Contemporary Educational Psychology 71 (2022) 102114

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

Variables versus numbers: Effects of symbols and algebraic knowledge on students' problem-solving strategies

Jenny Yun-Chen Chan[a,b,*], Erin R. Ottmar[a], Hannah Smith[a], Avery H. Closser[a,1]

[a] Social Science and Policy Studies, Worcester Polytechnic Institute, Worcester, MA 01609, USA

[b] Department of Early Childhood Education, The Education University of Hong Kong, Tai Po, Hong Kong Special Administrative Region

ARTICLE INFO

Keywords:
Variables
Algebra
Problem solving
Mathematical structure
Online learning

ABSTRACT

To efficiently solve mathematical expressions and equations, students need to notice the systemic structure of mathematical expressions (e.g., inverse relation between 3 and − 3 in 3 + 5 – 3). We examined how symbols—specifically variables versus numbers—and students' algebraic knowledge impacted seventh graders' problem-solving strategies and use of systemic structures within an online algebra game where students could dynamically transform expressions. We found that on simple problems, symbols did not impact students' strategy efficiency, although students with higher vs lower algebraic knowledge were more efficient at solving these problems. On complex problems, students with higher algebraic knowledge were *more* efficient at solving variable vs numerical problems, whereas students with lower algebraic knowledge were *less* efficient at solving variable vs numerical problems. The visualizations and examination of first steps further revealed that whether students leveraged systemic structures during their problem-solving varied across problems. The findings have implications for research on cognitive processes of symbols and practices on teaching mathematics.

1. Introduction

Symbols are cognitive tools that support reasoning, problem solving, and higher-level thinking (DeLoache, 2004; Koedinger et al., 2008; Vygotsky, 1978). However, learners' struggles with abstract symbols are documented across development and domains (Fyfe et al., 2014). In mathematics, a defining challenge of algebraic learning is the shift from working with concrete numbers that can be calculated to reasoning about the abstract notion of variables—letters that represent a range of unknown values and denote a systemic relation within expressions (Küchemann, 1981). For instance, a previous study has revealed that middle school students struggled with problems presented in variables more so than problems presented in numbers (Chan, Smith, Closser, Drzewiecki, & Ottmar, 2021). While the transition from numbers to variables is typically challenging for students, introducing and using algebraic symbols, such as variables, in early years can promote a smooth transition (Blanton & Kaput, 2005; Carpenter et al., 2005). Prior work has emphasized instructional and developmental sequences that use concrete numbers to broadly support how students think about abstract unknown variables (Fyfe et al., 2014; Koedinger & Anderson, 1998). However, less is known about how individual differences in students' mathematical knowledge may moderate learning and problem solving with variables. In the current study, we investigate how seventh-grade students' strategies may differ when solving matched problems presented in variables versus numbers and explore how their algebraic knowledge may interact with the symbols presented in notation to influence their algebraic problem solving. By focusing on seventh-grade students who are just beginning to work with variables on a regular basis (NGA Center & CCSSO, 2010), this work provides insights into how students may solve problems as they transition from reasoning about numbers to variables, as well as contributes new findings that inform instructional practices to better support this transition.

1.1. Mathematical structures and strategies

Noticing structures, especially systemic structures, in mathematics is an important foundation for learning algebra (Kaput, 1998; Venkat et al., 2019). There are two different types of "structure": *surface structure* and *systemic structure* (Kieran, 1989). Surface structure refers to how the terms and operands are presented within an expression. Surface structure includes individual features that make up an expression, such as the specific symbols used and the locational proximity between terms

* Corresponding author.
E-mail address: chanjyc@eduhk.hk (J.Y.-C. Chan).

[1] Dr. Closser is currently a Postdoctoral Research Associate in Human Development and Family Studies at Purdue University.

https://doi.org/10.1016/j.cedpsych.2022.102114

Available online 22 September 2022

within an expression (i.e., whether the terms to be combined are adjacent to each other). For instance, "3 + 5 – 3" has the surface structure of 3 followed by + 5 then − 3. Systemic structure refers to the underlying mathematical properties within an expression, such as commutativity, associativity, and distributivity. For instance, leveraging the systemic structure of "3 + 5 – 3" involves recognizing the inverse relation between 3 and − 3, and identifying the opportunity to apply the commutative property in order to simplify the expression to 5.

Understanding and noticing the systemic structure of expressions are important aspects of flexible and efficient problem solving (NGA Center & CCSSO, 2010; Rittle-Johnson & Star, 2007; Schneider et al., 2011). Mathematicians and mathematics teachers appreciate the elegance of efficient strategies involving the fewest number of steps and use them frequently in problem solving (Star & Newton, 2009). Efficient strategies allow students to solve problems with minimal time, effort, or cognitive resources, and to reserve these resources for learning and solving more challenging problems. However, many students struggle to leverage systemic structures while solving problems (Newton et al., 2020), especially when the terms to be combined are not adjacent to each other and the surface structure regarding the locational proximity of terms does not support uses of efficient strategies (Lee, Hornburg, Chan, & Ottmar, 2022). For instance, 3 and − 3 are non-adjacent to each other in "3 + 5 – 3", and students may calculate from left to right in response to the surface structure (3 + 5 – 3 = 8 – 3, 8 – 3 = 5), simplifying the expression in two steps. Alternatively, recognizing that 3 and − 3 are opposite numbers, students can use the systemic structure and combine the opposites to efficiently simplify to 5 in one step.

Although students can use the surface structure to compute and simplify numerical expressions or the systemic structure to combine the opposites, only the latter approach applies to expressions involving variables. For example, in "$x + y - x$", students cannot simply calculate from left to right. Rather, they need to notice the systemic structure and combine the opposites (i.e., x and $-x$) in order to isolate y. Because understanding the properties of opposites and inverses is an important aspect of mathematics and this understanding supports algebraic reasoning as well as efficient problem solving (Vamvakoussi & Vosniadou, 2004), we focus on opposites and inverses (hereafter opposites/inverses) as a systemic structure that students can leverage during problem solving. We examine how variation in symbols (i.e., presenting problems with identical systemic structure in either variables or numbers) impacts students' problem-solving strategies.

1.2. Variables and numbers

Unlike numbers that represent specific, known, and concrete values, variables are symbols that can be used to represent unknown values, relations between terms, or generalized patterns or structures in arithmetic (Usiskin, 1988). For example, in $5x + 2 = 12$, the variable x represents an unknown value to be solved. In the formula, $A = LW$, variables are used to indicate relations between area, length, and width, and they can be substituted with specific numbers when solving area problems (Malisani & Spagnolo, 2009). Further, in $a + b = b + a$, the variables are used to illustrate a generalized principle of commutativity in addition (Bardini et al., 2005). These varying uses of variables combined with their abstract conceptualization contribute to the challenges students face when learning algebra (Usiskin, 1988).

Students can solve a numerical problem by relying on the surface structure of the notation and carrying out operations with numbers; however, they need to notice the systemic structure within variable problems to simplify them (Malisani & Spagnolo, 2009). Students often struggle to solve variable problems because these problems require students to understand the meaning of variables, operate with unknowns, and make explicit relations between the terms (Filloy & Rojano, 1989; Herscovics & Linchevski, 1994; Malisani & Spagnolo, 2009; Philipp, 1992). While students can reason algebraically without the use of variables (Radford, 2006; Sfard & Linchevski, 1994), they may struggle to make sense of the symbolic letters and fail to recognize that, in certain situations, these letters represent variables which can be replaced by *any* number (Bardini et al., 2005; Malisani & Spagnolo, 2009). As variables can be used to represent generalized relations and can be substituted with varying or multiple values (Philipp, 1992; Schoenfeld & Arcavi, 1988), variables and their complex conceptualization may influence the ways in which students leverage the systemic structure of expressions when solving algebraic problems that move beyond concrete numbers.

While much work has shown that students struggle with variables and algebraic notations, presenting expressions with variables instead of numbers may actually facilitate learning of algebraic concepts because it removes the opportunities and impulse to calculate (Givvin et al., 2019; Stacey & MacGregor, 1999). Many students view mathematics as a set of numbers to calculate or procedures to memorize, rather than a set of concepts to understand (Garofalo, 1989). Prior work has demonstrated that presenting problems in numbers as opposed to variables draws students' attention to calculation procedures rather than the common structure between algebraic problems, hindering their learning of the underlying concepts (Lawson et al., 2019). On the contrary, presenting problems in variables instead of numbers limits calculations, and may encourage students to focus on the systemic structure and strategically leverage the structure in problem solving.

Although there are notable differences between variables and numbers, they share many of the underlying mathematical principles (Booth, 1981, 1984). Students can leverage systemic structures and apply mathematical properties, such as commutativity, to transform or simplify both variable and numerical expressions (Kieran, 1989); however, they often struggle to do so (Warren, 2003). In a preliminary study, sixth- and seventh-grade students rarely leveraged systemic structures to simplify variable or numerical expressions, and their struggles were particularly apparent on problems presented in variables (Chan, Smith, Closser, Drzewiecki, & Ottmar, 2021). Specifically, students took more steps and attempted problems more times when the problems were presented in variables as opposed to numbers, aligning with the literature on students' struggles with variables. Further, students' step and attempt counts did not vary by their mathematical knowledge. However, the sample was relatively small, most students were in advanced sixth-grade mathematics class, and only four problems were included in the prior analyses, warranting further investigation on the roles of symbols and students' mathematical knowledge in their problem-solving strategies.

1.3. Mathematical knowledge and strategy efficiency

In the context of mathematics equation-solving, *strategy efficiency* has been operationalized as using a strategy with the fewest steps and/or the computation that involves simple (e.g., small or whole numbers) rather than complex numbers (e.g., large numbers, fractions; Xu et al., 2017). To implement efficient problem-solving strategies to a mathematical problem flexibly and adaptively, students need to have the knowledge of underlying mathematical concepts and procedures (Robinson et al., 2018; Star & Rittle-Johnson, 2008). Conversely, students' strategy efficiency may also influence their performance on algebra assessments, especially when the assessments tap students' procedural knowledge in algebraic equation solving. This potentially bidirectional association between knowledge and strategy use has been demonstrated across different age groups and mathematics topics (Star & Rittle-Johnson, 2008; Torbeyns et al., 2006). For instance, on the topic of algebraic equation solving, middle school students with higher mathematics achievement are more likely to use a more efficient strategy that involves fewer steps compared to those with lower mathematics achievement (Newton et al., 2020; Wang et al., 2019).

One potential mechanism underlying the association between mathematical knowledge and strategy efficiency may be students' abilities to identify and leverage systemic structures (Rittle-Johnson &

Star, 2007; Schneider et al., 2011; Xu et al., 2017). In particular, students with higher vs lower mathematical knowledge may be more likely to notice the systemic structures and the opportunities for applying mathematical properties that support efficient problem solving (Chi et al., 1982) as well as transfer of knowledge across problems (Novick, 1988). Although noticing structural similarities between problems supports mathematical learning and problem solving, fifth- and sixth-grade students tend to focus on surface structures rather than systemic structures of equations and expressions (Sidney & Alibali, 2015, 2017).

1.4. The current study

In the current study, we aim to extend prior work by focusing on how symbols—variables vs numbers—impact students' strategy efficiency, as measured by the number of problem-solving steps. In particular, we examine whether seventh-grade students are more or less efficient at solving variable vs numerical problems by quantitatively comparing the number of steps that students take to solve these problems. We also investigate whether students' algebraic knowledge moderates the effect of symbol on their problem-solving strategies. In addition to the quantitative analyses, we qualitatively compare students' strategies by visualizing the problem-solving processes as well as the mathematical principles they use on these two types of problems.

