 needs to be 3 feet longer. They are asked to determine how much cloth is needed if they are to construct flag 2 while maintaining the same ratio of length to height as flag 1. Figure 5.4 shows the table of information as seen on the test instrument.
<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Height</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flag 1</td>
<td>3 ft</td>
<td>2 ft</td>
<td>6 ft²</td>
</tr>
<tr>
<td>Flag 2</td>
<td>6 ft</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.4 – Item 10 of the Proportional Reasoning Assessment Instrument

Around 31% (22) of students scored a 1 on item 10 indicating these students may have applied an incorrect strategy, possessed a misconception, showed no work, or obtained a completely incorrect answer. It is interesting to note 23 students (32.9%) scored a 2 on the test, suggesting these students may have some understanding of the topic, possess a misconception, applied an incorrect strategy, showed no work, or obtained a correct answer possibly through guessing. Review of actual student responses reveal many students received a score of 2 because they did not show any work.

While 23 students (32.9%) scored a 4 on item 10, several misconceptions were present among those students receiving a score of 1 or 2 (see Appendix H). Among those scoring a 2, the most common misconception was found in calculating the area of flag 2. The students correctly realized the height of flag 2 is twice the height of flag 1, however, incorrectly concluded the area of flag 2 must also be twice the area of flag 1. These students failed to realize area increases in a non-linear manner. Among those scoring a 1, the most common misconception is students applied an incorrect additive strategy. These students found the length of flag 2 by adding 3 to the length of flag 1. They repeated this process to determine the height of flag 2. Most of these students correctly found the area by multiplying the length and height to obtain an area of flag 2 of 30 ft².

Item 5 is part of a collection of mixture problems adapted from research by Noelting (1980). Students are shown quantities of orange juice and water both pictorially
and as an ordered pair. Item 5 has the amount of orange juice the same with the amount of water different; therefore, students are comparing second terms. Figure 5.5 outlines the context of these mixture problems and shows item 5 as seen on the test instrument.

You and your friend are going to make orange juice for a party. You will be given three different situations. For each situation, you will be presented with the contents of two trays. Each tray contains various amounts of orange juice and water. The shaded box represents the orange juice and the unshaded box represents the water. The goal for each is to determine which drink will have the stronger orange taste or if the two drinks will taste the same. Each mixture will be expressed as an ordered pair (e.g. (1, 3)) with the first term corresponding to the number of glasses of orange juice and the second term to the number of glasses of water. Show any calculations and explain your thinking.

<table>
<thead>
<tr>
<th>Tray A</th>
<th>Tray B</th>
</tr>
</thead>
<tbody>
<tr>
<td>□ □ □</td>
<td>□ □ □ □ □</td>
</tr>
</tbody>
</table>

(1,2) -Vs- (1,5)

Figure 5.5 – Item 5 on the Proportional Reasoning Assessment Instrument

Item 5 has a difficulty index of $p = .28$ and a discrimination index of $D = .56$, making it a very difficult item yet an excellent discriminator. Due to the high discrimination index, it is unlikely there are strong misconceptions among the students. Results of the study show 44 students (62.9%) scored a 2 on this item. According to the grading rubric (see Appendix D), these students may have some understanding of the topic, a misconception, applied an incorrect strategy, showed no work, or obtained a correct answer possibly through guessing. Careful analysis reveals most students receiving this score failed to show their work and therefore received a score of 2. Because of this, it is impossible to determine if misconceptions are present, if students
obtained the correct answer through guessing, or if they found the problem easy enough to answer without showing their work.

In addition to the misconceptions revealed through the item analysis and careful review of student responses, I explored the study conducted by Clark and Berenson (2000). Clark and Berenson (2000) calculated the correlation between the scores on the Proportional Reasoning Assessment Instrument and the scores on the State of North Carolina End-of-Grade test. This correlation was used to establish criterion validity. The correlation coefficient between the two sets of scores is $r = .69$. This value indicates a strong, positive correlation. According to Clark and Berenson (2000), this strong association supports the assertion by the Girls on Track program that differences in proportional reasoning skills explain a significant portion of variation in mathematical achievement among middle school students.

Recommendations for Practice

From the findings of this study, the following recommendations for practice are made:

1. Teachers may use the Proportional Reasoning Assessment Instrument before beginning a unit on ratios and proportion. This may provide teachers with insight regarding potentially weak areas among their group of students.

