

## University of Portland Pilot Scholars

---

Education Faculty Publications and Presentations

School of Education

---

2016

# Elementary School Students and their Knowledge about 'Variable'

Nicole C. Ralston

*University of Portland*, [ralston@up.edu](mailto:ralston@up.edu)

Follow this and additional works at: [http://pilotscholars.up.edu/edu\\_facpubs](http://pilotscholars.up.edu/edu_facpubs)



Part of the [Education Commons](#)

---

### Citation: Pilot Scholars Version (Modified MLA Style)

Ralston, Nicole C., "Elementary School Students and their Knowledge about 'Variable'" (2016). *Education Faculty Publications and Presentations*. 38.

[http://pilotscholars.up.edu/edu\\_facpubs/38](http://pilotscholars.up.edu/edu_facpubs/38)

This Conference Paper is brought to you for free and open access by the School of Education at Pilot Scholars. It has been accepted for inclusion in Education Faculty Publications and Presentations by an authorized administrator of Pilot Scholars. For more information, please contact [library@up.edu](mailto:library@up.edu).

**Title:** Elementary School Students and their Knowledge about ‘Variable’

**Topic Area:** Mathematics Education

**Presentation Format:** Paper Session

**Description:** This mixed methods research design investigated the knowledge held by elementary school students in Grades 1-5 about variables. To explore this issue, 1,745 students completed paper pencil assessments and 73 students participated in think aloud interviews. This presentation will describe the results, including the misconceptions students exhibited and suggestions for remediation.

**Paper Author:** Nicole C. Ralston, Ph.D.  
University of Portland  
School of Education  
5000 N Willamette Blvd  
Portland, OR 97203  
[ralston@up.edu](mailto:ralston@up.edu)

### **Elementary School Students and their Knowledge about ‘Variable’**

The term *variable* is quite elusive and often difficult to define, especially for elementary school students. A common exercise used to understand the complexity of the term involves trying to define variable with only one word (Schoenfeld & Arcavi, 1988). Because “the meaning of variable is variable,” the wide variety of different acceptable meanings can make the term difficult for students to understand (Schoenfeld & Arcavi, 1988, p. 425). Researchers have long agreed that understanding variables is an extremely important topic in middle and high schools, especially when making the leap to algebra, but the knowledge requirements are confusing and appear to change over time (Usiskin, 1988).

Kuchemann (1978; 1981) developed a framework of what he considered the six different student interpretations of variables:

1. Letter evaluated (i.e., the letter is a specific number, for example  $a+3=5$ ),
2. Letter ignored (i.e., the letter itself is not given meaning, for example  $a+b=43$  so  $a+b+2=?$ ),
3. Letter as object (i.e., the letter stands for an object, for example  $s$  stands for students),
4. Letter as a specific unknown (i.e., the letter is a specific yet unknown number, for example add 4 onto  $n+5$ ),
5. Letter as generalized number (i.e., the letter can represent several numerical values), and
6. Letter as variable (i.e., the letter can represent a large range of unknown numerical values).

Although historically reserved for students in official algebra courses in middle and high schools, the importance of teaching these meanings of variable earlier is becoming more widespread. The new Common Core State Standards (CCSS; 2011), for example, recommends that “students in Grade 3 begin the step to formal algebraic language by using a letter for the unknown quantity in expressions or equations” (p. 27). When meeting this standard, students’ first introduction to variables would often therefore involve adding in literal symbols (i.e.,  $x$  or  $y$ ) into simple open number sentences or equivalence problems instead of blanks or boxes (i.e.,  $x+8=23$  instead of  $\_\_+8=23$ ) (Fujii & Stephens, 2008). This Grade 3 CCSS definition of variable may be one of the more rudimentary definitions, because it does not represent varying quantities, (Fujii & Stephens, 2008) however it may be the appropriate starting point for elementary school students. Despite variable use being stressed in the CCSS starting in Grade 3, there is very little research surrounding the conceptions held by elementary school students in this area. The

purpose of this study is therefore to investigate this further – what knowledge do elementary school students possess about variable and what misconceptions are occurring?