Further, other non-systemic factors, such as the locational proximity of symbols (Lee, Hornburg, Chan, & Ottmar, 2022), the complexity of the problems (e.g., addition/subtraction or multiplication/division inversion problems; Robinson et al., 2006), and students' prior experience working with problems of the same structure (Sidney, 2020) may also impact their use of systemic structures as well as strategy efficiency. Therefore, we account for these factors to examine how symbols and students' algebraic knowledge independently impact their problem-solving strategies. Specifically, we designed four pairs of problems that were identical in systemic structure but varied in whether they were presented with variables (e.g., $x + y - x$) or numbers (e.g., $3 + 5 - 3$). We also systematically varied the complexity of problems, the locational proximity (hereafter proximity) of symbols, and the order in which variable and numerical problems were presented. Our specific research questions (RQs) and the corresponding hypotheses are pre-registered on the Open Science Framework (https://osf.io/tpnu2/) and are as followed:

(RQ1) Confirmatory: Does students' strategy efficiency, as indicated by their number of problem-solving steps, vary on problems presented in variables vs numbers?

Hypothesis 1: Students may use less efficient strategies by taking more steps on problems presented in variables as opposed to numbers.

(RQ2) Confirmatory: Does the effect of symbol on students' strategy efficiency vary by students' algebraic knowledge?

Hypothesis 2: Students with lower algebraic knowledge may use less efficient strategies by taking more steps on variable problems as opposed to numerical problems; this effect may be smaller or insignificant among students with higher algebraic knowledge.

(RQ3) Exploratory: Do the effects of symbol and algebraic knowledge on students' strategy efficiency remain after controlling for other non-systemic factors, such as the proximity of the terms to be combined, the order in which the problems were presented, and the complexity of the problems?

Hypothesis 3: Based on related prior research, we expect that students may use more efficient strategies when the opposites/inverses are adjacent vs non-adjacent to each other, when they solve the second vs first of the paired problems, and when the problems are simple vs complex. We will explore whether the effects of symbol and its interaction with algebraic knowledge remain after accounting for these potential influences.

(RQ4) Confirmatory: How do symbols impact students' use of systemic structures during problem solving?

Hypothesis 4: Based on prior research suggesting that students are more efficient on numerical vs variable problems, we hypothesize that students may be more likely to use systemic structures by leveraging mathematical properties on problems presented in numbers as opposed to variables.

We quantitatively address our first three research questions using the number of steps students take to complete the problems as the focal dependent variable. We qualitatively address our final research question by visualizing and comparing students' problem-solving processes on variable and numerical problems, paying close attention to whether their first step demonstrates leveraging the systemic structure of the problem. We first address our research questions on two pairs of simple problems involving addition and subtraction, then repeat the analyses on another two pairs of complex problems involving all four arithmetic operations in order to test the robustness of the findings.

2. Methods

2.1. Participants

The sample for this within-subjects study was drawn from a larger randomized controlled trial conducted with ten in-person schools and one virtual school in one large, suburban district in the Southeastern United States during the 2020–2021 school year amidst the COVID-19 pandemic. The larger study aimed to examine the efficacy of three educational technologies (see Decker-Woodrow et al., forthcoming). In the current study, we focused our analysis on the 445 seventh-grade students who were randomly assigned to playing From Here to There!, completed a pretest on algebraic knowledge, and completed the four pairs of problems in the From Here to There! game designed to address our research questions. In this game, students can manipulate algebraic symbols on the screen in real-time (Ottmar, Landy, Goldstone, & Weitnauer, 2015); all students' actions are recorded for each problem so we can examine their problem-solving processes and strategy efficiency. The other two technologies (i.e., online problem sets and an equation-solving game) did not provide the appropriate log data for our research questions, thus were not included in the current study.

A total of 911 students started the game, however, only 445 students completed all four pairs of problems in the game designed for this study. The analytic sample of 445 was much smaller than the original sample because (a) two schools opted out during the first two weeks of the study due to instructional concerns related to COVID-19 and (b) the paired problems were embedded in the middle of the game, requiring students to play at their own pace for several hours (problem 79 to 119; see From Here to There! section for more detail).

Among the 445 students (54 % male, 46 % female), 36 % were in advanced mathematics classes that were designed for students who excelled in mathematics, and 64 % were in on-level classes that implemented a typical mathematics curriculum. We received race and ethnicity information on 444 students from the school district. Of the 444 students, 51 % were White, 30 % were Asian, 12 % were Hispanic, 3 % were Black, and 4 % were of other races/ethnicities. Due to COVID-19, 70 % of the students (and their families) chose to receive their instruction in-person at the start of the Fall 2020, and the remaining 30 % of students (and their families) chose to receive virtual instruction.

This research was approved by the Institutional Review Board at a University in the Northeastern United States. This research involved typical educational practices and did not require parental consent. Instead, parents received a letter informing them about the research and the data collected from the educational technologies. The parents had the opportunity to opt their child out of this study.

2.2. Procedure

The larger study consisted of nine 30-minute intervention sessions, and four 45-minute assessment sessions to be completed before, during, immediately after, and two months after the intervention, respectively

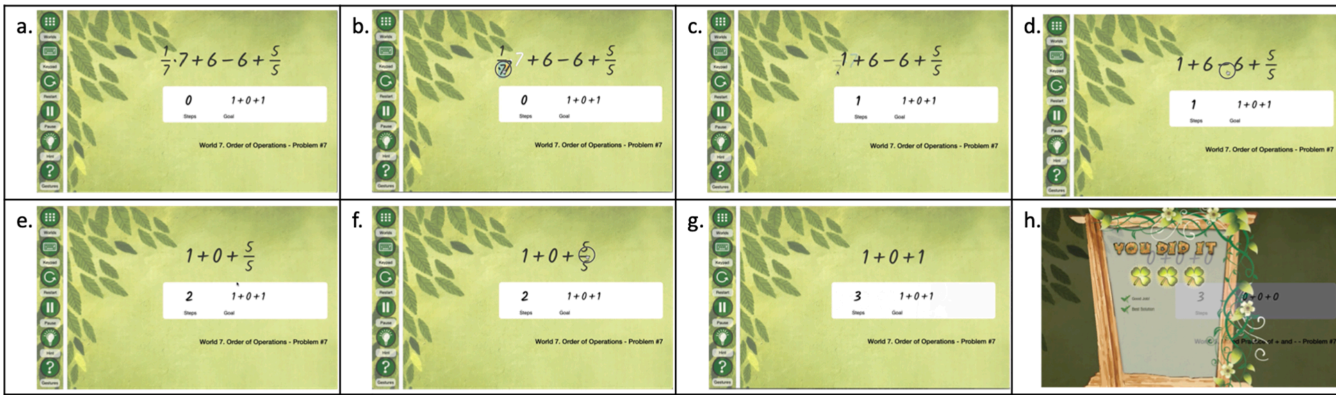

Fig. 1. A sample problem in From Here to There! (a) and a potential transformation process involving three steps (b, c, d, e, f) to reach the goal (g, h).

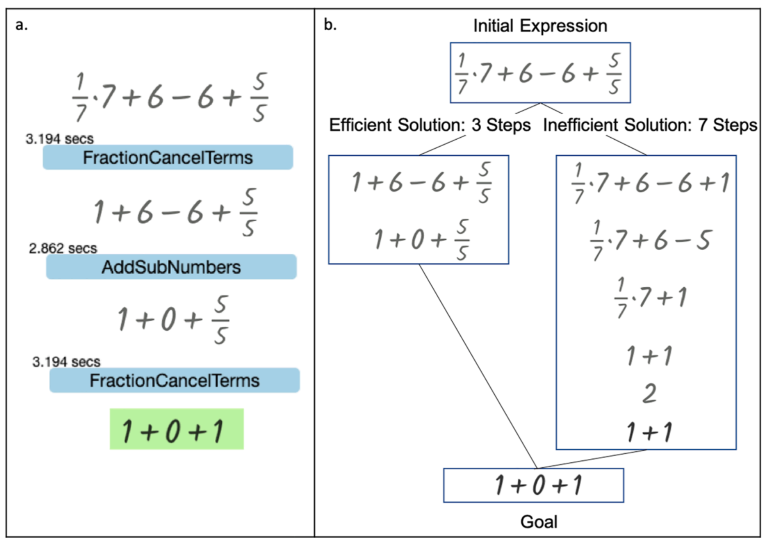

Fig. 2. (a) A visualization of the log data on a student's problem-solving process for a problem in From Here to There! (b) An illustration of two different problem-solving processes.

(see details in Decker-Woodrow et al., forthcoming). All sessions were administered online as a part of students' regular mathematics instruction. During each intervention session, students worked individually through the problems within the game at their own pace using a device. The system automatically ended the game and saved students' progress after 30 min. In each subsequent session, students continued the game from the problem on which they left off. In the current study, we used the log data recorded on the four pairs of problems within the game and the algebra pretest to address our research questions, thus we only described the tasks relevant to the current study.

2.3. From Here to There!

In From Here to There! game, mathematical topics were organized into 14 worlds, each containing 18 problems that provided practices on a mathematical topic (e.g., factoring). All students worked on the problems in the same order starting from basic arithmetic operations to more complex topics, such as fractions, distributions, and algebraic equations. For each problem, students were presented with an initial expression and a mathematically equivalent goal (Fig. 1a). The objective was to transform the expression into the specified goal using a series of gesture-actions (e.g., tapping or dragging) that transform expressions from one state to another. A sample problem with a series of steps—gesture-actions that lead to valid transformations—between the initial expression and the goal is illustrated in Fig. 1. In this example, students were asked to transform $1/7 \cdot 7 + 6 - 6 + 5/5$ into $1 + 0 + 1$ (Fig. 1a). The student first *multiplied 7 with* $1/7$ by dragging 7 on top of the denominator (Fig. 1b and 1c). Next, the student *subtracted 6 from 6* by tapping the subtraction sign (Fig. 1d and 1e). Last, the student *simplified* $5/5$ by tapping the division bar (Fig. 1f), reaching the goal $1 + 0 + 1$ (Fig. 1g). A rewards board appeared showing the number of clovers (one, two, or three) awarded to the student based on their strategy efficiency as measured by the number of steps taken to reach the goal (Fig. 1h). Students received more clovers for using more efficient strategies that involved fewer steps. Further, the step count (in the bottom right next to the goal; Fig. 1) turned red when students' steps exceeded the minimum number required to complete the problem. In these ways, students received feedback on strategy efficiency and were encouraged to take the fewest possible steps.

All student actions and the corresponding transformations of mathematical expressions were time-stamped and recorded (Fig. 2a), allowing us to compare students' equation-solving processes in ways not accessible in answer-based learning systems or paper-and-pencil tasks. Furthermore, students could take any series of mathematically valid steps that link the initial expression and the goal (Fig. 2b). The game thus provided an ideal context for examining variation in problem-solving processes and strategy efficiency (See detailed information on the game in Chan, Lee, Mason, Sawrey, & Ottmar, 2022).

Table 1
Mean, standard deviation, minimum, maximum, skewness, and kurtosis of students' total steps on each focal problem in the game.