2. Teachers may give the Proportional Reasoning Assessment Instrument at the end of a unit on ratios and proportion to identify possible misconceptions present among their group of students or areas where students may need additional instruction.
Recommendations for Future Research

From the findings of this study, the following recommendations for future research are made:

1. Future research in this area could be field-testing a new version of item 9. Careful analysis revealed students either possessed a misconception or applied an incorrect strategy to answer this problem. In addition, there were many students who gave only a partial solution. Rewording the item may improve the overall clarity of the problem. This may increase the discrimination index for this item and increase the instrument’s overall reliability.

2. Future research in this area could be expanding the Proportional Reasoning Assessment Instrument to include more test items. This would increase the reliability of this instrument.

3. Future research in this area could be administering the Proportional Reasoning Assessment Instrument to a larger population of students including students of different genders and varying ability levels. This would improve reliability and possibly allow for both group and individual measurement.

Summary

In this chapter I have discussed the results of the reliability, validity, and item analyses. I highlighted items 5, 8, 9, and 10 and presented student misconceptions on these items. Finally, I outlined recommendations for future practice and research.
REFERENCES

Behr, M., Lesh, R. & Post, T.: 1988, Proportional reasoning. In M. Behr and J. Hiebert (Eds.), Number concepts and operations in the middle grades, Lawrence Erlbaum Associates, Reston, VA.


APPENDICES
APPENDIX A

Approval Letter for Research on Human Subjects
Gary A. Mirka, PhD, Chairman
Institutional Review Board
on Research Involving Human Subjects

Date: April 10, 2000

Project Title: Girls on Track: Increasing Middle Grades Girls' Interest in Math-Related Careers by Engaging Them in Computer-Based Mathematical Explorations of Urban Problems in their Communities

IRB#: 1479h

Dear Dr. Berenson,

Approval for the project listed above has been extended (through April 10, 2001).

NOTE:
1. This committee complies with requirements found in Title 45 part 46 of The Code of Federal Regulations. For NCSU the Assurance Number is: M1263; the IRB Number is: 01XM.
2. Review de novo of this proposal is necessary if any significant alterations/additions are made.
3. For those projects that require an extension beyond the one year limitation of the IRB approval, the principle investigator must submit a letter to the IRB Chair stating their intention to continue the research as outlined in the NCSU IRB Policies and Procedures Document.

Sincerely,

[Signature]
APPENDIX B

Proportional Reasoning Assessment Instrument
1. Sally bought 3 pieces of gum for 12 cents and Anna bought 5 pieces of gum for 20 cents. Who bought the cheaper gum or were they equal? 
   **Show all of your work.**

2. To make coffee, David needs exactly 8 cups of water to make 14 small cups of coffee. How many small cups of coffee can he make with 12 cups of water? 
   **Show all of your work.**
3. There are 7 girls with 3 pizzas and 3 boys with 1 pizza. Who gets more pizza, the girls or the boys? *Show all of your work.*

4. There are two egg cartons. The shaded circles represent brown eggs and the unshaded circles represent white eggs. The blue carton contains 8 white eggs and 4 brown eggs. The red carton contains 10 white eggs and 8 brown eggs. Which carton contains more brown eggs relative to white eggs? Explain your thinking.

- Blue Carton
- Red Carton
You and your friend are going to make orange juice for a party. You will be given three different situations. For each situation, you will be presented with the contents of two trays. Each tray contains various amounts of orange juice and water. The shaded box represents the orange juice and the unshaded box represents the water. The goal for each is to determine which drink will have the stronger orange taste or if the two drinks will taste the same. Each mixture will be expressed as an ordered pair (e.g. (1, 3)) with the first term corresponding to the number of glasses of orange juice and the second term to the number of glasses of water. Show any calculations and explain your thinking.

5.

<table>
<thead>
<tr>
<th>Tray A</th>
<th>Tray B</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,2)</td>
<td>-Vs-</td>
</tr>
</tbody>
</table>

6.

<table>
<thead>
<tr>
<th>Tray A</th>
<th>Tray B</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2,5)</td>
<td>-Vs-</td>
</tr>
</tbody>
</table>

7.