### **Common Variable Misconceptions**

A review of the research reveals that students often experience misconceptions surrounding variables that appear to persist with age. MacGregor and Stacey (1997) found that common misinterpretations of variables for students ages 11-15 included believing the letter to stand for 1, for an abbreviated word, for an alphabetical value, for a particular numerical value, or for a label for an object. Further, when presented with the task ‘ $2n+3$ , what does the symbol stand for?’, the percentage of 6<sup>th</sup>, 7<sup>th</sup>, and 8<sup>th</sup> grade students who understood was 46%, 63%, and 76%, respectively (Asquith et al., 2007; Knuth et al., 2005). The majority of students who answered incorrectly either did not know the answer, or believed it to stand for an object, word, or a specific digit. These results indicate that middle school students may still be interpreting letters as letters (or specific corresponding numbers) instead of understanding their purpose in representing numbers or a range of numbers.

One of the most popular misconceptions held by middle and high school students appears to be misunderstanding literal symbols as labels (i.e.,  $c$  stands for cat, so  $4c$  might mean 4 cats) (Booth, 1984; MacGregor & Stacey, 1997). McNeil, Weinberg, and colleagues (2010) investigated how these misconceptions manifest when utilizing variables using mnemonic letters (i.e.,  $c$  for price of a cake), non-mnemonic letters (i.e.,  $x$  or  $y$ ), or Greek letters (i.e.,  $\Phi$  or  $\Psi$ ). They presented 322 middle school students with a commonly used problem adapted from Kuchemann (1978, 1981) using these three different types of variables as conditions: ‘Cakes cost  $c$  dollars each and brownies cost  $b$  dollars each. Suppose I buy 4 cakes and 3 brownies. What does  $4c + 3b$  stand for?’ Students in the mnemonic letters condition misinterpreted the

expressions the most often, and often considered the labels as standing for objects.

Approximately 37% of students in the mnemonic (i.e., c and b) condition interpreted the variables correctly while approximately 56% of students in the non-mnemonic conditions interpreted the variables correctly. There was no difference between the non-mnemonic letters and the Greek letters groups (McNeil, Weinberg et al., 2010). It seems clear that students experience misconceptions around variables that could be remedied through additional experience with variables and through teacher preciseness in choosing which variables to use (i.e., choosing non-mnemonic letters as variables).

### **Importance of Variable Knowledge**

Mastering an understanding of variables and work with variables is significant because, if not remedied early, misconceptions appear to persist into high school and even into adulthood. A seminal study utilizing the notoriously famous ‘Student-Professor Problem’ (Kaput & Clement, 1979) demonstrated that even mathematically-proficient adults (i.e., college students pursuing traditionally mathematically based majors) often seriously struggle with representing mathematical relationships with variables. The most common misconception exhibited by the adults involved committing ‘reversal errors,’ which occur when the variables are reversed in formulas and has been found to be highly prevalent in both high school and college students (MacGregor & Stacey, 1993; Fisher, Borchert, & Bassok, 2010). This reversal error was demonstrated when between 40% and 60% of adults solved the following problem incorrectly: “Write an equation using the variables  $S$  and  $P$  to represent the following statement: ‘There are six times as many students as professors at this university.’ Use  $S$  for the number of students and  $P$  for the number of professors” (p. 288). The most common error made involved reversing the solution: ‘ $6S=P$ ’. A follow-up study revealed that a large proportion (i.e., 40-43%) of college

students could not identify the  $P$  to mean number of professors or the  $S$  to mean number of students, revealing the common misconception that occurs when students misunderstand the  $S$  to mean students (instead of number of students) and therefore reading  $S=6P$  as “one student for every six professors” (Rosnick, 1981, p. 419). Because of these misconceptions that even adults appear to possess, it is crucial that students gain experience working with variables at an early age.

### **Purpose of this Study**

Because of the significance and the accompanied research gap, this research sought to investigate the knowledge elementary school students possess surrounding variables.

### **Methods**

This investigation of student variable knowledge was pursued using a two-pronged, mixed-methods research design. The first, quantitative prong, investigated student knowledge through paper pencil assessments while the second, qualitative prong, investigated student knowledge through think-aloud student interviews.

### **Student Assessments**

Student assessments were conducted to investigate student knowledge surrounding variable across a large number of students in each elementary grade, Grades 1-5. No research found in the area had used random samples of students; therefore a random sample was utilized to reduce bias and increase generalizability of the results.