Problem Pair	Problem Number	Problem	Order	Proximity	M (SD)	Min-Max (min steps)[a]	Skewness	Kurtosis
Simple Problems – World 5								
1.1	79	$9-9+8-8+7-7 \rightarrow 0+0+0$	Second	Adjacent	3.38 (1.00)	3 – 6 (3)	2.22	2.93
1.1	78	$x-x+y-y+z-z \rightarrow 0+0+0$	First	Adjacent	4.05 (1.83)	3 – 8 (3)	1.31	−0.03
1.2	81	$3+4+5-3-4-5 \rightarrow 5-5$	First	Non-adjacent	6.42 (3.35)	4 – 15 (3)	1.14	−0.05
1.2	82	$a+b+c-a-b-c \rightarrow c-c$	Second	Non-adjacent	5.96 (2.95)	4 – 14 (3)	1.65	1.72
Complex Problems – World 7								
2.1	115	$1/7 \cdot 7+6-6+5/5 \rightarrow 1+0+1$	First	Adjacent	8.14 (4.73)	3 – 20 (3)	1.21	0.56
2.1	116	$1/c \cdot c+b-b+a/a \rightarrow 1+0+1$	Second	Adjacent	5.08 (2.41)	3 – 11 (3)	1.48	1.25
2.2	119	$(5 \cdot 8 \cdot 1/5-8)/7 \cdot 0 \cdot 7 \rightarrow 0 \cdot 0$	Second	Non-adjacent	6.99 (4.47)	2 – 19 (4)	1.11	0.74
2.2	118	$(x \cdot y \cdot 1/x-y)/z \cdot 0 \cdot z \rightarrow 0 \cdot 0$	First	Non-adjacent	11.39 (9.29)	2 – 36 (4)	1.40	1.23

Note. The problem number indicates the order in which the problem is presented to students within the game.
Order indicates the order of the problems within the pair.
Proximity indicates whether the opposites and inverses are adjacent or non-adjacent to each other. Abbreviations: M = Mean, SD = Standard Deviation, Min = Minimum, Max = Maximum.
[a] The values in parentheses represent the minimum number of steps required to complete each problem by combining opposites/inverses.

2.3.1. *Problem structures*

We designed and embedded two simple pairs and two complex pairs of problems into the game. The two simple pairs only involved the opposites, and they were placed in World 5: *Mixed Practice of Addition and Subtraction*; the two complex pairs involved both opposites (addition and subtraction) and inverses (multiplication and division), and they were placed in World 7: *Order of Operations*. Within each pair, problems were matched on the systemic structure of the initial expression and the goal; the paired problems only varied on whether the symbols were numbers or variables (Table 1). For example, the numerical problem in Pair 1.1 was to transform "$9-9+8-8+7-7$" into "$0+0+0$", and its variable counterpart was to transform "$x-x+y-y+z-z$" into "$0+0+0$". In the numerical problem, students could tap on the three subtraction signs to subtract 9 from 9, 8 from 8, and 7 from 7 to reach the goal of $0+0+0$. Similarly, students could tap on the three subtraction signs in the variable problem to reach the goal. In this example, the paired problems both required three steps—combining the opposites three times.

The proximity of the opposites/inverses to be combined (adjacent vs non-adjacent) was counterbalanced between pairs of problems. The presentation order of numerical and variable problems was also counterbalanced between pairs. The order in which the students received the four pairs of problems (all eight problems) was the same for all students. For primary analyses, we dummy coded the paired problems and used the numerical problems as the reference group. We used this dummy coded variable to test the effect of symbol (numbers vs variables) on students' strategy efficiency for the first research question, and its interaction with algebraic knowledge for the second research question.

2.3.2. *Total steps in From Here to There!*

All student actions, including clicks and drags, were logged within the game. Using the log data, we computed the total number of steps students took to complete a problem and used it as a measure of strategy efficiency. As an example, the student in Fig. 1 took a total of three steps (also see Fig. 2a). We used students' total steps to complete a problem as the dependent variable in our analyses for the first three research questions. Further, we used the log data to visualize students' problem-solving processes for the final research question.

2.4. *Algebraic knowledge assessment at pretest*

Students' algebraic knowledge was assessed with 10 multiple-choice items selected from a previously validated measure (Star et al., 2014; Cronbach's $\alpha = 0.89$). The 10 items were selected because they measured students' conceptual knowledge, procedural knowledge, and procedural flexibility in algebraic equation solving, capturing aspects of students' mathematical knowledge relevant to the intervention. Two sample items were "solve for y in the equation $5(y-2) = 3(y-2)+8$" and "identify the expressions that are equivalent to $4(n+3)$". While the paired problems in the game focused specifically on simplifying expressions using systemic structures of opposites/inverses, the pretest focused more broadly on students' knowledge relevant to algebraic equation solving, tapping into students' algebraic understanding beyond opposites/inverses. Each item was scored as correct (1) or incorrect (0). The reliability of the assessment was good, Kuder-Richardson Formula 20 = 0.74. Kuder-Richardson Formula 20 is a reliability measure developed for binary variables; the values range between 0 and 1, and higher values represent higher reliability (Kuder & Richardson, 1937). The total score on the assessment at pretest was included as a covariate for the first research question, and its interaction with symbol was included as a focal fixed effect for the second research question.

2.5. *Approach to analysis*

First, we conducted descriptive and correlational analyses to examine the distribution of and relations between variables, as well as to inform our primary analyses. Because students could take as many steps as needed to reach the goal, we used the interquartile-range methods to replace outliers (Walfish, 2006). This method extracted the top and bottom 25 % of values from the data. Within these two quartiles, the values that were beyond 1.5 times the interquartile range were considered as outliers. These values were then replaced with observed values at either the fifth or ninety-fifth percentile. This method allowed us to retain all participants in the analyses while avoiding the results being distorted by the influential cases.

We addressed the first three research questions by conducting a series of mixed-effect regression models using the *lme4* package (Bates et al., 2015) with maximum likelihood estimation in R. In each model, we included students as a random effect to account for the repeated measures (i.e., all students completed all paired problems in the game). We considered the nesting structure of students within teachers or schools; however, the intra-class correlation was 0.01 for teacher and 0.003 for school, indicating that 1 % or less of the variance in step count, our primary dependent variable, was attributable to teacher or school, respectively. Because the values were well below the 0.07 threshold for including the nesting structure (Niehaus et al., 2014) and we did not have teacher- or school-level fixed effects, we proceeded with parsimonious models without nesting.

To address the first research question on the effect of symbol, we conducted a mixed-effect model with total steps as the dependent variable, symbol (variables vs numbers) as the problem-level fixed effect, and algebraic knowledge as the student-level fixed effect. To address the second research question on the potential moderating effect of algebraic knowledge, we built on the first model by adding the Symbol × Algebraic Knowledge interaction as a fixed effect. To address the third research question on the unique influences of symbol and algebraic knowledge, we built on the second model by adding proximity and

Table 2
Beta coefficients, standard errors, and exponents of the coefficients in the mixed-effect models predicting students' total steps on simple and complex problems.

Fixed Effects	Simple Problems – World 5 Model 1.1 *B* (*SE*) exp(*B*)	Model 1.2 *B* (*SE*) exp(*B*)	Model 1.3 *B* (*SE*) exp(*B*)	Complex Problems – World 7 Model 2.1 *B* (*SE*) exp(*B*)	Model 2.2 *B* (*SE*) exp(*B*)	Model 2.3 *B* (*SE*) exp(*B*)
Intercept	1.60 (0.02)*** 4.93	1.60 (0.02)*** 4.93	1.56 (0.02)*** 4.76	2.02 (0.02) *** 7.54	2.02 (0.02) *** 7.52	1.98 (0.02) *** 7.21
Algebraic Knowledge	−0.03 (0.01) *** 0.97	−0.03 (0.01) *** 0.97	−0.03 (0.01)*** 0.97	−0.01 (0.01) 0.99	−0.01 (0.01) 0.99	−0.01 (0.01) 0.99
Symbol: Variable	0.2 (0.2) 1.02	0.02 (0.02) 1.02	0.05 (0.02)* 1.06	0.05 (0.02)** 1.05	0.04 (0.02)** 1.04	−0.04 (0.02)* 0.96
Symbol × Algebraic Knowledge	–	0.001 (0.01) 1.001	0.001 (0.01) 1.001	–	−0.05 (0.01) *** 0.95	−0.05 (0.01)*** 0.95
Order: First	–	–	0.13 (0.02) *** 1.14	–	–	0.46 (0.02)*** 1.58
Proximity: Non-adjacent	–	–	0.52 (0.02)*** 1.68	–	–	0.37 (0.02) *** 1.45
Marginal R^2	0.03	0.03	0.25	0.01	0.02	0.22
Conditional R^2	0.25	0.25	0.42	0.62	0.62	0.70

Note. Symbol: Variable = 0.5, Number = −0.5; Order: First = 0.5, Second = −0.5; Proximity: Non-adjacent = 0.5, adjacent = −0.5.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

problem order as problem-level fixed effects to account for their potential influences on students' strategy efficiency. We first conducted the analyses with the simple paired problems in World 5, then repeated the analyses with the complex paired problems in World 7 to explore the consistency of the findings. We also explored whether the pattern of results remained when we included all four pairs of problems in one model and added problem complexity (i.e., World) as a problem-level fixed effect. To aid the interpretation of the results, we grand-mean centered the problem-level (i.e., symbol, proximity, order, complexity) and student-level (i.e., algebraic knowledge score) fixed effects.

To address the fourth research question on how symbols impacted students' problem-solving strategies, we visualized and qualitatively compared students' problem-solving processes on problems presented in variables vs numbers. Specifically, for each problem, we created a Sankey diagram to illustrate the variation in the problem-solving paths taken by students. In a Sankey diagram, each vertical line represents a step in the problem-solving process; the height of the vertical lines represents the number of students taking that step; the horizontal paths between vertical lines connect the steps to form the processes that students took to solve the problem (see Lee, Stalin, Drzewiecki, Trac, & Ottmar, in press for details). Using the Sankey diagrams, we graphed and descriptively compared the percentage of students who combined the opposites/inverses or moved them adjacent to each other as their first step on each problem. By focusing on students' first step, we illustrated the proportion of students who initially leveraged the systemic structure on each problem and whether their steps varied by the symbols in which the problems were presented.

3. Results

3.1. Preliminary analyses

A review of the descriptive statistics revealed that students scored an average of 5.76 points ($SD = 2.65$) out of 10 on the algebraic knowledge pretest, indicating that the scores were not subject to floor or ceiling effects. Further, the algebraic knowledge scores were widely (minimum = 0, maximum = 10) and normally distributed (skewness = −0.11, Kurtosis = −1.06), indicating that the sample captured a wide range of students at varying knowledge levels. Students' total step count across the eight problems, as well as their algebraic knowledge score descriptively varied by their course level, biological sex, and race/ethnicity (see Appendix A Table A1). To examine the relative contribution of algebraic knowledge, course level, biological sex, and race/ethnicity on students' total step count, a regression was conducted. After accounting for students' algebraic knowledge—a planned focal fixed effect in primary analyses—the demographic variables did not significantly predict step count (see Appendix A Table A2). Therefore, for all primary analyses, we conducted parsimonious models including algebraic knowledge and excluding demographic variables.

Next, a descriptive analysis of students' total step count on each focal problem revealed that their total steps were widely distributed, as indicated by the minimum, maximum, skewness, and Kurtosis (Table 1). Because students' total steps were skewed towards the minimum, as indicated by the means, we conducted Poisson instead of linear models to address the first, second, and third research questions. We noted that, descriptively, students tended to take fewer steps on the second as opposed to first of the paired problems. They also tended to take fewer steps on problems where the opposites/inverses were adjacent to each other as opposed to non-adjacent. Because of the prior research and these descriptive findings, we included proximity and order as covariates for our third research question.

3.2. Does students' strategy efficiency vary on problems presented in variables vs numbers?

To address our first research question, we conducted a mixed-effect model with the two simple problem pairs in World 5, then repeated the analyses on the complex problem pairs in World 7. In each model, we included the symbol of the problem (variables vs numbers) as a dummy-coded problem-level fixed effect, and pretest algebraic knowledge as a student-level fixed effect.

We found that whether the problems were presented in variables or numbers did not significantly impact students' strategy efficiency on simple problems, $p = .343$. However, students with higher algebraic knowledge tended to use more efficient problem-solving strategies that involved fewer steps, $p < .001$ (Table 2, Model 1.1). In particular, students who scored one point higher than average on the algebraic knowledge assessment took 3 % (i.e., 1 – 0.97) fewer steps to solve these problems. The fixed effects together accounted for 3 % of the variance in students' total steps.