<table>
<thead>
<tr>
<th>Tray A</th>
<th>Tray B</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2,3)</td>
<td>-Vs-</td>
</tr>
</tbody>
</table>
8. Sarah took a bike ride this weekend. Below is a graph of her journey. The variable labeled Distance represents the distance Sarah is away from her starting point and the variable Time represents the amount of time that has passed since she began her journey. The graph is divided into three intervals: A, B and C. What information can you deduce from the graph about how fast she was traveling in each interval?
9. Two trees were measured five years ago. Tree A was 8 feet high and tree B was 10 feet high. Today, tree A is 14 feet high and tree B is 16 feet high. Over the last five years, which tree’s height increased the most relative to its initial height? Show any calculations that lead you to your answer.

10. You are shown a flag that measures 3 feet in length and 2 feet in height. It uses 6 square feet of cloth. If you wanted to make it 3 feet longer while maintaining the same ratio of length to height, how much cloth would you need? Show your work.

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Height</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flag 1</td>
<td>3 ft</td>
<td>2 ft</td>
<td>6 ft²</td>
</tr>
<tr>
<td>Flag 2</td>
<td>6 ft</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C

Key of Problem Types
# Key of Problem Types

<table>
<thead>
<tr>
<th>Test Item</th>
<th>Symbol</th>
<th>Problem Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>fs/c</td>
<td>familiar settings; comparison</td>
</tr>
<tr>
<td>2</td>
<td>mv</td>
<td>missing value</td>
</tr>
<tr>
<td>3</td>
<td>as</td>
<td>associated sets</td>
</tr>
<tr>
<td>4</td>
<td>ppw</td>
<td>part-part-whole</td>
</tr>
<tr>
<td>5</td>
<td>m-1</td>
<td>mixture; like first terms with comparison of second terms</td>
</tr>
<tr>
<td>6</td>
<td>m-2</td>
<td>mixture; any ratio</td>
</tr>
<tr>
<td>7</td>
<td>m-3</td>
<td>mixture; equivalence of any ratio</td>
</tr>
<tr>
<td>8</td>
<td>c/g</td>
<td>comparison; graphical interpretation</td>
</tr>
<tr>
<td>9</td>
<td>s</td>
<td>stretcher</td>
</tr>
<tr>
<td>10</td>
<td>c/s</td>
<td>comparison; non-linear stretcher</td>
</tr>
</tbody>
</table>
APPENDIX D

4-Point Grading Rubric
<table>
<thead>
<tr>
<th>Score</th>
<th>Details</th>
</tr>
</thead>
</table>
| 4     | Demonstrates understanding of the concept  
        Applies an appropriate strategy to solve the problem  
        Obtains the correct answer |
| 3     | Demonstrates understanding of the concept  
        Applies an appropriate strategy to solve the problem  
        Obtain an incorrect answer possibly due to a math error |
| 2     | Possesses some understanding of the concept or has a misconception  
        Applies an incorrect strategy to solve the problem or shows no work  
        Incomplete Answer or obtained the correct answer possibly through guessing |
| 1     | Possesses a misconception  
        Applies an incorrect strategy to solve the problem or shows no work  
        Obtains an incorrect answer |
APPENDIX E

Comments From Preliminary Interviews
1. Sally bought 3 pieces of gum for 12 cents and Anna bought 5 pieces of gum for 20 cents. Who bought the cheaper gum or were they equal? Show the calculations that lead you to your answer.

   Sally...Anna...Equal

   Is this necessary?

   Caroline: Found the highlighted words confusing; what’s their purpose
            Easier than #2

   Kristina: No comment
            Easy

2. To make coffee, David needs exactly 8 cups of water to make 14 small cups of coffee. How many small cups of coffee can he make with 12 cups of water? Show the calculations that lead you to your answer.

   Caroline: Straightforward
            Easy

   Kristina: No comment
            Easy
3. There are 7 girls with 3 pizzas and 3 boys with 1 pizza. Who gets more pizza, the girls or the boys? Explain your thinking.

Caroline: No comment
More difficult than the first 2 problems

Kristina: Not good with fractions
3 boys : 1 pizza          7 girls : 3 pizzas
6 boys : 2 pizzas         (More for girls)
Medium

4. There are two egg cartons. The shaded circles represent brown eggs and the unshaded circles represent white eggs. The blue carton contains a dozen eggs consisting of 8 white eggs and 4 brown eggs. The red carton contains one and a half dozen eggs consisting of 10 white eggs and 8 brown eggs. Which carton contains more brown eggs relative to white eggs? Explain your thinking.

