2022-03-18

Variability in Variables: An Analysis of Variable Use in A Middle School Mathematics Textbook

Annalisa G. Lavin

Honors Thesis

VARIABILITY IN VARIABLES: AN ANALYSIS OF VARIABLE USE IN A MIDDLE SCHOOL MATHEMATICS TEXTBOOK

by

Annalisa Lavin

Submitted to Brigham Young University in partial fulfillment of graduation requirements for University Honors

Mathematics Education Department

Brigham Young University

April 2022

Advisor: Daniel Siebert

Honors Coordinator: Keith Leatham

ABSTRACT

VARIABILITY IN VARIABLES: AN ANALYSIS OF VARIABLE USE IN A

MIDDLE SCHOOL MATHEMATICS TEXTBOOK

Annalisa Lavin

Mathematics Education Department

Bachelor of Science

Variables are an important part of mathematics as they allow mathematicians to generalize and more easily express mathematical ideas, however research has shown that students struggle with understanding variables. One way to study variable use is to classify variables into different types based on how they are used. This study uses variable types to investigate variable usage in a middle school mathematics textbook. This study identified 9 variable types and 3 algebraic processes that variable types are tied to. This analysis also found evidence of a growing complexity in variable use, starting with sophisticated variable use in grade 6 and growing from there.

ACKNOWLEDGEMENTS

Dedicated to my ever-patient family, loving husband, and kind friends who supported and encouraged through this entire project. I wouldn't have made it through with you all.

Many thanks to Dr. Dan Siebert for his extensive help on this project and his tireless pursuit of making mathematics education better.

This project is also dedicated to all the educators who got me to where I am today. Thank you for inspiring me to keep learning every day.

TABLE OF CONTENTS

LIST OF FIGURES

I. Introduction

Middle school mathematics is an important transitionatory time for students. It is at this point that students begin to move away from basic arithmetic and focus on more complex algebraic topics. Understanding algebra and being confident in algebraic skills is essential to student success as they progress to high school and higher education. However, there are many facets of algebra which may impede student progression and success.

Algebra presents many unique challenges, one of which is the use of variables. Using symbols to represent values and performing operations on them is a different and new skill set for students. Students may struggle with variable use because there are many different reasons and ways that variables are used in algebra. The meaning of variables changes based on context and how the reader interprets how the variable is being used. For example, the variable x is being used very differently when an author asks for the reader to solve for x then when the author uses x to express a property like $x + y = y + x$. These differences are rarely, if ever, addressed when teaching algebra, so students are left to decode the differences on their own and create their own interpretations of variable types.

Despite the fact that variables have been identified by researchers as an area where middle school students struggle, variable usage remains understudied (Carpenter, Franke, & Levi, 2003; Usiskin, 1988). This lack of research puts both students and educators at a disadvantage. Variables give mathematicians the power to generalize, making them extremely important to use but also very difficult to understand. The abstract nature of variables can create a barrier between higher mathematics and students who are developing fluency with mathematical symbols and expressions. Mathematics teachers know how and when to use different types of variables, but this knowledge may be mostly implicit. The different uses may not have been made explicit to them, so teachers do not think to make these uses explicit to students. Highlighting the different uses through research will assist teachers in making these important mathematical distinctions clear to their students.

II. Literature Review

Some of the most notable research done on variable types is by Usiskin (1988) in his article "Conceptions of School Algebra and Uses of Variables." This literature review outlines Usiskin's four conceptions of algebra and the variable types he and other researchers have identified, and an example of the complexity of variable usage.

Different Algebra Conceptions

Usiskin (1988) argued that algebra cannot be easily defined because there are several conceptions of algebra, each with its own purposes and set of relevant procedures, that make up algebra. These different conceptions impact the importance and use of variables.

The first conception is "algebra as a generalized arithmetic" (p. 11). In this conception of algebra, the focus is on generalizing patterns in mathematics and representing these patterns in terms of variables. Representing patterns can also be used in mathematical modeling. Here variables can be used to express a pattern that is observed in the real world and to predict future outcomes. Usiskin points out that the key instructions given to students when involved in this conception of algebra are to *translate* and *generalize*. For example, the pattern seen in $-1 \cdot 5 = -5$ and $-2 \cdot 5 = -10$ can be generalized as $-x \cdot y = -xy$. Competency with this conception allows learners to take sentences in English and express them in algebraic symbols in order to make patterns and quantitative relationships clearer and easier to follow.

The next conception that Usiskin identifies is “algebra as a study of procedures for solving certain kinds of problems” (p. 12). This conception of algebra is focused on taking algebraic equations involving an unknown quantity and arriving at a solution. The key instructions are to *simplify* and *solve*. An example of a typical problem seen in this conception is solving $5x + 3 = 40$. As learners engage in simplifying and solving, they create a series of equivalent equations or expressions until an answer or most simplified form is reached.

The third conception is “algebra as the study of relationships among quantities” (p. 13). With this conception of algebra, the emphasis is on describing relationships. An important distinction in this conception is that the value of variables can and will vary. For instance, questions like “what happens to the value of $\frac{1}{x}$ as x gets larger and larger?” are investigated in this conception. The key instructions are to *relate* and *graph.* Usiskin argues that this is the only conception where independent and dependent variables exist.

The last conception identified by Usiskin is “algebra as the study of structures” (p. 15). This conception of algebra is most clearly seen in advanced college courses such as abstract algebra but can also be seen in high school algebra as well. This view of algebra is used when students are asked to analyze the structure of algebraic expressions to factor, expand, or simplify the expressions. These terms *factor*, *expand*, or *simplify* are also common instructions associated with this conception. For example, “factor $3x^2 + 4ax - 132a^2$” is a typical problem type. When using this conception of algebra, variables are often viewed as arbitrary symbols because the learner does not think about variables as representing numeric values, but rather as “marks on paper” to be manipulated using a set of widely accepted procedures.