We repeated the analyses with complex problems in World 7 and found a different pattern of results. The effect of symbol was significant, $p = .004$, but the effect of algebraic knowledge was not, $p = .083$. Compared to complex problems presented in numbers, students took 5

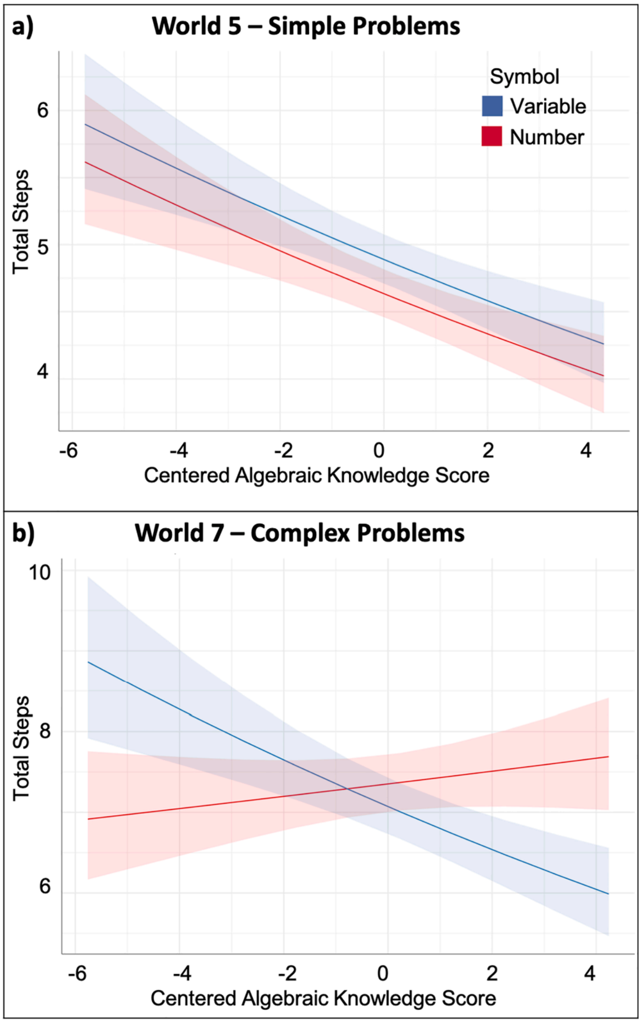

Fig. 3. The interaction of symbol and algebraic knowledge on students' total problem-solving steps on simple problems in World 5 (a) and complex problems in World 7 (b). The bandwidths represent 95% confidence intervals of the estimates.

Table 3
Beta coefficients, standard errors, and exponents of the coefficients in the mixed-effect model including all four pairs of problems.

Fixed Effects	Model 3 *B* (*SE*)	exp(*B*)
Intercept	1.79 (0.02)***	5.97
Algebraic Knowledge	−0.02 (0.01) ***	0.98
Symbol: Variable	0.03 (0.01) *	1.03
Symbol × Algebraic Knowledge	−0.02 (0.005) ***	0.98
Order: First	0.33 (0.01) ***	1.39
Proximity: Non-adjacent	0.43 (0.01) ***	1.53
Complexity: Complex	0.50 (0.01) ***	1.65
Algebraic Knowledge × Complexity	0.02 (0.005) ***	1.02
Symbol × Complexity	−0.12 (0.03) ***	0.89
Algebraic Knowledge × Symbol × Complexity	−0.05 (0.01) ***	0.95
Marginal R^2	0.37	
Conditional R^2	0.63	

Note. Symbol: Variable = 0.5, Number = −0.5; Order: First = 0.5, Second = −0.5; Proximity: Non-adjacent = 0.5, adjacent = −0.5.
* $p < .05$.
*** $p < .001$.

% more steps on complex problems presented in variables (Table 2, Model 2.1). The fixed effects accounted for 1 % of the variance in students' total steps.

3.3. *Does the effect of symbol on students' strategy efficiency vary by algebraic knowledge?*

We built on the models for RQ1 by adding an interaction term between symbol and algebraic knowledge. We found that the Symbol × Algebraic Knowledge interaction was not a significant fixed effect of students' strategy efficiency on simple problems, $p = .916$ (Table 2, Model 1.2; Fig. 3a). However, the interaction was a significant fixed effect of students' strategy efficiency on complex problems, $p < .001$ (Table 2, Model 2.2; Fig. 3b). Specifically, students with lower algebraic knowledge took more steps on complex variable vs numerical problems, whereas students with higher algebraic knowledge took fewer steps on complex variable vs numerical problems. Post-hoc analyses further revealed that, while students with higher algebraic knowledge took significantly fewer steps on complex variable problems compared to students with lower algebraic knowledge, $p = .001$, the relation between total steps and algebraic knowledge was not statistically significant for complex numerical problems, $p = .578$. In other words, students' strategy efficiency on complex variable problems varied by their algebraic knowledge level, but they were equally efficient at solving complex numerical problems regardless of their algebraic knowledge level. The fixed effects accounted for 2 % of the variance in students' total steps.

3.4. *Do the effects of symbol and algebraic knowledge on strategy efficiency remain after controlling for the order of problems and proximity of terms to be combined?*

To explore whether the order of the problems and the proximity of the terms to be combined impacted the pattern of our findings, we built on the models from our second research question by adding problem order and term proximity as problem-level fixed-effects. We found that both order and proximity significantly impacted students' strategy efficiency. Specifically, on simple problems, students took 14 % more steps on the first vs second of the paired problems, $p < .001$, while other fixed effects were held constant. Students took an average of 68 % more steps when the opposites were non-adjacent vs adjacent to each other, $p < .001$. The fixed effects of order and proximity accounted for 22 % of the variance in the model (Table 2, Model 1.3). Similarly, on complex problems, order and proximity significantly predicted students' strategy efficiency. Students took an average of 58 % more steps on the first vs second of the paired problems, $p < .001$; they took an average of 45 % more steps when the opposites/inverses were non-adjacent vs adjacent to each other, $p < .001$ (Table 2, Model 2.3). These two fixed effects accounted for 20 % of the variance in the model. Although these non-systemic factors significantly impacted students' strategy efficiency, the effects of algebraic knowledge and Symbol × Algebraic Knowledge interaction remained significant for simple and complex problems, respectively.

Finally, we explored whether the pattern of the results remained when both the simple and complex problems were included in one model. We found that the Symbol × Complexity, Algebraic Knowledge × Complexity, as well as the Symbol × Algebraic Knowledge × Complexity interactions were significant, *ps* $< .001$. These interactions indicated that the effects of symbol, algebraic knowledge, and Symbol × Algebraic Knowledge interaction differed between simple and complex problems (Table 3 and Fig. 3). Together, the findings aligned with the results reported for the first and second research questions.

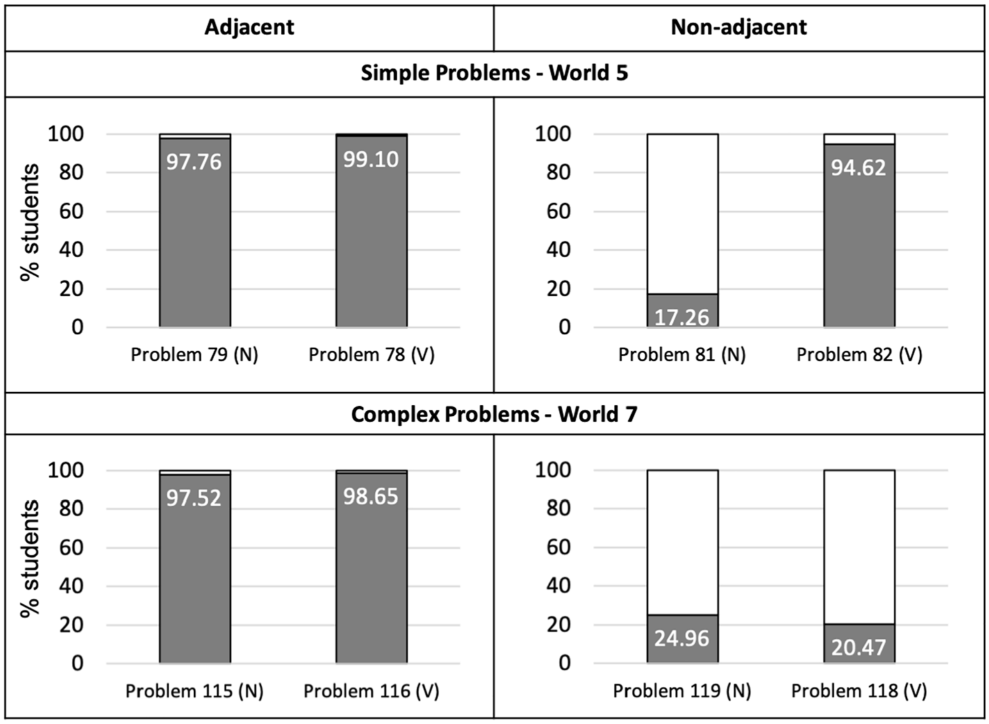

Fig. 4. The percentage of students combining opposites/inverses or moving opposites/inverses adjacent to each other as their first step on numerical and variable problems. *Note.* Gray bars represent the percentage of students who either combine the opposites/inverses or move them adjacent to each other. White bars represent the percentage of students who take other types of steps. *Abbreviations*: N = Number, V = Variable.

3.5. How do symbols impact students' use of systemic structures during problem solving?

To qualitatively examine how students' problem-solving strategies varied by problems presented in numbers or variables, we first created Sankey diagrams for each problem to visualize students' problem-solving process (Appendix B). Within our sample, the diagrams showed that students' problem-solving processes were highly variable in terms of the number of steps taken by the students, as indicated by the number of vertical lines (i.e., nodes) within each path, as well as the type of steps, as indicated by the label of each vertical line within each column. Examining the leftmost paths of each diagram provided insights into the most common first steps students took on each problem and whether they leveraged systemic structures in their first step of problem-solving.

For each pair of problems, we graphed and descriptively compared the percentage of students who combined the opposites/inverses or moved them adjacent to each other as their first step (Fig. 4; see Table 4 for the steps that are categorized as combining the opposites/inverses or moving them adjacent to each other). On simple problems, over 97 % of students combined or moved the opposites together when they were adjacent to each other (Problems 78 and 79), regardless of the symbols in which the problems were presented. When the opposites were not adjacent to each other, only 17 % of students combined or moved the opposites together on the numerical problem (Problem 81), whereas 95 % of students did so on the variable problem (Problem 82). Similarly, on complex problems, over 97 % of students combined or moved the opposites/inverses together on both numerical and variable problems when opposites/inverses were adjacent to each other (Problems 115 and 116). When the opposites/inverses were not adjacent to each other, only 20 % (Problem 118) to 25 % (Problem 119) of students leveraged the systemic structure of the problem by combining or moving opposites/inverses together.

4. Discussion

In the current study, we examined how problems that were matched in systemic structures but presented in either variables or numbers influenced students' problem solving. We tested this effect by presenting four problem pairs in variables as opposed to numbers within the context of a mathematical game and examined how this within-student variation in symbols influenced students' strategy efficiency and problem-solving processes on simple and complex problems. Three main findings emerged. First, the effects of symbol on students' problem-solving strategies were not straightforward. Students' strategy efficiency did not vary by variables vs numbers on simple problems, although students with higher algebraic knowledge tended to use more efficient strategies on these problems. Second, students with lower algebraic knowledge were more efficient on complex numerical vs variable problems, whereas students with higher algebraic knowledge were more efficient on complex variable vs numerical problems. Third, both order and proximity significantly impacted students' strategy efficiency, yet the Symbol × Algebraic Knowledge interaction on students' strategy efficiency remained significant for complex problems. We discuss these findings and their implications in detail.