Blue Carton
Red Carton

Caroline: Leave out “dozen” and “one-half dozen” b/c its redundant
Help to make it less wordy
Same difficulty as #3

Kristina: 4/12 = 1/3
8/18 = 4/9
“more fractions;” unsure how to compare these
drew pictures to arrive at the red carton having more brown eggs
hard
You and your friend are going to make orange juice for a party. You will be given three different situations. For each situation, you will be presented with the contents of two trays. Each tray contains various amounts of orange juice and water. The shaded box represents the orange juice and the unshaded box represents the water. The goal for each is to determine which drink will have the stronger orange taste or if the two drinks will taste the same. Each mixture will be expressed as an ordered pair (e.g. (1, 3)) with the first term corresponding to the number of glasses of orange juice and the second term to the number of glasses of water. Show any calculations and explain your thinking. Are the instructions too wordy?

5.  
   
   (1,2)  -Vs-  (1,5)

Caroline: medium (of three)

Kristina: same amount of orange; second tray has more water; easy

6.  
   
   (4,1)  -Vs-  (1,4)

Caroline: easiest (of three); like the gum problem

Kristina: same amount of everything (ingredients); easy

7.  
   
   (2,3)  -Vs-  (4,6)

Caroline: Lot of words; What do you mean by trays? Maybe draw trays; Are the OJ and water mixed together? Use a key to explain shaded boxes? None of them too hard

Kristina: same because (2,3) = (4,6); easy
8. Sarah took a bike ride this weekend. Below is a graph of her journey. The variable labeled Distance represents the distance Sarah is away from her starting point and the variable Time represents the amount of time that has passed since she began her journey. The graph is divided into three intervals: A, B and C. Compare the relationship between distance and time for each interval.

This is by far the most difficult problem. Should we emphasize “how fast was she traveling” during each of the three intervals?

Caroline: What do you do for the problem? Explain the graph?
Harder (maybe because there are no units.)

Kristina: I just don’t understand the graph; Is it a hill or how far she has gone? I don’t know how to put it into words.
Hard!
9. Two trees were measured five years ago. Tree A was 8 feet high and tree B was 10 feet high. Today, tree A is 14 feet high and tree B is 16 feet high. Over the last five years, which tree’s height has increased the most? Show any calculations that lead you to your answer.

Which increased by the greater %? Is what we really want to test? Caroline said this would make the problem harder.

Caroline: All you had to do is subtract; had to take numbers and think
Really easy but harder than #1

Kristina: No comment; noticed she set up proportion incorrectly
Medium

10. You are shown a flag that measures 3 feet in length and 2 feet in height. It uses 6 square feet of cloth. If you wanted to make it 3 feet longer while maintaining the same ratio of length to height, how much cloth would you need? Show your work.

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Height</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flag 1</td>
<td>3 ft</td>
<td>2 ft</td>
<td>6 ft²</td>
</tr>
<tr>
<td>Flag 2</td>
<td>6 ft</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Caroline: Maybe add “maintaining the same proportion of length to height.”
Pretty easy

Kristina: Just doubled height and applied area formula to get answer
Pretty easy
APPENDIX F

Sample of Student Responses to Item 8
Sample of Correct Responses

8. Sarah took a bike ride this weekend. Below is a graph of the distance she traveled during the duration of her journey. The variable labeled Distance represents the distance Sarah is away from her starting point and the variable Time represents the amount of time that has passed since she began her journey. The graph is divided into three equal time intervals: A, B and C. What information can you deduce from the graph about how fast she was traveling during each time interval? Show all of your work.

Distance

A medium speed

she must have been sitting on a park bench

C fastest speed

Time
8. Sarah took a bike ride this weekend. Below is a graph of the distance she traveled during the duration of her journey. The variable labeled **Distance** represents the distance Sarah is away from her starting point and the variable **Time** represents the amount of time that has passed since she began her journey. The graph is divided into three equal time intervals: A, B and C. What information can you deduce from the graph about how fast she was traveling during each time interval? Show all of your work.

At Interval A Sarah increased her distance as time went by.
At Interval B Sarah showed where she was and did not increase her distance as time went by.
At Interval C Sarah increased her distance more than she did at Interval A.
3. Sarah took a bike ride this weekend. Below is a graph of the distance she traveled during the duration of her journey. The variable labeled Distance represents the distance Sarah is away from her starting point and the variable Time represents the amount of time that has passed since she began her journey. The graph is divided into three equal time intervals: A, B and C. What information can you deduce from the graph about how fast she was traveling during each time interval? Show all of your work.