Different Types of Variables

Usiskin (1988) identifies several types of variables, each tied to specific conceptions of algebra. Other researchers, such as Blanton et al. (2011), have also identified several types of variables, but have not connected these types to conceptions of algebra or purposes for using algebra as clearly as Usiskin. Moreover, the variable types identified by these other scholars have an equivalent in the types that Usiskin describes. Thus, as I describe variable types, I limit my discussion to Usiskin’s work. The variable types identified by Usiskin appear in Table 1 along with descriptions and examples pulled from his paper.

The first type of variable Usiskin (1988) describes is a *pattern generalizer*, which is used under the first conception of algebra. This type of variable is used to express algebraic properties or patterns and is also used to translate numerical sentences from English to algebra. Usiskin considers variables used to represent real world quantities when mathematically modeling to also be pattern generalizers. This type of variable is similar to Blanton’s et al.’s (2011) definition of variables representing numbers in generalized patterns. Blanton et al. agree that variables are used to express algebraic patterns and add that variables used in this way symbolize arbitrary numbers.

The next variable types that Usiskin identifies are *unknowns* and *constants*, both used in the second conception of algebra, which focuses on solving equations. Unknowns are variables that the reader is trying to solve for, and any other variable involved is typically conceived of as a constant, which represents a known, fixed, but unspecified number. Blanton et al. describes unknowns as variables used to represent a fixed but unknown number. They emphasize that the value of these variables must be fixed and make the equation they appear in true.

In the third conception, where the focus is on functions and covariation, variables are referred to as either *arguments* or *parameters*. Arguments represent numeric values from the domain of a function. Usiskin describes parameters as numbers on which other numbers depend. Blanton et al. include arguments in their description of variable types but refer to them as *independent variables*. They also add an additional variable type, *dependent variables*, that refer to numeric values from the range of a function. They note that an important way of thinking about independent and dependent variables is considering how the dependent variable varies in relation to variation in the independent variable. Blanton et al. describe parameters as representing a quantity that, when varied, "determines the characteristics or behavior of other quantities" (p. 33), which is very similar to how Usiskin defined parameters.

The only type of variable Usiskin identified that is used in the last conception, where the focus is on the structure of algebraic objects, is variables used as *marks on paper*. Variables of this type are not thought of as numbers or having any value, but rather are thought of as arbitrary objects. Marks on paper are manipulated and used to explore or represent algebraic structures. This variable type connects to Blanton's idea of a variable that represents an arbitrary or abstract placeholder.

Blanton et al. uses the example of factoring to illustrate this variable type, the same type of example that Usiskin uses to illustrate marks on paper.

Table 1

Usiskin's Variable Types

Variable type name	Description	Example
Pattern generalizer	These variables are used to express an algebraic pattern or property, or a real-world trend	$a + b = b + a$ $T = -0.4Y + 1020$
Unknown	This variable is one the reader wants to solve for and find the value for	$5x + 3 = 40$
Constant	This variable is replaced by a specific value and does not need to be solved for	x =7.4
Argument	This variable stands for a domain value of a function	The variable x in the function: $f(x) = 3x + 5$
Parameter	This variable stands for a number on which other numbers depend	The variables m and b in the equation: $y = mx + b$
Arbitrary marks on paper	These variables are not thought of as having a value. They are used when performing operations and manipulating algebra equations	Factor $3x^2 + 4ax - 132a^2$

An Example of Complexity: $y = mx + b$

Usiskin notes that often many variable types are involved in solving algebraic problems, and that variables might change type during the problem solution. This could explain why students struggle to understand and work with variables. To illustrate this complexity, he analyzes the solution to the problem, "Find an equation for the line through (6,2) with slope 11" (p. 14). He walks through the typical solution for this problem by using the equation $y = mx + b$, which is treated as both a pattern and a formula, so the variables could be considered pattern generalizers or parameters and arguments. In order to solve this problem, we start by substituting 11 into the equation for m, indicating that the variable m is being used as a constant, not a parameter. This results in the equation $y = 11x + b$. The next step is to solve for b, which changes the variable type of b from a parameter into an unknown, because we are interested in the value of the variable.

In order to solve for b we use the given ordered pair, which gives us a value for the argument x and the y that is related to the x. Usiskin points out that the reason we may substitute in values for x and y is because $y = mx + b$ describes a general pattern. This means that x and y must be reconceptualized as pattern generalizers. Substituting in the given ordered pair we get the equation $2 = 116 + b$ which simplifies to $b = -64$. We now substitute in the known value for b and reach our solution, which is $y = 11x - 64$.

In the course of this one problem and solution, variables were thought of as pattern generalizers, arguments, parameters, constants, and unknowns. Variables changed type throughout the problem based on what operations were being done or what was needed to be done next. This demonstrates how complicated variable use can become.

Role of the Reader

While working with variables it is important to note the role of the reader. “The reader” could be the writer, the user, the learner, the teacher—in other words, the person who is making sense of the variables. The meaning and type of variable is not embedded in the variable itself, but rather is constructed by the individual interacting with it. Both Usiskin and Blanton et al. seem to be aware of this dependency on the reader. This is particularly evident in the above section where Usiskin explored how the variables in one problem take on several meanings. However, neither article fully articulates this dependency on the reader or considers the role of the reader in variable use and type. My analysis will explicitly attend to the role of the reader in interpreting variable type and meaning. Because of the active role that the reader plays in interpreting variables there could be multiple ways to read a symbol. In order to decide which is the best interpretation, readers must carefully consider the context in which the variable is found and draw upon their understanding of mathematical norms for using and interpreting variables. Note that the “correctness” of a reader’s interpretation of variables is not determined by whether the interpretation matches an inherent meaning in the symbols themselves, but rather by the degree to which other experts in the mathematical community would agree with the reader’s interpretation.

Research Question

From the literature that exists, it is clear that there are different types of variables that are used in different ways. The literature also suggests this variety of types and the combination of variable types in a single problem contribute to student confusion when learning algebra. However, these papers do not specifically address which variable types are used in middle school, which is where students will first encounter the consistent use of variables. More work is needed to understand how variables are used in middle school because middle school is where students will form their initial understanding of variables. This understanding will be what students will draw on as they progress in their mathematics careers. If students' understanding of variables is not developed properly, they will struggle to succeed in mathematics.