4.1. Nuanced effects of symbol on strategy efficiency and use of systemic structures

We hypothesized that students may use less efficient strategies, indicated by taking more steps, on problems presented in variables as opposed to numbers. Although symbols did not impact students' strategy efficiency on simple problems, we did find an effect of symbol on complex problems and this effect varied by students' algebraic knowledge. Some prior studies indicate that middle school students with higher mathematics achievement are more likely to use efficient strategies involving fewer steps when solving algebraic equations compared to students with lower mathematics achievement (Newton et al., 2020; Wang et al., 2019). Although students' strategy use seems to be tightly

Table 4
A list of the first steps that are categorized as combining or moving opposites/inverses together for each problem.

Number		Variable	
Simple Problems – World 5			
Problem 79.	$9-9+8-8+7-7$	Problem 78.	$x-x+y-y+z-z$
•	$0+8-8+7-7$	•	$0+y-y+z-z$
•	$9-9+0+7-7$	•	$x-x+0+z-z$
•	$9-9+8-8+0$	•	$x-x+y-y+0$
Problem 81.	$3+4+5-3-4-5$	Problem 82.	$a+b+c-a-b-c$
•	$0+4+5-4-5$ or $4+5-0-4-5$	•	$0+b+c-b-c$ or $b+c-0-b-c$
•	$3+0+5-3-5$ or $3+5-3-0-5$	•	$a+0+c-a-c$ or $a+c-a-0-c$
•	$3+4+0-3-4$ or $3+4-3-4-0$	•	$a+b+0-a-b$ or $a+b-a-b-0$
•	commute such as $3-3+4+5-4-5$	•	commute such as $a-a+b+c-b-c$
Complex Problems – World 7			
Problem 115.	$\frac{1}{7}\cdot 7+6-6+\frac{5}{5}$	Problem 116.	$\frac{1}{c}\cdot c+b-b+\frac{a}{a}$
•	$1+6-6+\frac{5}{5}$	•	$1+b-b+\frac{a}{a}$
•	$\frac{1}{7}\cdot 7+0+\frac{5}{5}$	•	$\frac{1}{c}\cdot c+0+\frac{a}{a}$
•	$\frac{1}{7}\cdot 7+6-6+1$	•	$\frac{1}{c}\cdot c+b-b+1$
•	$\frac{1\cdot 7}{7}+6-6+\frac{5}{5}$ or $\frac{7\cdot 1}{7}+6-6+\frac{5}{5}$	•	$\frac{1\cdot c}{c}+b-b+\frac{a}{a}$ or $\frac{c\cdot 1}{c}+b-b+\frac{a}{a}$
Problem 119.	$\frac{5\cdot 8\frac{1}{5}-8}{7}\cdot 0\cdot 7$	Problem 118.	$\frac{x\cdot y\frac{1}{x}-y}{z}\cdot 0\cdot z$
•	$\frac{8\cdot 1-8}{7}\cdot 0\cdot 7$	•	$\frac{y\cdot 1-y}{z}\cdot 0\cdot z$
•	$(5\cdot 8\cdot\frac{1}{5}-8)\cdot 0$	•	$(x\cdot y\cdot\frac{1}{x}-y)\cdot 0$
•	commute such as $\frac{5\cdot\frac{1}{5}\cdot 8-8}{7}\cdot 0\cdot 7$	•	commute such as $\frac{x\cdot\frac{1}{x}\cdot y-y}{z}\cdot 0\cdot z$

related to their mathematical knowledge in prior studies, this relation has been evidenced as weak or inconsistent across studies (Rittle-Johnson & Star, 2007; Schneider et al., 2011; Xu et al., 2017). Our finding extends prior work by showing that students' strategy efficiency is influenced by their algebraic knowledge as well as symbol and other non-systemic factors of the problems. This finding contributes to the larger theory of perceptual learning (Goldstone et al., 2017; Kellman et al., 2010), which suggests that thinking, learning, and problem solving are grounded in the environment and that the ways in which problems are presented influence how individuals process mathematical information.

Visualizations of students' problem-solving processes also do not seem to align with our fourth hypothesis that students may be more likely to leverage systemic structures on numerical vs variable problems. However, the visualizations do provide insights into when students attend to systemic structures, and these insights build on prior research in an important way. Prior research has reported that fifth- and sixth-grade students tend to focus on the surface features of problems, but they can leverage the systemic similarities between problems to support mathematical learning when problems are grouped by systemic structures (Sidney & Alibali, 2015, 2017). Using the Sankey diagrams, we found that students used a wide variety of strategies (Appendix B), and they did not consistently leverage the systemic structures by combining the opposites/inverses to simplify expressions. Even so, they were more likely to leverage the systemic structures when simple problems were presented in variables and when the opposites/inverses were adjacent to each other (Fig. 4).

When the opposites were non-adjacent to each other, 95 % of students combined the opposites as their first step on the simple variable problem, whereas only 17 % of students did so on the simple numerical problem. Further, among the students who did not combine the opposites, about half of them performed calculations to combine 3 and 4 as their first step on the simple numerical problem. One possible explanation for this finding is that students are often trained to solve problems quickly (Schoenfeld, 1992), contributing to their tendency to rush through problems without pausing. Indeed, when presented with numerical problems, students tend to perform operations from left to right (Blando et al., 1989; Bye et al., 2022; Gunnarsson et al., 2016); they may also be inclined to immediately perform calculations instead of noticing the systemic structures (Givvin et al., 2019; Lawson et al., 2019; Stacey & MacGregor, 1999). Conversely, students who pause longer vs shorter before problem solving tend to complete problems with higher strategy efficiency (Chan, Ottmar, & Lee, 2022).

Because students cannot combine unlike terms in the game (e.g., tapping the addition sign in $x + y$ makes the expression shake and the expression remains as $x + y$), presenting problems in variables as opposed to numbers may help students pause and notice the systemic structures, at least on simple problems. The shaking feedback in the game may further provide an attentional cue that combining unlike terms is an invalid mathematical action, prompting students to try a different action. Together, the qualitative findings on the visualizations extend the quantitative results on students' strategy efficiency and provide additional evidence on how symbols may influence students' strategies across problems.

4.2. Differential effects of variable vs numbers based on algebraic knowledge

Partially supporting our second hypothesis, we did find a significant interaction effect between symbol and algebraic knowledge. While students with lower vs higher algebraic knowledge were equally efficient at solving complex numerical problems, students' strategy efficiency varied by their algebraic knowledge on complex variable problems. Specifically, students with lower algebraic knowledge were *less* efficient at solving complex variable vs numerical problems, whereas students with higher algebraic knowledge were *more* efficient at solving complex variable vs numerical problems. While lower vs higher algebraic knowledge is relative to our overall high-performing sample, these findings extend prior work that shows (a) students tend to struggle with problems presented in variables (Filloy & Rojano, 1989; Herscovics & Linchevski, 1994; Malisani & Spagnolo, 2009; Philipp, 1992), and (b) middle schoolers tend to use less efficient strategies by taking more steps on complex variable vs numerical problems (Chan, Smith, Closser, Drzewiecki, & Ottmar, 2021). In particular, our findings suggest that students may be more susceptible to struggling with problems presented in variables when they have lower vs higher algebraic knowledge and when the problems are complex vs simple.

One possible explanation is that students with lower algebraic knowledge may not yet have sufficient knowledge or skills to leverage systemic structures, especially when numbers are not available to perform calculations on complex problems. In contrast, students with higher algebraic knowledge may have sufficient knowledge or skills to leverage systemic structures across problems, leading to higher strategy efficiency when solving complex variable vs numerical problems. This finding extends prior work on the selective benefits of variables in learning of mathematical concepts (Lawson et al., 2019; Stigler et al., 2010); it suggests that presenting problems in variables can improve students' strategy efficiency in simplifying complex problems, albeit that this benefit may be limited to students who already have some foundational knowledge of algebra. This hypothesis warrants further investigation, but the current finding provides a starting point for future research on ways to leverage numbers and variables in learning and instruction to support students across knowledge levels.

4.3. Other non-systemic factors influence students' strategy efficiency

In addition to symbols, the four problem pairs varied in the proximity of the terms to be combined as well as the order in which the numerical and variable problems within each pair were presented. Although proximity and order were counterbalanced between pairs, they might still impact students' strategy efficiency. Therefore, in our follow-up exploratory analyses, we statistically accounted for these potential influences. We found that, aligning with prior research (Lee, Hornburg, Chan, & Ottmar, 2022; Sidney, 2020) and supporting theories of perceptual learning (Goldstone et al., 2017; Kellman et al., 2010), students tended to complete problems more efficiently when the terms to be combined were adjacent to each other and on the second of the paired problems. These two factors accounted for 20 % or more of the variance in students' strategy efficiency, six to ten times larger than the effects of symbol and algebraic knowledge, which together only accounted for 2 or 3 % of the variance. We also tested whether the pattern of the results on simple and complex problems were significantly different by combining all problems in one model and including problem complexity as an additional fixed effect. Results revealed significant interactions between symbol, students' algebraic knowledge, and problem complexity, suggesting that the effects of symbol, algebraic knowledge, and their interaction differed between simple and complex problems.

Together, these exploratory analyses support our main findings on the important, albeit small, effects of symbols. Students' strategy efficiency may be influenced by more than just whether the problem is presented in variables or numbers; factors such as students' knowledge level, the complexity of the problems, the order in which the problems were presented, and the proximity of the terms to be combined may also influence students' strategy efficiency. With more problem pairs and more participants, future research should systematically examine whether and how these factors interact to influence students' strategy efficiency. Understanding the ways in which these factors operate in tandem may inform the design of instructional materials that help students develop efficient and flexible problem-solving skills.

4.4. Limitations and future directions

The current study has several limitations. First, the analytic sample was a subset of the students participating in a larger randomized controlled trial. Although 911 students started the game-based intervention, only 445 students (49 %) completed all four problem pairs within the game designed to address our research questions on mathematical symbols. This was likely because the four pairs of problems were embedded in the middle of the game (Problems 78 to 119 out of 252 problems), students solved problems at their own pace, and the study was conducted during the 2020–2021 school year amidst many disruptions to schooling due to COVID-19 outbreaks. As one would suspect, the 445 students who completed these four pairs of problems embedded in the game (i.e., progressed to Problem 119) had a descriptively higher average algebraic knowledge score ($M = 5.76$) compared to the larger sample of 911 students ($M = 4.78$) who started the game but did not complete enough of the intervention to reach these problems. Although the analytic sample was relatively high performing compared to the larger sample, students' algebraic knowledge scores in the current analytic sample were not subject to the ceiling effect. Further, students' algebraic knowledge scores were widely and normally distributed. While the analytic sample might not be representative of the students in the larger study, school district, state, or the country, it nonetheless allowed us to examine the influence of students' algebraic knowledge as well as its interaction with symbols on students' strategy efficiency.

Second, the study was conducted in the context of a larger randomized controlled trial. Given the primary aims and the constraints of the larger study, we only included two pairs of simple problems and two pairs of complex problems. Although the number of problems included in the analyses were small, we still observed a main effect of algebraic knowledge and a Symbol × Algebraic Knowledge interaction on students' strategy efficiency. Further, the larger study was conducted in an authentic classroom learning environment rather than a laboratory and we used the log files of students' moment-to-moment actions during problem solving in a mathematical game, providing some ecological validity to the current findings.

Third, the current study did not allow us to disentangle students' conceptualizations of variables and why problems presented in variables were challenging, especially for students with lower algebraic knowledge. Specifically, the variables in our problems can represent generalized arithmetic relations between terms, be interpreted as unknowns, and be substituted with varying or multiple values. Difficulty grasping one or more of these abstract concepts likely contributed to students' struggles with the problems presented in variables. More research with different types of variable problems that specifically target each concept is needed to identify students' core challenges with variables and whether these challenges vary by students' algebraic knowledge.