She started at zero and traveled a few miles in time A. Then she didn't go anywhere in time B. Then she traveled more miles to get to her stopping point at the end of time C.

\[
24 \div 3 = 8
\]

A, B, C all = 8 hrs each
8. Sarah took a bike ride this weekend. Below is a graph of the distance she traveled during the duration of her journey. The variable labeled Distance represents the distance Sarah is away from her starting point and the variable Time represents the amount of time that has passed since she began her journey. The graph is divided into three equal time intervals: A, B and C. What information can you deduce from the graph about how fast she was traveling during each time interval? Show all of your work.

Intervals A and C, her speed either increased or decreased because her distance and time were climbing up.

Intervals B she was going the same speed each time since neither the time or distance increased or decreased.
8. Sarah took a bike ride this weekend. Below is a graph of the distance she traveled during the duration of her journey. The variable labeled Distance represents the distance Sarah is away from her starting point and the variable Time represents the amount of time that has passed since she began her journey. The graph is divided into three equal time intervals: A, B and C. What information can you deduce from the graph about how fast she was traveling during each time interval? Show all of your work.

\[
\begin{align*}
\text{Distance} & \quad \text{Time} \\
\hline
A & \quad B & \quad C
\end{align*}
\]

- In interval B she was probably traveling fast since she went in a straight line.
- In interval A & C she went slower because it looked like she was going down something.
8. Sarah took a bike ride this weekend. Below is a graph of the distance she traveled during the duration of her journey. The variable labeled Distance represents the distance Sarah is away from her starting point and the variable Time represents the amount of time that has passed since she began her journey. The graph is divided into three equal time intervals: A, B and C. What information can you deduce from the graph about how fast she was traveling during each time interval?
Show all of your work.

Distance

A: there was an increase in speed
B: she had a steady pace the whole time
C: major increase in speed
8. Sarah took a bike ride this weekend. Below is a graph of the distance she traveled during the duration of her journey. The variable labeled Distance represents the distance Sarah is away from her starting point and the variable Time represents the amount of time that has passed since she began her journey. The graph is divided into three equal time intervals: A, B and C. What information can you deduce from the graph about how fast she was traveling during each time interval?

Show all of your work.
APPENDIX G

Sample of Student Responses to Item 9
9. Two trees were measured five years ago. Tree A was 8 feet high and tree B was 10 feet high. Today, tree A is 14 feet high and tree B is 16 feet high. Over the last five years, which tree's height increased the most relative to its initial height?

Show all of your work.
Incorrect Responses

9. Two trees were measured five years ago. Tree A was 8 feet high and tree B was 10 feet high. Today, tree A is 14 feet high and tree B is 16 feet high. Over the last five years, which tree's height increased the most relative to its initial height? Show all of your work.

\[
\frac{8}{14} = \frac{16}{16}
\]

Both of them increased the same amount most relative to its initial height.

9. Two trees were measured five years ago. Tree A was 8 feet high and tree B was 10 feet high. Today, tree A is 14 feet high and tree B is 16 feet high. Over the last five years, which tree's height increased the most relative to its initial height? Show all of your work.

Tree A

\[
\begin{align*}
A &= 8 \text{ ft} + 6 \\
A &= 14 \text{ ft}
\end{align*}
\]

Tree B

\[
\begin{align*}
B &= 10 \text{ ft} + 6 \\
B &= 16 \text{ ft}
\end{align*}
\]

They are both
9. Two trees were measured five years ago. Tree A was 8 feet high and tree B was 10 feet high. Today, tree A is 14 feet high and tree B is 16 feet high. Over the last five years, which tree’s height increased the most relative to its initial height? Show all of your work.

\[
\begin{align*}
A &: 8 = 14 \\
B &: 10 - 16
\end{align*}
\]

They both grew the same.

\[
\begin{align*}
\frac{14 - 8}{8} &= \frac{16 - 10}{16}
\end{align*}
\]
9. Two trees were measured five years ago. Tree A was 8 feet high and tree B was 10 feet high. Today, tree A is 14 feet high and tree B is 16 feet high. Over the last five years, which tree's height increased the most relative to its initial height? Show all of your work.

A: 8 to 14

B: 10 to 16

Both trees' height increased by the same amount, 6, over the past 5 years.