One way to identify the variable types middle school students might encounter is through an examination of middle school mathematics textbooks. Since these textbooks reflect the mathematical content that students will likely study, they also provide insight into the types of variables students are likely to encounter. A review of the literature suggests that no studies have been conducted of middle school mathematics textbooks to identify how variables are used in middle school mathematics. There is a clear need for research to investigate how variables are being used in a critical time of students' mathematical learning. This leads to the research question: What are the different ways that variables are used in middle school mathematics textbooks?

III. Methods

Data and Research Personnel

The curriculum used in this study is Eureka Math (2015), created by Great Minds, and accessed through EngageNY.org. This is a free, online mathematics curriculum for grades pre-kindergarten through grade 12 and is widely used across the United States. The authors note that the curriculum was designed to align with the Common Core State Standards (2010) and emphasize key concepts that build from year to year rather than procedural memorization. According to Great Minds, Eureka Math has been downloaded more than 13 million times. I selected this curriculum because of its popularity and ease of access.

I took samples from grade 6 module 4, grade 7 module 3, and grade 8 module 4. Grade 6 module 4 students extend their arithmetic understanding to working with variables that "stand-in" for numbers. Grade 7 module 3 expands on student understanding of equivalent expressions by using properties of operations. Grade 8 module 4 focuses on creating and solving linear equations. In each grade I selected the module that aligned with the Common Core topic of Expression and Equations, as this is where variables are used frequently. I sampled the first 6 lessons of each module. Grade 6 module 4 lesson 6 does not appear in the analysis as there were no variables used in the lesson.

The research team consisted of three researchers, myself included. I am a senior in mathematics education. Ashlyn Rounds is a research assistant and junior in mathematics education. Dan Siebert has a Ph.D. in mathematics education and 25 years of research experience.

Data Analysis

We began data analysis by analyzing a small selection of the sample that spanned all three grades. We coded the data by highlighting every variable in the sample, identifying what variable type each use was by comparing it to the variable types I identified in the literature review (see Table 1), and writing the corresponding abbreviation by the variable. We took several lessons from our sample at a time to code. After we had each coded the section individually, we met together to compare and discuss the codes we assigned. During this meeting, we identified any disagreements we had and discussed possible reasons for the disagreements. We made modifications to the codes so that the descriptions fit the data better, and then individually recoded the same section of data in order to test the modifications made. We met again and compared our codes on the same data until we were in agreement and felt satisfied with our code descriptions.

We repeated this process of individual coding and comparing codes, adding a new small section of the sample data each time. Changes made to the coding scheme often required us to return to previously coded data and recode so that the codes were used consistently across all of the data. We continued to add small sections until we had coded the whole sample and were in agreement on all the codes.

Each meeting led to modifications of the codes. The most common modification made was adjusting the definitions of the variable types. This included adding distinguishing details based on examples we found in the sample or making note of keywords or phrases that typically accompanied a specific type of variable. The more of the sample we coded, the more specific and precise the definitions became. Another modification that was made to the existing codes was the addition of new codes. These new codes were created to capture and match distinctions we observed in variable usage. As we coded the sample, we would also collect examples that illustrated the typical use of a variable type. These examples were added next to the definitions to be used as a reference.

After all the data in the sample had been coded, I counted all the variable uses and types and created a spreadsheet. I separated the uses by grade and lesson, so it was easier to see which lessons and grades used which variable types the most. I also created a tally for all the uses across all three grades so that I was able to see which types were most common. This made it so I could find patterns in the data and draw conclusions from the counts.

The last step in the data analysis process was to write descriptions of each of the codes with enough detail that readers would understand the defining features of each type of variable. I wrote these final descriptions by drawing from the definitions my team created and the examples I made note of earlier. The descriptions identified the most important features of each of the variable types, as well as information that would help the reader differentiate between variable types that were similar.

IV. Results

As I coded types of variables in the textbooks, I noticed that the uses of variables were tightly connected to the type of algebraic processes in which they were situated. The different types of variables can be organized around three main algebraic processes: signify and operate, constrain and solve, and define or name. There is also a more localized process, reinterpreting symbols, which can happen within or across the main three processes. In this section, I first define the algebraic process and then discuss the variables associated with that process. As the different types of variables are sometimes hard to distinguish from one another, each section includes examples found in the text and justifications for why each variable was coded the way it was.

Signify and Operate

The *signify and operate* process focuses on representing numbers and performing operations on them for any purpose other than solving for an unknown. Numbers are represented by a variable in order to make it easier to perform operations. Sometimes the meanings for the numbers associated with the variables are temporarily ignored to reduce cognitive load and make computations easier. Note that finding the value associated with the variables is not the focus of the problem solving process in the signify and operate process; rather, the focus is on using variables to represent, explore, and simplify relationships between quantities. The types of variables that are most commonly used in this process are placeholders, objects, and specified numbers.

Placeholders

Placeholders are variables that represent unspecified values and are often used to show relationships between these unspecified quantities. The reader keeps in mind that the variable is representing a number and could be replaced with a specific number. In grade 7 lesson 1, the authors write "Let t represent the number of triangles, and let q represent the number of quadrilaterals" (p. 15). Here the reader sees that t and q are representing numbers, namely the number of triangles and quadrilaterals in a class activity. The variable is being directly tied to a value, and while this value is unspecified, it could be replaced with a specific number. It is important to note that the value of the placeholder is unspecified, not arbitrary. In grade 8 lesson 1, students are asked to translate phrases into symbolic expressions for the sole purpose of representing relationships between quantities. The following example in the lesson demonstrates what students are expected to do:

> Problem: The sum of four consecutive even integers is –28.
>
> Answer: Let x be the first even integer. Then, $x + x + 2 + x + 4 + x + 6 = -28$. (p. 17)

Here the value of x is not arbitrary, as replacing x with 2 makes the statement incorrect. So, while the value of x is not the focus of the problem, the value of x is important to how the variable is being used. In other cases, the value of a placeholder could have various different values. Grade 7 lesson 2 discusses an equation that "is true for every value of x" (p. 29). Here the variable can take on any value, but the reader is aware and acknowledges that the variable is representing a number.