Future studies should include a more diverse sample as well as more problems systematically targeting different concepts of variable and varying in symbols, order, proximity, complexity, systemic structures, and mathematical properties. Doing so will extend the findings beyond inverse operations and allow further investigation of how different factors may interact with symbols to influence students' problem-solving strategies. These future directions will also contribute new findings on when and for whom variables support efficient problem-solving strategies.

4.5. Implications for teaching and learning

Despite the limitations, the current findings have implications for instructional practices that guide students' attention to the systemic structures of expressions in classroom and online learning contexts. One practical implication is that students, especially those with lower algebraic knowledge, may benefit from explicit instruction or targeted practice on identifying systemic structures before solving complex variable problems. For example, providing students with ample opportunities to solve both variable and numerical problems of identical structure may help them notice the systemic structure. Explicit instruction on identifying the identical structure between variable and numerical problems may also help students improve their attention to systemic structures, comfort with variables, and strategy efficiency. This implication aligns with prior work showing that drawing students' attention to the analogous systemic structure across problems (e.g., fraction vs whole number division) can improve their mathematical reasoning (Sidney & Alibali, 2017). Guiding students' attention to the systemic structures through instruction and practice with variable problems that prevent impulsive calculation may potentially help strengthen flexibility and efficiency in problem solving among students with lower prior knowledge. In summary, presenting problems in variables instead of numbers as well as mapping relations between problems presented in variables vs numbers may be effective at helping students inhibit their impulse to calculate and attend to the systemic structures of the expressions.

4.6. Conclusion

Overall, this study explores the ways that variations in the symbols used in mathematical expressions influence students' problem-solving strategies. Students' strategy efficiency is influenced by their algebraic knowledge and its interaction with the symbols (i.e., numbers and variables) in which the problems are presented. Other factors, such as the proximity of inverses/opposites in problems, overall problem complexity, and problem order within the game, also influence students' strategy efficiency and the use of systemic structure. If these factors impact whether students notice systemic structures and subsequently their problem-solving strategies, researchers, educators, and content developers should consider these factors when designing studies, instruction, learning materials, and assessments. Further, providing students with different variations of problem structures using a combination of symbols may provide students with greater opportunities to notice systemic structures and improve their mathematical understanding through problem-solving practice.

CRediT authorship contribution statement

Jenny Yun-Chen Chan: Conceptualization, Methodology, Formal analysis, Visualization, Project administration, Writing - original draft, Writing - review & editing. **Erin Ottmar:** Conceptualization, Resources, Funding acquisition, Writing - original draft, Writing - review & editing. **Hannah Smith:** Methodology, Writing - original draft, Writing - review & editing. **Avery H. Closser:** Writing - original draft, Writing - review & editing, Visualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through an Efficacy and Replication Grant (R305A180401) to Worcester Polytechnic Institute. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. We thank the participating teachers and students for their support with this study. We also thank Ji-Eun Lee, David Brokaw, Yveder Joseph, Kathryn Drzewiecki, and members of the Math Abstraction Play Learning and Embodiment Lab for their work on this project.

Appendix A. Analyses by demographic variables

See Tables A1 and A2.

Table A1
Mean and standard deviation of algebraic knowledge scores and total step count across the eight problems by course level, biological sex, and race/ethnicity.

	n	Algebraic Knowledge Score	Total Steps
Course level			
On-Level	281	4.40 (2.11)	56.49 (21.62)
Advanced	164	8.10 (1.68)	49.79 (19.69)
Biological sex			
Male	239	5.87 (2.71)	53.47 (20.99)
Female	206	5.64 (2.59)	54.65 (21.38)
Race/ethnicity			
White	228	4.75 (2.39)	52.94 (18.34)
Asian	133	7.97 (1.78)	51.61 (20.44)
Hispanic	55	4.60 (2.20)	64.75 (29.93)
Black	12	4.42 (2.23)	52.58 (11.71)
Other	16	6.81 (2.61)	54.81 (25.25)

Table A2
Beta coefficients and standard errors of the coefficients in a linear regression model predicting students' total steps across the eight problems.

Predictors	*B* (SE)
Algebraic knowledge	−0.06 (0.07)
Course level: On-level	0.25 (0.13)
Biological sex: Female	0.05 (0.09)
Race/ethnicity: White	−0.19 (0.26)
Race/ethnicity: Asian	−0.06 (0.26)
Race/ethnicity: Hispanic	0.37 (0.28)
Race/ethnicity: Black	−0.23 (0.38)
R^2	0.06

Note. Course level: Advanced was the reference group; Biological sex: Male was the reference group; Race/ethnicity: Other was the reference group.

Appendix B. Sankey diagrams

To improve the legibility of the diagrams, the diagrams only illustrate up to the first 10 steps of students' problem-solving strategies.

For full images, please visit the project site on OSF at https://osf.io/tpnu2/ (see Figs. B1 to B8).

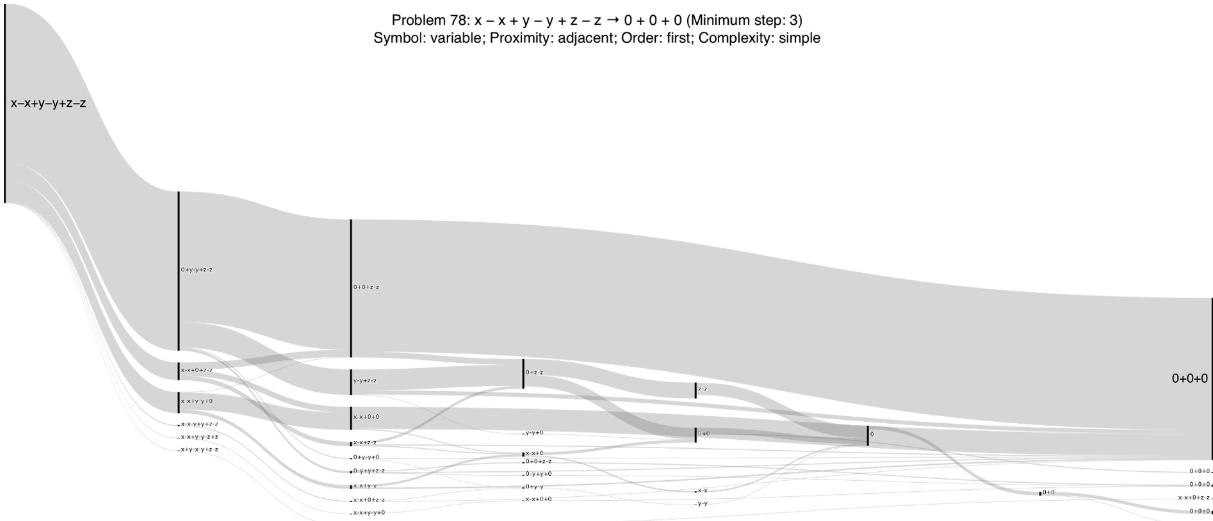

Fig. B1. The Sankey diagram for Problem 78.

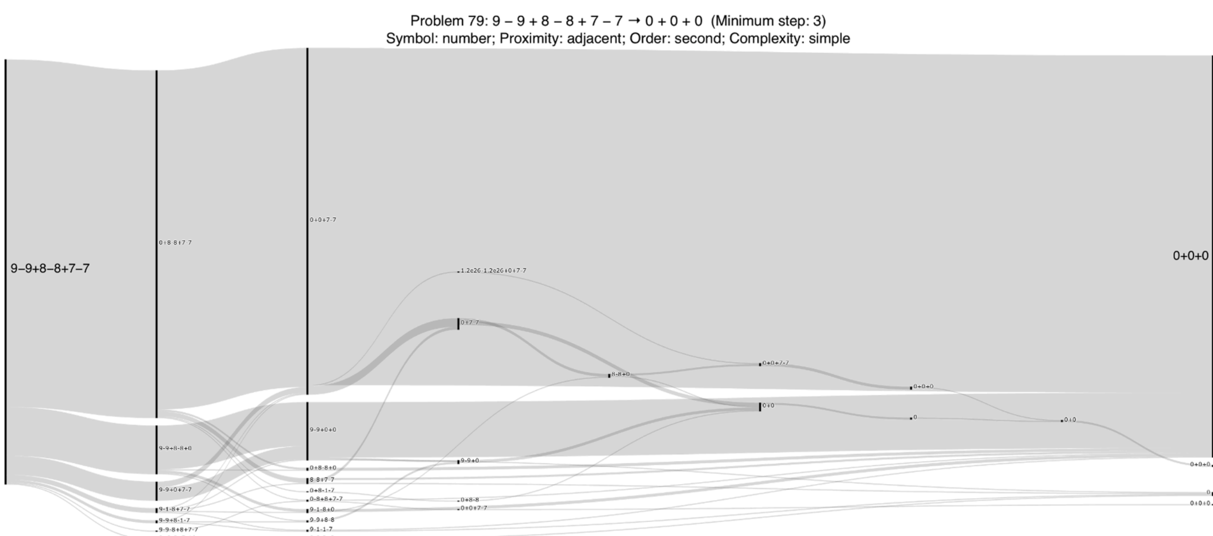

Fig. B2. The Sankey diagram for Problem 79.

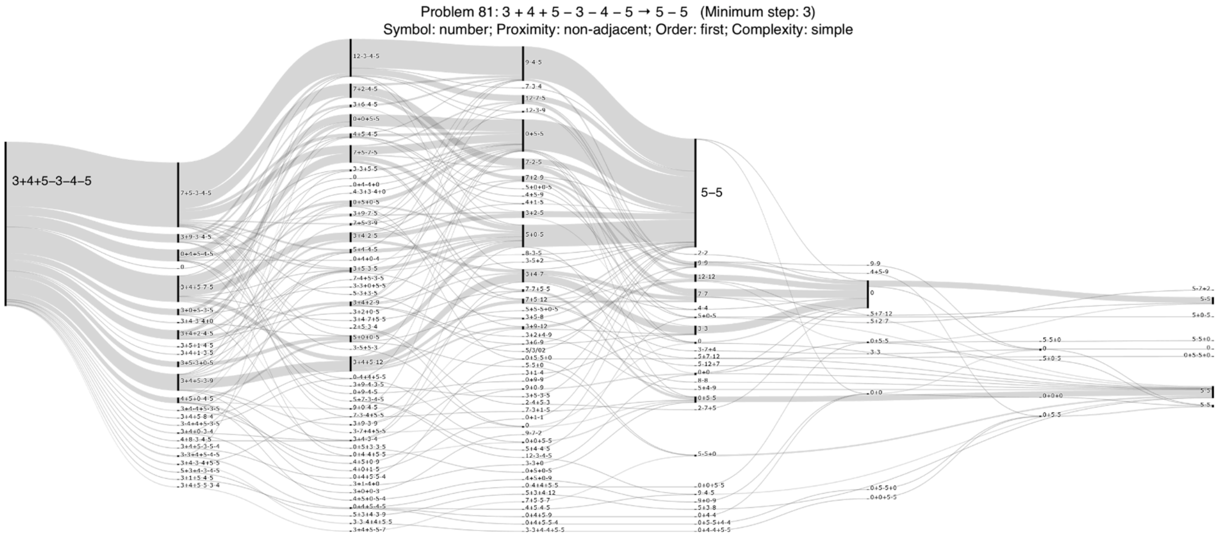

Fig. B3. The Sankey diagram for Problem 81.