**Student demographics.** A stratified cluster random sample was used in this study, with schools stratified by mathematics achievement (i.e., high achievement, medium, and low) and students clustered by school. Of the 397 elementary schools in two urban counties of Washington State, six elementary schools from six different school districts were randomly

selected to participate: two high achieving schools, two medium achieving schools, and two low achieving schools. Schools were stratified by 3<sup>rd</sup> grade mathematics achievement: percent passing the 3<sup>rd</sup> grade state standardized test by school ranged from 37% to 83% with a mean of 65% ( $SD = 17\%$ ). Percent free or reduced price meals also varied widely, ranging from 30% to 77% with a mean of 45% ( $SD = 18\%$ ).

A total of 1,745 students participated by completing the assessment. The students were relatively equally distributed across the grade levels: 351 (20%) first grade students, 309 (18%) second grade students, 336 (19%) third grade students, 384 (22%) fourth grade students, and 365 (21%) fifth grade students. Approximately 51% of students were male, the average age of students was 8.96 ( $SD = 1.49$ ), and approximately 22% of students were English Language Learners (ELL).

**Assessment.** All participating students were administered a newly developed and validated diagnostic assessment of algebraic thinking skills. Reliability and validation efforts of the instrument are currently underway; however internal reliability coefficients (i.e., Cronbach's Alpha) for all versions of the assessment are at or above 0.80 and evidence of validity has been collected. Each assessment included at least three items investigating variable knowledge, primarily in the rudimentary first level of 'variable stands for a specific number.' Eight different variable items were used across the different grade levels. The items varied on the letter used as the variable (i.e., g, b, a, etc.), whether one or two of the same variables appeared, whether one or two different variables appeared, whether there was an 'equivalence context' (i.e.,  $7+4+5=7+e$ ) or not (i.e.,  $6+b=9$ ), and the type of operation (i.e., addition, subtraction, multiplication, or division).

**Analysis.** Classical Test Theory statistics (i.e., descriptive statistics, frequency counts, difficulty or p-values, etc.) were used to analyze the student responses. Response process analysis using coding and frequency counts was further used to provide common answers to each of the items. Further, analysis of variance (ANOVA) was used to investigate the differences in knowledge levels across grade levels.

### **Think-Aloud Interviews**

A think-aloud protocol study was conducted to further investigate student knowledge surrounding variable and better understand student thinking and any misconceptions students are experiencing.

**Student demographics.** Seventy-three students of varying abilities from two different elementary schools in King County of Washington State were selected to participate in the think-aloud interview protocol. Grade level participants included 18 1<sup>st</sup> grade students (25%), 12 2<sup>nd</sup> grade students (16%), 11 3<sup>rd</sup> grade students (15%), 10 4<sup>th</sup> grade students (14%), and 22 5<sup>th</sup> grade students (30%).

**Interviews.** In the think-aloud protocol interview, students were asked to verbalize their cognitive processes when answering each of the assessment questions (described above). Using a think-aloud protocol can provide an in-depth view of how students perceive the assessment items (Ericsson & Simon, 1984; Ginsburg, 1997). In this think-aloud protocol, interview students were asked to verbalize their cognitive processes when answering each of the questions. By asking the students what they were thinking when answering the assessment items or why they answered in the way they did, “the experimenter seeks to learn directly from them [the students] the underlying cognitive structure that produced the overt behavior” (Ericsson & Simon, 1984, p. 42).



**Analysis.** The think-aloud data was analyzed using a four-step process. First, all think-aloud interviews were audiotaped and transcribed. Second, several readings of the interview transcriptions were conducted to gather first pass “codes,” or common thought processes and conceptions. Third, qualitative codes or themes were finalized and all interview transcriptions were coded to discover a pattern of themes or codes (Miles & Huberman, 1994). Fourth and finally, these themes were analyzed in conjunction with the results of the response processes analysis discovered in the large assessment sample described above to look for themes across a larger range of data.

### Results and Discussion

The difficulty (p-values) for all items are displayed in Table 1.