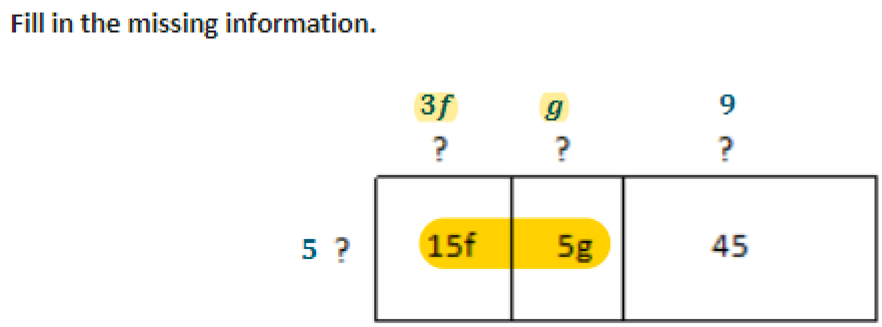
Objects

Objects are variables that can be thought of as “marks on the page” and are often used during symbolic manipulations. While these variables technically represent numbers, the reader temporarily thinks of them as objects with no meaning so that she can operate on them using symbolic manipulations. Objects often appear in the context of expressions where the expression is being acted on as a whole, without paying attention to the individual variable. For example, grade 7 lesson 2 has students “find the sum of -3 x and 8 x.” The solution given is “5 x” (p. 41). Here the variable x is being manipulated without thought or mention of x representing a number. When variables are being used as objects, the focus is on the various algebraic manipulations and operations being done to the variable, not on what the value of the variable is or could be.

Placeholders and objects can be difficult to tell apart. The main reason for this confusion is that it can be difficult to know if the reader is thinking about the variable as a number or as a “mark on the page.” This is further complicated by the fact that placeholders and objects are involved in the same algebraic process. The major difference between placeholders and objects is the implication of the variable having a value. In grade 7 lesson 4, the authors use variables in a rectangle to represent an area and ask students to find the side lengths as seen in Figure 1. This setup is used several times to the point that most readers would consider this a procedure and it would be easy to view the variables used as objects. However, the context of the rectangle implies that the variables represent an area, which means that they have an unspecified value attached to them; thus, these variables are placeholders.

Figure 1

Diagram Representing Area and Side Lengths



(7.3.4 p. 71)

Objects appear without reference to their value and there is no attempt to imply that the variable has a value. If the author instructs the reader to perform an operation on an expression or lists a string of operations done on expressions containing variables, then these variables are typically objects. Distinguishing placeholders from objects sometimes comes down to a judgment call made by the reader. The reader should pay careful attention to any attempt made by the authors to assign value to a variable.

Note that objects cannot be replaced directly by a number and need to be reconceptualized as a different type of variable before the reader can replace the variable with number. For example, in grade 7 lesson 4, the authors ask the reader to "rewrite the expressions as a product of two factors," then pose the expression "$72t + 8$" which is rewritten as "$8(9t+1)$" (p. 63). In this case the variable is being manipulated without thinking of it as a number, making it an object. A request to substitute a value in for t would require the reader to reconceptualize the variable as representing a number, thus changing the variable type.

Specified Numbers

Specified numbers are variables that are assigned a specific numerical value. Specified numbers are similar to placeholders as they represent a value, but they differ from placeholders as they represent a specified, rather than unspecified, numeric value. For example, in grade 6 lesson 5, readers are instructed to “evaluate both expressions when $x = 2$” (p. 60). The variable x is given a specific value for the sake of evaluating an expression. While specified numbers are often denoted by a variable with an equal sign, other key words can let the reader know that a variable is a specified number. In grade 7 lesson 2 the authors instruct, “Have students check the equivalency of the expressions by substituting 2 for x and 6 for y” (p. 30). Here the variables are being specified by the word “for.” Another sign that these variables are specified numbers is the use of the word “substituting,” as specified numbers can be replaced in expressions by a number. For example, in grade 7 lesson 2, after the instructions of “have students check the equivalency of the expressions by substituting 2 for x and 6 for y,” there are two equations, $(3x + 5y - 4) - (4x + 11)$ and $-x + 5y - 15$ (p. 30). Both x and y are classified as specified numbers because these variables were assigned a specific value in the instructions above and are immediately replaced with those specific values in the next line. These variables are not classified as placeholders because the value has been decided and expressed.

Table 2

Signify and Operate Variables

Variable name	Definition	Example
Placeholder	Represent unspecified values and show relationships between quantities	Let q represent the number of quadrilaterals in an envelope
Object	"Marks on the page" often used during algebraic operations	$3(x + 2) = 3x + 6$
Specified number	Variables that are assigned a specific value	Let $x = 16$

Constrain and Solve

The purpose of the *constrain and solve* process is focused on finding values for unknowns. This process typically involves constructing equations that constrain the values of unknowns and then performing symbolic manipulations to transform the initial equation into one where the variable representing the unknown is on one side of the equal sign and a single number is on the other. The types of variables involved in the constrain and solve process are unknown number, unknown object, and solution.

Unknown Numbers

Unknown numbers are variables that represent a value that the reader is interested in knowing but has not yet specified. These variables are similar to the placeholder variables used in the signify and operate process, but the main difference is that the reader wants to know the value of the variable rather than merely express relationships between quantities. This desire to know the value may be indicated by the framing of the problem. For example, in grade 6 lesson 4 the question is posed, "What is *x, y, z?*" (6.4.4 p. 43). This example demonstrates that the variables are being used in a way where the reader knows that they represent unknown values, and that finding the values is the focus of the problem. The value of the unknown number is not yet specified, but the value that replaces the unknown number must satisfy the given equation or context. This could be one specific value, like seen in the earlier example, but it could also be multiple values, like in grade 8 lesson 3 when the question asks, "What number(s) x satisfy the equation?" (8.4.3 p. 30). Here, there could be multiple values that could replace x, and the reader is interested in all of them.