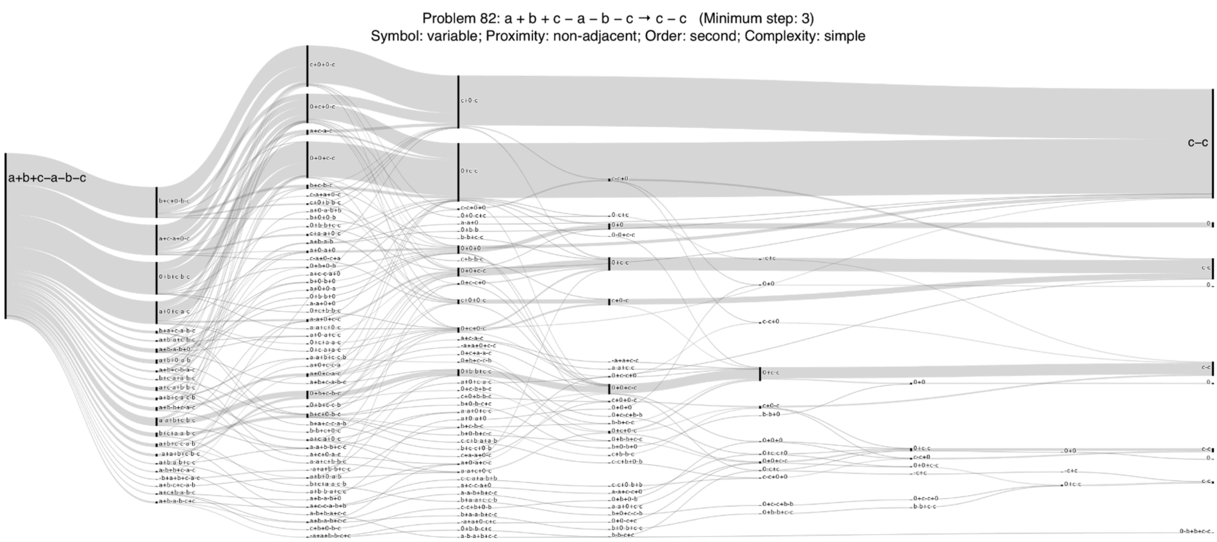

Fig. B4. The Sankey diagram for Problem 82.

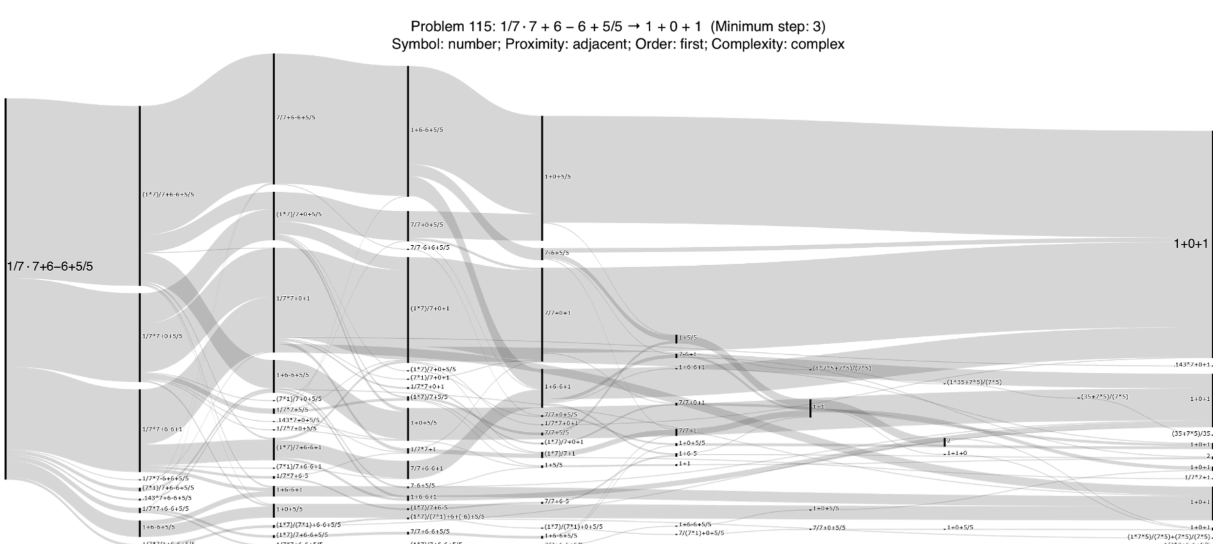

Fig. B5. The Sankey diagram for Problem 115.

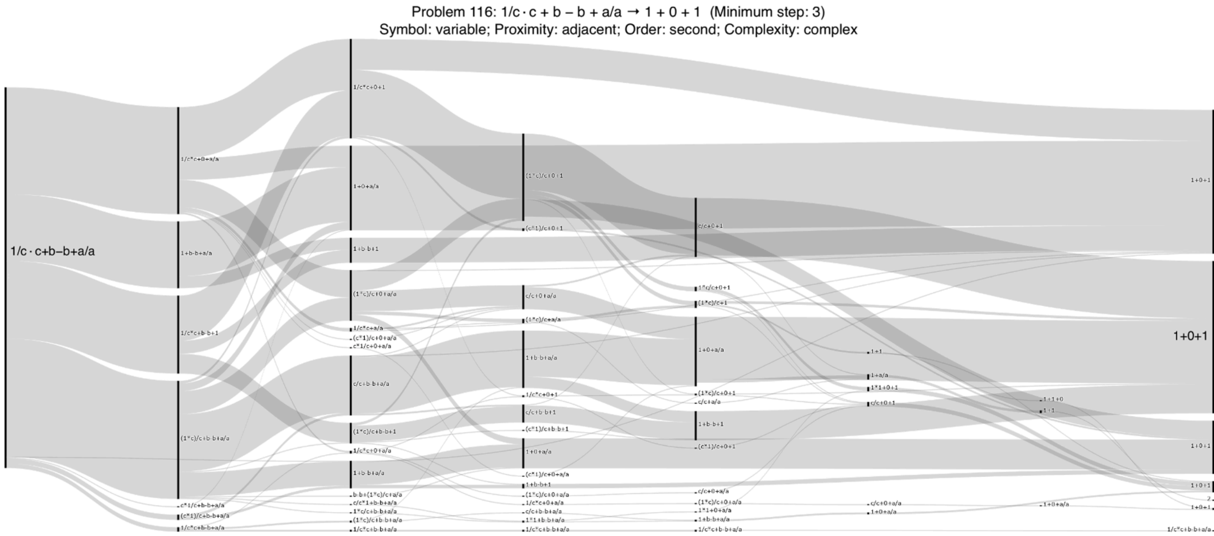

Fig. B6. The Sankey diagram for Problem 116.

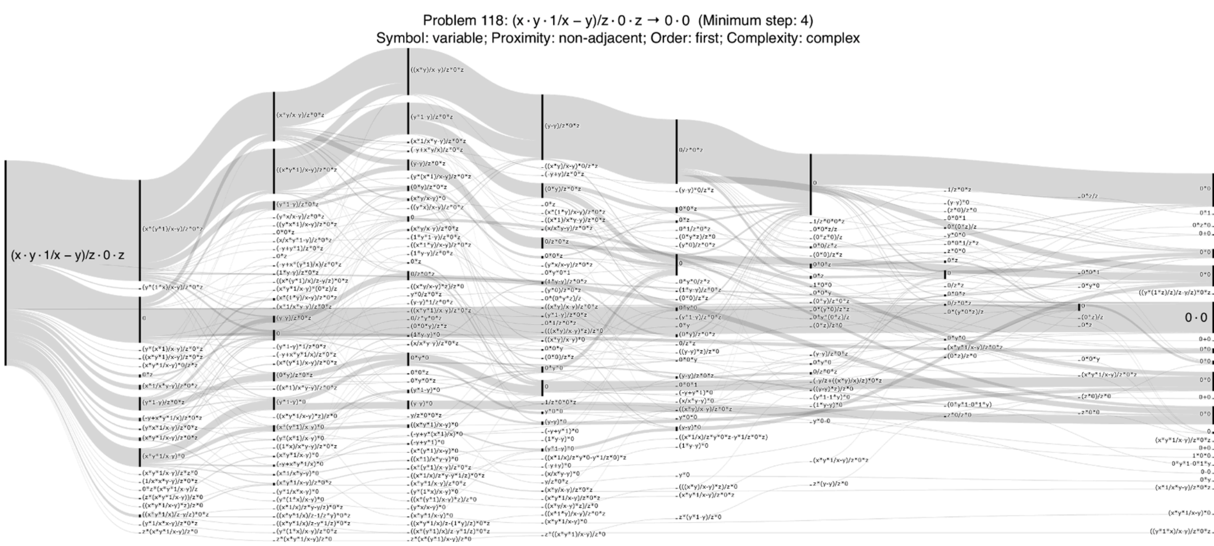

Fig. B7. The Sankey diagram for Problem 118.

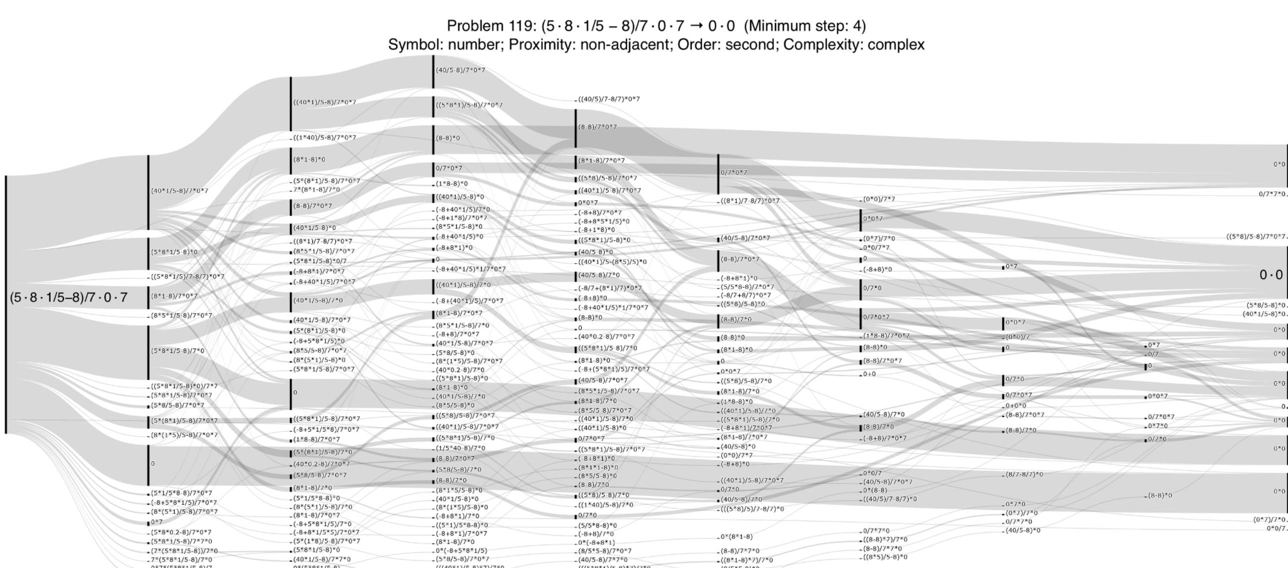

Fig. B8. The sankey diagram for Problem 119.

References

Bardini, C., Radford, L., & Sabena, C. (2005). *Struggling with variables, parameters, and indeterminate objects or how to go insane in mathematics* (pp. 129–136). Melbourne, Australia: PME.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., … Fox, J. (2015). Package "lme4". *Convergence, 12*(1), 2.

Blando, J. A., Kelly, A. E., Schneider, B. R., & Sleeman, D. (1989). Analyzing and modeling arithmetic errors. *Journal for Research in Mathematics Education, 20*(3), 301–308.

Blanton, M. L., & Kaput, J. J. (2005). Characterizing a classroom practice that promotes algebraic reasoning. *Journal for Research in Mathematics Education, 36*(5), 412–446.

Booth, L. R. (1981). Child-methods in secondary mathematics. Retrieved from *Educational Studies in Mathematics, 12*(1), 29–41 https://www.jstor.org/stable/3482379.