Table 1

*Difficulty Values (p-Values) for Variable Items*

	1 <sup>st</sup> Grade	2 <sup>nd</sup> Grade	3 <sup>rd</sup> Grade	4 <sup>th</sup> Grade	5 <sup>th</sup> Grade
$6+b=9$	0.38	0.51	0.81	-	-
$10-g=2$	0.40	0.54	0.88	-	-
$4+a+a=10$	0.20	0.34	0.73	0.86	0.91
$c+c+3=15$	0.14	0.32	0.68	0.69	0.80
$7+4+5=7+e$	0.12	0.25	0.54	0.62	0.82
$n+n+n=n+12$	-	-	0.20	0.26	0.28
$4n+5=21$	-	-	0.50	0.70	0.79
$x+y+y=10;$	-	-	-	0.55	0.62
$x+y=6$	-	-	-	0.55	0.62

Although students often initially struggled to solve one variable items like  $6+b=9$  in 1<sup>st</sup> and 2<sup>nd</sup> grades (i.e., with p values of 0.38 and 0.51), they were fairly proficient by 3<sup>rd</sup> grade (i.e., with p values of 0.81). By 5<sup>th</sup> grade, students were also fairly proficient (i.e., with p values of 0.80 and above) with both two variable items (i.e.,  $4+a+a=10$ ) and variables used in an

equivalence context (i.e.,  $7+4+5=7+e$ ). Older students were also asked to solve more complex items involving two different variables (i.e., if  $x+y+y=10$  and  $x+y=6$  what is  $x$  and  $y$ ?). Students surprisingly mastered these problems fairly well (i.e., with  $p$  values of 0.55 to 0.62), but continued to struggle throughout the grade levels (i.e., with  $p$  values of 0.20 to 0.28) when the variables occurred on both sides of the equal sign (i.e.,  $n+n+n=n+12$ ).

There was a statistically significant difference by grade for the three items measured across all five grade levels: for  $4+a+a=10$  ( $F(4, 872) = 116.821, p < .001$ ), for  $c+c+3=15$  ( $F(4, 847) = 72.880, p < .001$ ), and for  $7+4+5=7+e$  ( $F(4, 856) = 73.106, p < .001$ ). Post-hoc Tukey analyses for  $4+a+a=10$  and  $7+4+5=7+e$  revealed that 1<sup>st</sup> and 2<sup>nd</sup> grade students significantly underperformed all other grades, while 3<sup>rd</sup> grade students significantly underperformed 5<sup>th</sup> grade students. Similarly, analyses for  $c+c+3=15$  revealed that 1<sup>st</sup> and 2<sup>nd</sup> grade students significantly underperformed all other grades; however there were no other differences.

### Conception Analysis

It appears that students faced a variety of different misconceptions when solving problems involving variables, with younger students more likely to be confused. The conception analysis and interview results of three specific items are highlighted below.

**6+b=9.** For the item  $6+b=9$ , some students understood immediately: for example a 2<sup>nd</sup> grader said, “Oh, that’s kind of like algebra where there’s like an  $x$  and you try and figure out what it is!” Others students were not familiar with the use of the variable and left it blank or just listed numbers in the problem such as 6 or 9: “9? Because it’s [written] right there.” However, certain students who were unfamiliar with problems such as these were able to work them out on their own. One 1<sup>st</sup> grade student, for example, said, “I’ll skip it. Wait, I think I know how to do

this one. 6 plus what number equals 9... b must be 3!” The most common conceptions for the item  $6+b=9$  are displayed below in Table 2.

Table 2

*Common Answers to  $6+b=9$*

	Blank	Number in problem	'b' is the 2 <sup>nd</sup> letter	Correct	Number in problem	Add or subtract numbers
$6+b=9$	-	6	2	3	9	15
1 <sup>st</sup> Grade	<b>15%</b>	3%	4%	38%	<b>14%</b>	4%
2 <sup>nd</sup> Grade	<b>16%</b>	3%	8%	51%	<b>14%</b>	1%
3 <sup>rd</sup> Grade	3%	2%	3%	81%	5%	2%
All	<b>11%</b>	3%	5%	56%	<b>11%</b>	2%

**10-g=2.** Similar conceptions occurred in the item  $10-g=2$  (see Table 3 below). Several students had not experienced these types of items. One 1<sup>st</sup> grade student said, “I can’t figure these out. The letter ones I can’t do.” Other students thought that the ‘g’ in  $10-g=2$  was actually a 9: “9? Because they kind of look like 9s.” Another student thought ‘g’ was a 6 because “if you turn the g upside down it’s going to be a 6.” Others simply replaced the answer with numbers from the problem like 2 or 10. Others still thought that the letter stood for a specific word, like ‘b’ for box or ‘g’ for Grady (i.e., the student’s name). Like before, some students who were not sure about the problem were still able to talk themselves through it successfully. One 3<sup>rd</sup> grade student, for example, said: “What? Okay, that’s kind of weird. I’m going to skip it. Wait, I think I know. You have to find out what the g is. 10 minus what would equal 2... so it is 8.” It is clear that students held a wide variety of different conceptions about what the variable could stand for, especially in cases where they had not previously been exposed to such problems.