9. Two trees were measured five years ago. Tree A was 8 feet high and tree B was 10 feet high. Today, tree A is 14 feet high and tree B is 16 feet high. Over the last five years, which tree's height increased the most relative to its initial height? Show all of your work.

Tree A

-8 ft

-14 ft

Tree B

-10 ft

16 ft

The trees are equal

+6
APPENDIX H

Sample of Student Response to Item 10
Correct Response

10. You are shown a flag that measures 3 feet in length and 2 feet in height. It uses 6 square feet of cloth. If you wanted to make it 3 feet longer while maintaining the same ratio of length to height, how much cloth would you need? Show all of your work.

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Height</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flag 1</td>
<td>3 ft</td>
<td>2 ft</td>
<td>6 ft²</td>
</tr>
<tr>
<td>Flag 2</td>
<td>6 ft</td>
<td>4 ft</td>
<td>24 ft²</td>
</tr>
</tbody>
</table>

\[
\frac{3}{2} = \frac{6}{x} \Rightarrow x = 4
\]

10. You are shown a flag that measures 3 feet in length and 2 feet in height. It uses 6 square feet of cloth. If you wanted to make it 3 feet longer while maintaining the same ratio of length to height, how much cloth would you need? Show all of your work.

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Height</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flag 1</td>
<td>3 ft</td>
<td>2 ft</td>
<td>6 ft²</td>
</tr>
<tr>
<td>Flag 2</td>
<td>6 ft</td>
<td>4 ft</td>
<td>24 ft²</td>
</tr>
</tbody>
</table>
Incorrect Responses

10. You are shown a flag that measures 3 feet in length and 2 feet in height. It uses 6 square feet of cloth. If you wanted to make it 3 feet longer while maintaining the same ratio of length to height, how much cloth would you need? Show all of your work.

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Height</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flag 1</td>
<td>3 ft</td>
<td>2 ft</td>
<td>6 ft²</td>
</tr>
<tr>
<td>Flag 2</td>
<td>6 ft</td>
<td>5 ft</td>
<td>30 ft²</td>
</tr>
</tbody>
</table>

\[
\frac{3}{x} = \frac{2}{4}
\]
10. You are shown a flag that measures 3 feet in length and 2 feet in height. It uses 6 square feet of cloth. If you wanted to make it 3 feet longer while maintaining the same ratio of length to height, how much cloth would you need? Show all of your work.

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Height</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flag 1</td>
<td>3 ft</td>
<td>2 ft</td>
<td>6 ft²</td>
</tr>
<tr>
<td>Flag 2</td>
<td>6 ft</td>
<td>2 ft</td>
<td>12 ft²</td>
</tr>
</tbody>
</table>

10. You are shown a flag that measures 3 feet in length and 2 feet in height. It uses 6 square feet of cloth. If you wanted to make it 3 feet longer while maintaining the same ratio of length to height, how much cloth would you need? Show all of your work.

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Height</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flag 1</td>
<td>3 ft</td>
<td>2 ft</td>
<td>6 ft²</td>
</tr>
<tr>
<td>Flag 2</td>
<td>6 ft</td>
<td>2 ft</td>
<td>12 ft²</td>
</tr>
</tbody>
</table>

10. You are shown a flag that measures 3 feet in length and 2 feet in height. It uses 6 square feet of cloth. If you wanted to make it 3 feet longer while maintaining the same ratio of length to height, how much cloth would you need? Show all of your work.

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Height</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flag 1</td>
<td>3 ft</td>
<td>2 ft</td>
<td>6 ft²</td>
</tr>
<tr>
<td>Flag 2</td>
<td>6 ft</td>
<td>2 ft</td>
<td>12 ft²</td>
</tr>
</tbody>
</table>
10. You are shown a flag that measures 3 feet in length and 2 feet in height. It uses 6 square feet of cloth. If you wanted to make it 3 feet longer while maintaining the same ratio of length to height, how much cloth would you need? Show all of your work.

<table>
<thead>
<tr>
<th>Flag</th>
<th>Length</th>
<th>Height</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flag 1</td>
<td>3 ft</td>
<td>2 ft</td>
<td>6 ft²</td>
</tr>
<tr>
<td>Flag 2</td>
<td>6 ft</td>
<td>5 ft</td>
<td>30 ft²</td>
</tr>
</tbody>
</table>

\[ \frac{3}{5} \times 3 = 6 \]

\[ 6 + 5 = 30 \]