Booth, L. R. (1984). *Algebra: Children's strategies and errors*. Windsor, UK: NFER.

Bye, J, Lee, J. E., Chan, J. Y. C., Closser, A. H., Shaw, S., Ottmar, E. (2022, April) *Perceiving precedence: Order of operations errors are predicted by perception of equivalent expressions* [Poster]. The 2022 Annual meeting of the American Educational Research Association (AERA), San Diego, CA.

Carpenter, T. P., Levi, L., Franke, M. L., & Zeringue, J. K. (2005). Algebra in elementary school: Developing relational thinking. *ZDM - International Journal on Mathematics Education, 37*(1), 53–59. https://doi.org/10.1007/BF02655897

Chan, J. Y.-C., Lee, J.-E., Mason, C. A., Sawrey, K. B., & Ottmar, E. R. (2022). From Here to There!: A dynamic algebraic notation system improves understanding of equivalence in middle-school students. *Journal of Educational Psychology, 114*(1), 56–71. https://doi.org/10.1037/edu0000596

Chan, J. Y.-C., Ottmar, E., & Lee, J. E. (2022). Slow down to speed up: Longer pause time before solving problems relates to higher strategy efficiency. *Learning and Individual Differences*. https://doi.org/10.1016/j.lindif.2021.102109

Chan, J. Y.-C., Smith, H., Closser, A. H., Drzewiecki, K. C., & Ottmar, E. (2021). Number vs. variable: The effect of symbols on students' math problem solving. In T. Fitch, C. Lamm, H. Leder, & K. Teβmar-Raible (Eds.), *Proceedings of the Forty-Third Annual Meeting of the Cognitive Science Society* (pp. 2836–2842). Vienna, Austria: University of Vienna.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence*. NJ: Erlbaum.

Decker-Woodrow, L., Mason, C. A., Lee, J.-E., Chan, J. Y.-C., Sales, A., Liu, A., & Tu, S. (forthcoming). The impacts of three educational technologies on algebraic understanding in the context of COVID-19. *AERA Open*.

DeLoache, J. S. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences, 8*(2), 66–70. https://doi.org/10.1016/j.tics.2003.12.004

Filloy, E., & Rojano, T. (1989). Solving equations: The transition from arithmetic to algebra for the learning of mathematics. *For the Learning of Mathematics, 9*(2), 19–25.

Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review, 26*(1), 9–25. https://doi.org/10.1007/s10648-014-9249-3

Garofalo, J. (1989). Beliefs and their influence on mathematical performance. *The Mathematics Teacher, 82*(7), 502–505.

Givvin, K. B., Moroz, V., Loftus, W., & Stigler, J. W. (2019). Removing opportunities to calculate improves students' performance on subsequent word problems. *Cognitive Research: Principles and Implications, 4*(1), 1–13.

Goldstone, R. L., Marghetis, T., Weitnauer, E., Ottmar, E. R., & Landy, D. (2017). Adapting perception, action, and technology for mathematical reasoning. *Current Directions in Psychological Science, 26*(5), 434–441. https://doi.org/10.1177/0963721417704888

Gunnarsson, R., Sönnerhed, W. W., & Hernell, B. (2016). Does it help to use mathematically superfluous brackets when teaching the rules for the order of operations? *Educational Studies in Mathematics, 92*(1), 91–105.

Herscovics, N., & Linchevski, L. (1994). A cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics, 27*(1), 59–78.

Kaput, J. J. (1998). Representations, inscriptions, descriptions and learning: A kaleidoscope of windows. *The Journal of Mathematical Behavior, 17*(2), 265–281. https://doi.org/10.1016/s0364-0213(99)80062-7

Kellman, P. J., Massey, C. M., & Son, J. Y. (2010). Perceptual learning modules in mathematics: Enhancing students' pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science, 2*(2), 285–305. https://doi.org/10.1111/j.1756-8765.2009.01053.x

Kieran, C. (1989). The early learning of algebra: A structural perspective. In S. Wagner, & C. Kieran (Eds.), *Research Issues in the Learning and Teaching of Algebra: The Research Agenda for Mathematics Education* (pp. 33–56). Virginia, U.S.A: NCTM and Lawrence Erlbaum Associates. https://doi.org/10.4324/9781315044378-4.

Koedinger, K., Alibali, M., & Nathan, M. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science, 32*, 366–397. https://doi.org/10.1080/03640210701863933

Koedinger, K. R., & Anderson, J. R. (1998). Illustrating principled design: The early evolution of a cognitive tutor for algebra symbolization. *Interactive Learning Environments, 5*(1), 161–179. https://doi.org/10.1080/1049482980050111

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*(3), 151–160.

Küchemann, D. (1981). Algebra. In K. M. Hart (Ed.), *Children's Understanding of Mathematics: 11–16* (pp. 103–119). John Murray.

Lawson, A. P., Mirinjian, A., & Son, J. Y. (2019). Can preventing calculations help students learn math? *Journal of Cognitive Education and Psychology, 17*(2), 178–197. https://doi.org/10.1891/1945-8959.17.2.178

Lee, J.-E., Hornburg, C. B., Chan, J. Y.-C., & Ottmar, E. (2022). Perceptual and number effects on students' initial solution strategies in an interactive online mathematics game. *Journal of Numerical Cognition, 8*(1), 166–182. https://doi.org/10.5964/jnc.8323

Lee, J.-E., Stalin, A., Drzewiecki, K., Trac, C., & Ottmar, E. (in press). Show the flow: Visualization of students' solution strategies with Sankey diagrams in an online mathematics game. *Journal of Interactive Learning Research*.

Malisani, E., & Spagnolo, F. (2009). From arithmetical thought to algebraic thought: The role of the "variable". *Educational Studies in Mathematics, 71*(1), 19–41. https://doi.org/10.1007/s10649-008-9157-x

National Governors Association Center for Best Practices (NGA Center) and Council of Chief State School Officers (CCSSO). (2010). Common core state standards for mathematics. Washington, DC, USA: Author. Retrieved from http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf.

Newton, K. J., Lange, K., & Booth, J. L. (2020). Mathematical flexibility: Aspects of a continuum and the role of prior Knowledge. *The Journal of Experimental Education, 88*(4), 503–515. https://doi.org/10.1080/00220973.2019.1586629

Niehaus, E., Campbell, C. M., & Inkelas, K. K. (2014). HLM behind the curtain: Unveiling decisions behind the use and interpretation of HLM in higher education research. *Research in Higher Education, 55*(1), 101–122. https://doi.org/10.1007/s11162-013-9306-7

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 510–520. https://doi.org/10.1037/0278-7393.14.3.510

Ottmar, E. R., Landy, D., Goldstone, R., & Weitnauer, E. (2015). Getting From Here to There!: Testing the effectiveness of an interactive mathematics intervention embedding perceptual learning. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the Annual Conference of the Cognitive Science Society* (pp. 1793–1798). Cognitive Science Society. https://cogsci.mindmodeling.org/2015/papers/0311/paper0311.pdf.

Philipp, R. A. (1992). The many uses of algebraic variables. *Mathematics Teacher, 85*(7), 557–561.

Radford, L. (2006). In *Algebraic thinking and the generalization of patterns: A semiotic perspective* (pp. 2–21). Mérida, México: Universidad Pedagógica Nacional. https://doi.org/10.1007/1-4020-7910-9_99.

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology, 99*(3), 561–574. https://doi.org/10.1037/0022-0663.99.3.561

Robinson, K. M., Ninowski, J. E., & Gray, M. L. (2006). Children's understanding of the arithmetic concepts of inversion and associativity. *Journal of Experimental Child Psychology, 94*(4), 349–362. https://doi.org/10.1016/j.jecp.2006.03.004

Robinson, K. M., Price, J. A. B., & Demyen, B. (2018). Understanding arithmetic concepts: Does operation matter? *Journal of Experimental Child Psychology*, *166* (Supplement C), 421–436. https://doi.org/https://doi.org/10.1016/j.jecp.2017.09.003.

Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology, 47*(6), 1525–1538. https://doi.org/10.1037/a0024997

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334–370). New York: MacMillan. https://doi.org/10.1177/002205741619600202.

Schoenfeld, A. H., & Arcavi, A. (1988). On the meaning of variable. *The mathematics teacher, 81*(6), 420–427.

Sfard, A., & Linchevski, L. (1994). The gains and the pitfalls of reification - The case of algebra. *Educational Studies in Mathematics, 26*(2–3), 191–228. https://doi.org/10.1007/BF01273663

Sidney, P. G. (2020). Children's learning from implicit analogies during instruction: Evidence from fraction division. *Cognitive Development, 56*, Article 100956.

Sidney, P. G., & Alibali, M. W. (2015). Making connections in math: Activating a prior knowledge analogue matters for learning. *Journal of Cognition and Development, 16*(1), 160–185. https://doi.org/10.1080/15248372.2013.792091

Sidney, P. G., & Alibali, M. W. (2017). Creating a context for learning: Activating children's whole number knowledge prepares them to understand fraction division. *Journal of Numerical Cognition, 3*(1), 31–57. https://doi.org/10.5964/jnc.v3i1.71

Stacey, K., & MacGregor, M. (1999). Learning the algebraic method of solving problems. *The Journal of Mathematical Behavior, 18*(2), 149–167.

Star, J. R., & Newton, K. J. (2009). The nature and development of experts' strategy flexibility for solving equations. *ZDM - International Journal on Mathematics Education, 41*, 557–567. https://doi.org/10.1007/s11858-009-0185-5

Star, J. R., Pollack, C., Durkin, K., Rittle-Johnson, B., Lynch, K., Newton, K., & Gogolen, C. (2014). Learning from comparison in algebra. *Contemporary Educational Psychology, 40*, 41–54. https://doi.org/10.1016/j.cedpsych.2014.05.005

Star, J. R., & Rittle-Johnson, B. (2008). Flexibility in problem solving: The case of equation solving. *Learning and Instruction, 18*(6), 565–579. https://doi.org/10.1016/j.learninstruc.2007.09.018

Stigler, J. W., Givvin, K. B., & Thompson, B. J. (2010). What community college developmental mathematics students understand about mathematics. *MathAMATYC Educator, 1*(3), 4–16.

Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2006). The development of children's adaptive expertise in the number domain 20 to 100. *Cognition and Instruction, 24*(4), 439–465. https://doi.org/10.1207/s1532690xci2404

Usiskin, Z. (1988). Conceptions of school algebra and uses of variables. *The ideas of algebra, K-12, 8, 19*.

Vamvakoussi, X., & Vosniadou, S. (2004). Understanding the structure of the set of rational numbers: A conceptual change approach. *Learning and Instruction, 14*(5), 453–467.

Venkat, H., Askew, M., Watson, A., & Mason, J. (2019). Architecture of mathematical structure. Retrieved from *For the Learning of Mathematics, 39*(1), 13–17 https://flm-journal.org/Articles/15CDA7D47D380050F5587BF8C2BCB8.pdf.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Walfish, S. (2006). A review of statistical outlier methods. *Pharmaceutical Technology, 30* (11), 82–86.

Wang, J., Liu, R. D., Star, J., Liu, Y., & Zhen, R. (2019). The moderating effect of regulatory focus in the relationship between potential flexibility and practical flexibility. *Contemporary Educational Psychology, 56*(January), 218–227. https://doi.org/10.1016/j.cedpsych.2019.01.013

Warren, E. (2003). The role of arithmetic structure in the transition from arithmetic to algebra. *Mathematics Education Research Journal, 15*(2), 122–137. https://doi.org/10.1007/BF03217374

Xu, L., Liu, R. D., Star, J. R., Wang, J., Liu, Y., & Zhen, R. (2017). Measures of potential flexibility and practical flexibility in equation solving. *Frontiers in Psychology, 8* (AUG), 1368. https://doi.org/10.3389/fpsyg.2017.01368