Table 3

*Common Answers to  $10-g=2$*

	Blank	Number in problem	'g' is the 7 <sup>th</sup> letter	Correct	Number looks like a letter	Number in problem	Add or subtract numbers
10-g=2	-	2	7	8	9	10	12
1 <sup>st</sup> Grade	<b>17%</b>	<b>16%</b>	4%	40%	3%	5%	3%
2 <sup>nd</sup> Grade	8%	9%	4%	54%	1%	1%	2%
3 <sup>rd</sup> Grade	2%	4%	2%	88%	0%	1%	1%
All	9%	10%	3%	54%	1%	2%	2%

**c+c+3=15.** Students appeared to continue to experience misconceptions as the variable items became more difficult. When students were asked to solve items with two identical variables such as  $c+c+3=15$ , many of the younger students skipped this item: one 1<sup>st</sup> grader, for example, commented: “I don’t know. Can I just skip this one?” A 2<sup>nd</sup> grader said, “But I don’t know what these [points to ‘c’s] stand for.” A 3<sup>rd</sup> grader said, “This one’s a little too hard for me,” and a 4<sup>th</sup> grader said, “I’m going to skip that.” Other students were a little confused about having two variables in one problem. Many students put 12, which is what  $c+c$  equals. One 3<sup>rd</sup> grade student said, “Wait, for both ‘c’s or for only one c?” while a 5<sup>th</sup> grade student said, “15 minus 3 is 12.”

Table 4

*Common Answers to  $c+c+3=15$*

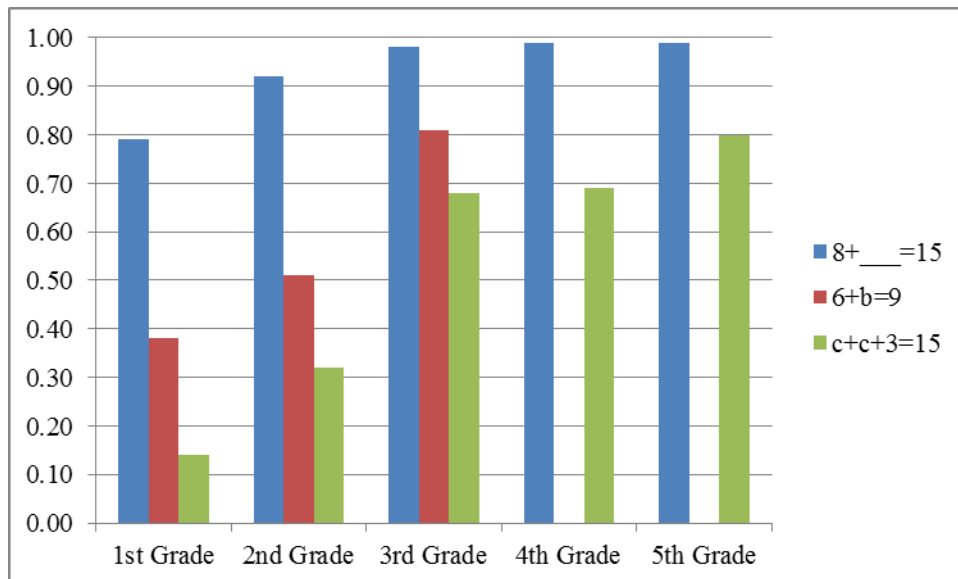
	Blank	Number in problem	Correct	Number in problem	Add or subtract numbers
$c+c+3=15$	-	3	6	15	12
1 <sup>st</sup> Grade	<b>11%</b>	<b>12%</b>	14%	9%	7%
2 <sup>nd</sup> Grade	<b>21%</b>	<b>14%</b>	32%	8%	8%
3 <sup>rd</sup> Grade	4%	7%	68%	2%	7%
4 <sup>th</sup> Grade	3%	8%	69%	2%	8%
5 <sup>th</sup> Grade	2%	4%	80%	1%	7%
All	<b>10%</b>	9%	52%	4%	7%

**Conclusion**

These findings supported the findings of previous researchers (MacGregor & Stacey, 1997), in that it does appear that students might benefit from more opportunities to experience variables. Students exhibited misconceptions when solving items with variables and appeared to struggle more when a variable was used in an open number sentence instead of a box or blank. Figure 1 below compares their results with the open number sentence  $8 + \_\_ = 15$ , a similar arithmetic problem that differs only in the use of the blank versus the use of a variable.