Identifying unknown numbers often requires the reader to attend to the entire context surrounding variables. For example, in grade 8 lesson 6, the following example appears: "The equation in this example can be modified to $2x + 1 = -(5x + 9)$ to meet the needs of diverse learners" (8.4.6 p. 67). At first glance, x could be considered an object or unknown object as the focus does not seem to be on the value of x or the fact that x is representing a value. However, this example appears as a modification for the original question "What value of x would make the following linear equation true: $12(4x + 6) - 2 = -(5x + 9)$?" (8.4.6 p. 67). In this passage, it is clear that x is an unknown number, so the use of x in the modification is also part of the constrain and solve process, making it an unknown number.

Unknown Object

Unknown objects are variables that are treated as objects during the constrain and solve process. Unknown objects appear when the reader temporarily thinks of the variables as having no meaning other than being “marks on paper” during the process of solving an equation. When using unknown objects, the emphasis is on using algebraic manipulations to reduce the equation to the form of having the variable on one side of the equal sign and a single number on the other. In grade 8 lesson 4, unknown objects are used when the authors are outlining the operations done on linear equations as follows:

$$x - 9 = \frac{3}{5}x$$

$$x - x - 9 = \frac{3}{5}x - x$$

$$(1 - 1)x - 9 = \left(\frac{3}{5} - 1\right)x$$

$$-9 = -\frac{2}{5}x$$

$$-\tfrac{5}{2} \cdot (-9) = -\tfrac{5}{2} \cdot -\tfrac{2}{5}x \qquad \text{(8.4.4 p. 46)}$$

Every use of x in this example is an unknown object, as the reader does not treat x as a number, but rather as a symbol to manipulate. Unknown objects are also used to give instructions, as seen in the same lesson, where the authors instruct “solve the linear equation $x - 9 = \frac{3}{5}x$” (8.4.4 p. 46). The use of x here is classified as an unknown object because the reader is not required to think of x as having a numeric value, but can instead immediately begin performing algebraic manipulations.

Solution

Solutions are variables that have a specific value attached to them, and that value has been deduced through the constrain and solving process. Solutions satisfy an equation. In grade 8 lesson 4, the reader is asked to solve a linear equation. After going through a series of algebraic operations, they are left with $x = -10$. The authors point out that "since the left side is equal to the right side, then $x = -10$ is the solution to the equation" (8.4.4 p. 45). The variable x is a solution as it is attached to a specific value which was found by solving, and it satisfies the equation. Solutions are similar to specified numbers as the variable is representing a specific and known value. The value of solutions can also be proposed, then tested. An example of this can be found in grade 8 lesson 3 where the question is posed "Is the number 3 a solution to the equation? That is, is this equation a true statement when x =3?" (8.4.3 p. 31). This use of x is classified as a specified number because the reader does not yet know if 3 would satisfy the equation. However, once this specified number has been tested, and it is seen that "the left side of the equation equals the right side of the equation" it is confirmed that "we can say that x =3 is a solution to the equation" (8.4.3 p. 31).

Occasionally students were required to check a value that did not satisfy an equation. Students were shown that they should report a negative result using the variable for the unknown number, an unequal sign, and the number that was not a solution (i.e., $x \neq$-1). We referred to this type of variable as a *non-solution*. For example, in grade 8 lesson 3, the authors ask the students to check whether x =5 is a solution for the linear equation $8x - 19 = -4 - 7x$. After substituting 5 in for x, the reader is left with 21=39. The authors conclude "since $21 \neq -39$, then $x \neq 5$" (8.4.3 p. 31). The correct value of x is still unknown, but the reader knows that a proposed value is not a solution.

Table 3

Constrain and Solve Variables

Variable name	Definition	Example
Unknown number	Represent a value that the reader is interested in knowing but has not been specified yet	What number x will satisfy the equation?
Unknown object	"Marks on the paper" used when the reader is performing operations to solve an equation	Solve the linear equation $5x + 3 = 6x - 2$
Solution	Variables with specific values that are obtained through solving an equation	$x = 5$ satisfies the above equation

Define or Name

The *define or name process* involves using variables to define properties and operations, or to distinguish mathematical objects from each other quickly and concisely. The define or name process happens when the reader wants to illustrate a pattern or assign a name to an object in order to reference it later. This process does not involve solving for variables or performing operations. The most common variables used in this process are generalized numbers, parameters, and mathematical objects.

Generalized Numbers

Generalized numbers are variables that are placeholders in an identity or definition for numbers that come from a common number set, like the integers or natural numbers. An example of generalized numbers is given in the identity $w - x + x = w$ (6.4.1 p. 15). The property or identity being described with these variables must be true for every number in the set that is being operated on. The variable is thought of as representing a number (as opposed to being thought of as merely a "mark on paper"). Generalized numbers can be replaced by any number in the defined set, but there is no "correct" number that the variable can be replaced by.

Generalized numbers are sometimes difficult to distinguish from placeholders. While generalized numbers represent an unspecified value, like placeholders do, generalized numbers are used to define properties or identities, while placeholders are used to represent relationships between numbers in specific cases. For example, grade 6 lesson 5, both x and n are generalized numbers in the equation

$$x^n = (x \cdot x \cdots x)\}n \text{ times}$$

as they are being used to provide a definition of what a number raised to a power means (6.4.5 p. 55). However, in the subsequent equation

$$4^n = (4 \cdot 4 \cdots 4)\}n \text{ times}$$

n is being used as a placeholder because this equation is a more specific case (6.4.5 p. 56).

Parameters

Parameters are variables that represent constants in equations or expressions that also have unknown numbers. Parameters allow the reader to express a general description for functions or problem types. In the equations $x + p = q$ and $px = q$, both p and q are parameters (8.4.1 p. 13). The reader understands that both p and q could be replaced with any number, but the focus of these equations is not on the value of p or q; instead, the focus is on understanding the structure of the equations to identify the set of functions and problem types the equations are being used to model. For example, the equation $px = q$, where $p \neq 0$, can be used to describe a family of equations that can all be solved using the same strategy, namely dividing both sides of the equation by the coefficient of the unknown.