Figure 1

*Proportion of Students Solving Open Number Sentence and Variable Items Correctly*



When solving these variable items, many students simply reused numbers from the problem in the answer or even utilized the numeric / alphabetic code students sometimes learn (i.e.,  $a=1$ ,  $b=2$ ,  $c=3$ , etc.). When problems became more complex and two identical variables were seen (i.e.,  $c + c + 3 = 15$ ), students did not seem to understand that when the variable is the

same and used more than once the same number will replace both letters. Although this analysis measured one of the first, most rudimentary meanings of the word variable, it is important this meaning becomes familiar and is mastered before progressing to more complex meanings. It is imperative this knowledge is fostered early, as misconceptions surrounding variable can persist throughout middle and high schools and even college and beyond. This issue is best summed up by MacGregor and Stacey (1997): “Students frequently base their interpretations of letters and algebraic expressions on intuition and guessing, on analogies with other symbol systems they know, or on a false foundation created by misleading teaching materials... Their misinterpretations lead to difficulties in making sense of algebra and may persist for several years if not recognized and corrected” (p. 15).

## References

- Asquith, P., Stephens, A. C., Knuth, E. J., & Alibali, M. W. (2007). Middle school mathematics teachers' knowledge of students' understanding of core algebraic concepts: Equal sign and variable. *Mathematical Thinking and Learning*, 9(3), 249-272.
- Booth, L. R. (1984). *Algebra: Children's strategies and errors. A report of the strategies and errors in secondary mathematics project*. Windsor, Berks: NFER-Nelson.
- The Common Core Standards Writing Team. (2011). *Progressions for the Common Core State Standards in Mathematics*.
- Ericsson, K. A. & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Fisher, K. J., Borchert, K., & Bassok, M. (2011). Following the standard form: Effects of equation format on algebraic modeling. *Memory & Cognition*, 39(3), 502-515.
- Fujii, T., & Stephens, M. (2008). Using number sentences to introduce the idea of variable. In C. Greenes & R. Rubenstein (Eds.) *Algebra and algebraic thinking in school mathematics: Seventieth Yearbook*, pp. 127-140. Reston, VA: National Council of Teachers of Mathematics.
- Ginsburg, H. P. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. Cambridge University Press: Cambridge, UK.
- Kaput, J., & Clement, J. (1979). Letter to the editor. *Journal of Mathematical Behavior*, 2, 208.
- Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. S. (2005). Middle school students' understanding of core algebraic concepts: Equality & variable. *International Reviews on Mathematical Education*, 37, 68-76.

- Kuchemann, D. (1978). Children's understanding of numerical variables. *Mathematics in School*, 7(4), 23-26.
- Kuchemann, D. (1981). Algebra, in Hart, K. (Ed.) *Children's understanding of mathematics 11-16*. London: John Murray.
- MacGregor, M. & Stacey, K. (1993). Cognitive models underlying students' formulation of simple linear equations. *Journal for research in mathematics education*, 24(3), 217-232.
- Macgregor, M. & Stacey, K. (1997). Students' understanding of algebraic notation: 11-15. *Educational Studies in Mathematics*, 33, 1-19.
- McNeil, N. M., Weinberg, A., Hattikudur, S., Stephens, A. C., Asquith, P., Knuth, E. J., & Alibali, M. W. (2010). A is for apple: Mnemonic symbols hinder the interpretation of algebraic expressions. *Journal of Educational Psychology*, 102(3), 625-634.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis* (2nd edition). Thousand Oaks, CA: Sage Publications.
- Rosnick, P. (1981). Some misconceptions concerning the concept of variable. *The Mathematics Teacher*, 74(6), 418-420.
- Schoenfeld, A., & Arcavi, A. (1988). On the meaning of variable. *The Mathematics Teacher*, 81, 420-427.
- Usiskin, Z. (1988). Conceptions of school algebra and uses of variables. In A. F. Coxford (Ed.s), *The Ideas of Algebra K-12* (pp. 8-19). National Council of Teachers of Mathematics: Reston, VA.