Mathematical Object

Mathematical objects are variables that act as names for objects that the reader wants to refer back to. These objects could be angles, points, axes, sets, and functions. In grade 8 lesson 5 a problem describes a triangle where "∠A is the largest angle; its measure is twice the measure of ∠B" (8.4.5 p. 63). Here both A and B are being used as names for the angles of the proposed triangle. By assigning names to the angles, the reader is able to easily track which angle has what measure and what their relationship is to the other angles. Mathematical objects help keep other algebraic processes organized and easy to read and follow.

Table 4

Define or Name Variables

Variable name	Definition	Example
Generalized number	Represent algebraic patterns or identities	$s + r = r + s$
Parameter	Represent constants in equations or expressions that also have unknown numbers	The variables p and q in the equation: $x + p = q$
Mathematical object	Act as a name for objects the reader will refer back to	∠A is the largest angle

Reinterpreting Symbols

Sometimes a single variable instance must be interpreted in more than one way, so the variable is classified as two types. These cases are when the reader must *reinterpret symbols*. A slash is used to indicate that a variable is being read as two types of variables. In this process, the reader will think of a variable as one type in order to make sense of the surrounding context or to perform operations, but then must reinterpret how they are thinking of the variable in order to continue in the next algebraic process.

Reinterpreting symbols can happen *within* one algebraic process, such as a variable being classified as an unknown object/solution. *Unknown object/solution* variables are common in the constrain and solve process. Once the reader has gone through the operations necessary to get an unknown object equal to a constant, they must reinterpret the unknown object as a solution. In grade 8 lesson 5 the reader is given the equation $x + 3x - 5 = 143$ and are asked to solve for x. The process of solving is then given:

$$x + 3x - 5 = 143$$

$$(1 + 3)x - 4 = 143$$

$$4x - 5 = 143$$

$$4x - 5 + 5 = 143 + 5$$

$$4x = 148$$

$$x = 37 \qquad \text{(8.4.5 p. 54)}$$

The first 5 lines all use x as an unknown object as algebraic manipulations are done without thinking of the variable as having a value, so the reader is thinking of x x as an unknown object when they reach the final line. However, once the reader reaches the final line, they also recognized that $x = 37$ is the solution to the equation, so they must reinterpret what type of variable x is.

This process can also happen *across* algebraic processes. One example of this transition is a placeholder/unknown number variable. In grade 6 lesson 4 a relation between division and repeated subtraction is explored in a table shown in Figure 2:

Figure 2

A Table used to Explore Division Properties

Division Equation	Divisor Indicates the Size of the Unit	Tape Diagram	What is x, y, z?
$12 \div x = 4$	$12 - x - x - x - x = 0$		$x = 3$

(6.4.4 p. 43)

The use of x in the first column is a placeholder as the reader knows that x has a value, but is not yet interested in what that value is. The x's in the second column represents the same value as in the first column, so they can still be thought of as placeholders. Note that while attending to the first two columns, the reader is focusing on understanding the relationship between division and repeated subtraction, and not on solving for x. However, in the third column the reader's purpose changes to finding the value of x. In order to transition to the solving process, the x's in the second column must be thought of as unknown numbers as well. Thus, the switch in how the reader interprets the variable accompanies the change in the reader's purpose.

Frequency of Variable Types

Because of the small size of the sample and the likelihood that variable use is dependent upon the mathematical topics being taught, I cannot make general claims about how common variable types were across the textbook series, nor can these results be generalized to middle school mathematics textbooks in general. Nevertheless, I present the tallies of the different types to show the frequency of variable types in the sample and emphasize the many different types of variables students must navigate even in a small sample of the textbook series (see Table 5).

Table 5

Variable Frequency by Grade

Algebraic process	Variable type	# of uses	Grade 6	Grade 7	Grade 8
Signify and operate	Object	1861	0	1490	349
Signify and operate	Placeholder	1057	243	432	408
Signify and operate	Specified Number	405	6	227	172
Constrain and solve	Unknown Object	763	0	0	763
Constrain and solve	Solution	150	38	0	112
Constrain and solve	Unknown Number	139	83	6	50
Define or Name	Generalized Number	437	325	18	94
Define or Name	Mathematical Object	26	0	0	26
Define or Name	Parameter	10	0	0	10
	Total	4,848	695	2,173	1,984

Object/Specified Number	153	0	131	22
Object/Placeholder	55	0	11	44
Unknown Object/Solution	53	0	0	53
Placeholder/Unknown Number	48	44	4	0
Placeholder/Specified Number	37	0	0	37
Unknown Object/Specified Number	22	0	0	22

Note. The slash codes were counted as one of each variable type they represented i.e., 55 object/placeholders were counted as 55 objects and 55 placeholders. Entries below the dotted line represent the slash codes.

In total, 4,848 variables were coded from 17 lessons. The most frequently used type of variable was object followed by placeholder. Both of these variables are part of the signify and operate process, which could mean that this process was the one that most of the sample was involved in. Because this is a fairly small sample, it is difficult to say if the signify and operate process is the most commonly found algebraic process across all texts. However, with 3,323 of the coded variables being involved in the signify and operate process, it is clear that this process is one of the most variable heavy processes and there is a significant need for students to understand how variables are being used in the process.

Objects and unknown objects were used 2,624 times, so 54.13% of the uses of variables were objects and unknown objects. One of the reasons objects were used so much is because the textbook laid out every step in the simplification or solving process. This is usually several lines of algebraic work, and each line contains several uses of an object or unknown object. When an object is used it typically appears multiple times in a small section. This frequency of use implies that this is a variable type that students are likely to encounter, so they need to be comfortable and confident in using variables as objects. With objects being used so frequently, there is a risk that students will begin to assume that all variables should be treated like objects, so the difference between objects and other variable types needs to be emphasized to students.

V. Discussion

There are several points of interest from the research that are explored more fully in this discussion. The first part explores the findings of the research, looking at the complexity of variable use and similarities across variable types. The next section compares my findings to those of Usiskin, as Usiskin's variable types were the original starting point of this study.

Research Findings

Complexity in Variable Types

The data produced by this research suggests a growing complexity in the use of variables throughout middle school. Despite the small sample size, it is clear that even in grade 6 variables are being used in complex ways that require students to frequently translate between types and processes. All three algebraic processes were used in all three of the grades and every grade had at least one use of a slash code. This illustrates how complicated variable use is in middle school mathematics. Being aware of different variable types and learning to recognize what variable type is being used is central to meaningful use of variables.

In grade 6 there are 5 different variable types used from 3 algebraic processes and 1 type of slash code used. Lessons contained from 1-5 variable types with an average of 2.2 variable types per lesson. This tells us that even in 6th grade the complexity level of variable use is high. The first three lessons only use 1 variable type, so these lessons may be considered simple in terms of variable use. However, grade 6 lesson 4 uses 5 variable types, including 44 uses of the placeholder/unknown number slash code. This slash code adds to the complexity as students not only have to reinterpret variable types in a problem but also must reinterpret which algebraic process they are working with. There are many different variable uses in this small sample which suggest that students are expected to transition between variable types and processes frequently. The variables used in grade 7 suggest a growing complexity as there are 8 types of variables used in the sample and 3 types of slash codes. Lessons have from 2-6 variable types with an average of 3.33 variable types per lesson. There is also a large jump in the number of variables used from 6th grade to 7th, with 695 variables used in 6th grade and 2,173 variables used in 7th grade. Thus, the data suggests that not only is the complexity increased from 6th to 7th grade, but also the frequency of variable use. Grade 8 makes use of all 9 variable types and 5 types of slash codes, making it the most diverse grade in the sample. Lessons have from 3-9 variable types with an average of 6.5 variable types per lesson. This increase of variable types used and the need to reinterpret symbols in multiple ways suggests an evolution of variable usage across middle school.

Variable Type Similarities Across Processes

While the variable types associated with the define and name process clearly differ in nature from other variable types, there are striking similarities between the variable types associated with the processes *signify and operate* and *constrain and solve Placeholders* are similar to *unknown numbers* because they both refer to unspecified numbers. *Objects* are similar to *unknown objects* because they both are treated as "marks on paper." And *specified numbers* are similar to *solutions* in that they both represent specified numbers. It appears that these variable types could be combined because of their similarities. However, these variable types are distinct and there is value in separating them.

Variable types are identified based on what algebraic process they are involved in. In the signify and operate process, the focus of the reader is on representing relationships and performing operations. The value of the variables used during this process is not the focus of the process. While value may be assigned to a variable, as with specified numbers, the emphasis is on the structure of expressions or equations and the operations done to them. In the constrain and solve process, the focus of the reader is on the variables and their numeric values. This shift of focus from representing to solving creates two distinct reasons for using variables. For example, objects and unknown objects are used in similar ways while performing operations but are embedded in different processes. Objects will remain objects or become specified numbers and the operations done to objects are for the purpose of simplification. The operations done to unknown objects are done with the intent to solve for the unknown object and find the solution. Although both objects and unknown objects have operations done to them the reader has very different goals and motives when using them, and consequently operates differently on them.

This distinction in variable type based on algebraic processes may not seem necessary to experienced mathematics users because their fluency with variable types and algebraic processes may cause them to miss subtle differences in variable types and uses. In contrast, students may be more likely to link variable types with algebraic processes as a way to make sense of variable usage. The distinctions made by this study are intended to aid student understanding of variable use.

Comparison to Usiskin

While my research was originally based on the variable types, definitions, and framework presented by Usiskin (1988), I eventually constructed a new framework and new variable types. My research and variable coding are distinct from Usiskin in my use of processes rather than conceptions, and in the specific variable types I identified.

Process Vs. Conception

Usiskin described four algebraic conceptions and categorizes variable types using these conceptions. These conceptions serve as a framework to understand variable use in school algebra. While these conceptions are useful to Usiskin and his purpose of his article, they ultimately are not the most practically useful way of distinguishing variable types. I chose to separate variable types into different algebraic processes rather than conceptions.

One reason I made this choice is because when teaching middle school mathematics, it makes more sense to teach students about different processes of algebra rather than conceptions. Processes are action based and easily identified in algebra lessons. Conceptions are useful when looking at mathematics education from a theoretical or historical standpoint but are less useful to teachers as they engage in the day-to-day teaching of algebra.

Another reason I use processes instead of conceptions is because processes are more directly linked to variable types. In order to determine what algebraic conception is being used, the reader must first think about what algebraic process is being used. For example, if a reader was trying to identify which algebraic conception is being used in the problem “solve $16x + 2 = 20$,” they would first need to identify the fact that the variable x is being solved for, which means it is in the constrain and solve process. From here the reader could conclude that this problem falls under Usiskin’s second conception of algebra. However, the extra step of identifying a conception seems to provide little additional understanding or insight beyond that provided by identifying the process. Simply put, it is a more direct path to go from variable type to algebraic process.

Variable Types

Usiskin identified six variable types while my research identified nine variable types. There is some overlap in variable types and definitions as well as some key differences, which are explored below and summarized in Table 6.

There were several similarities between Usiskin’s variable types and those used in this study. Parameters were used the same way in Usiskin’s study as in this study and retained the same name. Usiskin’s constants and my specified numbers are also used in similar ways across studies. Pattern generalizers are similar to generalized numbers, as both are used to express algebraic patterns and properties. Usiskin’s “marks on paper” are called objects in my study, but both terms emphasize that these variables are not thought of as numbers, but rather are thought of as objects with their own set of operations and procedures. I separated Usiskin’s variable type unknowns into two types, unknown numbers and unknown objects because sometimes unknowns are thought of as numbers and other times as objects.

Table 6

Comparison of Usiskin's Variable Types to Mine

Usiskin's variable type	Similar variable type	Key differences
Pattern generalizer	Generalized number and placeholder	Separated out variables that generalize properties and variables that represent specific cases
Unknown	Unknown number and unknown object	Distinguished between variables used to represent a value and variables that are treated as objects
Constant	Specified number	No observed differences
Argument	-	Arguments not observed in sample
Parameter	Parameter	No observed differences
Marks on paper	Objects	Objects are not strictly used in studying structure
-	Solution	No comparable variable type
-	Mathematical object	No comparable variable type

Many of the differences in variable types between my framework and Usiskin's (1988) reflect my attempt to be clearer about the algebraic processes in which the variables are typically used. For example, Usiskin's pattern generalizer is used to represent general algebraic patterns, but also patterns that are drawn from real world contexts. I separated these two uses by creating the variable types generalized numbers, which represent universal patterns, and placeholders, which represent more specific cases. A similar separation happened with Usiskin's unknowns. My variable types distinguish between unknowns that are thought of as a number and unknowns that are thought of as objects.

There are some variable types that are unique to each study. Usiskin's argument variable type was not observed in the sample, and thus, there is no comparable variable type in my framework. My framework includes the variable types solutions and mathematical objects, both of which were not discussed by Usiskin.

VI. Contributions

This study makes important contributions to the existing literature. One important finding from this study is the confirmation that variable type is tied to algebraic processes. This connection was implied by Usiskin, (1988) who claimed that variables were tied to conceptions of algebra, which, based on the examples that Usiskin gave, involved different algebraic processes. In my research I found that variable types were directly related to the algebraic processes that they were used for.

This research also created a list of variable types that are likely to be found in middle school mathematics classes. This list, while originally based on Usiskin's variable types, includes several new variable types. Two variable types, mathematical objects and solutions, are unrelated to Usiskin's variable types. These variable types describe nuances in variable usage that are not discussed in existing literature. My research also introduces several variable types that are modifications of Usiskin's types. My variable types capture a clearer picture of how variables are used in middle school text and make connections between variable type and algebraic process.

Another contribution made by this research is the recognition of the complexity of variable use, even early on in mathematics learning. Usiskin's variable types suggest that variable usage is complex in algebra in general, and do not investigate the complexity of variable use in any specific grade. My research suggests that this complexity starts as early as 6th grade and continues to increase from there.

VII. Implications

This research is meant to ultimately help mathematics educators be better prepared and able to support their students in learning how to interpret and use variables in meaningful and correct ways. The findings in this study are not meant to be taught directly to middle school students, as middle school mathematics teachers already face daunting curriculum to cover in the course of the school year. Adding more concepts to teach to middle school mathematics students is not practical. These findings serve as a resource for educators as they prepare and teach lessons.

Few students are taught about the nuances in variable use or the fact that there are many different types of variables with different uses. Even in higher level mathematics these differences are not taught explicitly; rather, learners infer them from the mathematical activity they see modeled by their teachers and textbook authors. Consequently, learners' knowledge of these distinctions become implicit understanding. This study takes implicit knowledge and makes it explicit. By becoming aware of variable types, mathematics teachers can deepen their mathematical understanding. This deeper understanding can help teachers anticipate student struggle better when it comes to variables. For example, experienced mathematics educators understand that when solving linear equations, the solving process is complete once they are left with the variable equal to a constant, as that means they have found the solution. However, students sometimes struggle to recognize when they are done solving an equation. This study offers one possible explanation for their struggle—when solving linear equations, the variable being used is thought of as an unknown object, but once the student reaches the point where they have the variable equal to a constant, they must reinterpret the variable as a solution. By understanding the different types of variables used, and especially the times when variable type changes within a problem, educators will be better prepared to anticipate and address students' struggles with variables.

Another way this research can assist mathematics teachers is by helping them anticipate what algebraic processes will be used in specific grades. For example, the data suggests that the most frequently used variable type in grade 6 is generalized number. From this information teachers can anticipate that the major algebraic process to focus on in grade 6 is define and name and can plan lessons accordingly.

VIII. Limitations and Future Research

There are clear limitations to this study, as well as clear paths to extend the research on variable types. The most obvious limitation of this study is the sample size. This exploratory study looked at one curriculum series and only part of one module for each of the grades. Consequently, only tentative conclusions can be made regarding variable types in middle school mathematics. This limitation gives clear direction for future research. The next step in continuing this research is to repeat the variable coding process with a larger sample from *Eureka Math* or with another textbook series. Researchers may wish to conduct a much larger study that would include multiple textbook series across grades 6 through 8. This large-scale study would help reveal what is common across curricula and patterns in middle school variable use.

Another limitation of this study is that it only looks at how variables are used in textbooks and does not examine how they are actually used in middle school mathematics classrooms. Mathematics teachers and students may be using variables in ways that are different from how the textbook uses variables. This provides another opportunity for continued research. A future study could research how variables are being used in middle school classrooms and compare that usage with the variable types used in the adopted textbook or curricular materials.

IX. Conclusion

Variable usage is one of the most powerful tools in mathematics and is critical for students to understand. It is important for students to be able to make sense of and use variables meaningfully in middle school mathematics because this is the time when they begin a serious study of variables. If students can gain a solid, foundational understanding of variable usage they will be more prepared for the increasingly technical and STEM driven world they find themselves in. Understanding variable types and the processes they are related to is a good first step towards understanding the complex and compelling world of variable use.

References

Blanton, M., Levi, L., Crites, T., Dougherty, B. J., & Zbiek, R. M. (2011). *Developing essential understanding of algebraic thinking for teaching mathematics in grades 3-5*. National Council of Teachers of Mathematics.

Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic and algebra in elementary school.* Heinemann.

Eureka Math. (2015). *Great minds*. https://greatminds.org/math

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Mathematics standards*. Mathematics Standards | Common Core State Standards Initiative. Retrieved March 16, 2022, from http://www.corestandards.org/Math/

Usiskin, Z. (1988). Conceptions of school algebra and uses of variables. In A. Coxford (Ed.), *Ideas of algebra: K-12* (pp. 8-19). National Council of Teacher of Mathematics.