Title
Teaching the concept of a variable with meaning and purpose: Connecting contextual mathematical thought to the abstract symbols and operations of Algebra

Author
Bye, Jeffrey Kramer

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Teaching the concept of a variable with meaning and purpose:

Connecting contextual mathematical thought

to the abstract symbols and operations of Algebra

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Psychology

by

Jeffrey Kramer Bye

2016

ABSTRACT OF THE DISSERTATION

Teaching the concept of a variable with meaning and purpose:

Connecting contextual mathematical thought

to the abstract symbols and operations of Algebra

by

Jeffrey Kramer Bye

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2016

Professor Patricia Cheng, Chair

The concept of an *algebraic variable* is both important in its own right and foundational for higher levels of math, but many students struggle to comprehend its *meaning* and *purpose*, demonstrating a variety of misconceptions about the interpretation of a variable and algebra's relation to arithmetic. Common educational practices fail to support a substantial portion of students in connecting their *intuitive* cognitive capabilities to the formal external representations (i.e., symbolic notation) of algebra, depriving these students of understanding how and why variables are used, as well as their relevance in solving real-world problems. Previous attempts at improving students' understanding of variables have focused on schematic induction across varied concrete examples or the generalization of relational thinking from arithmetic. While these efforts are important, the approaches do not fully elucidate the purpose of using formal symbols (e.g., letters) to represent unknown numbers. I posit that the clearest way to demonstrate the purpose of symbolic variables is through students' formulation and attempted solution of mathematical problems where multiple unknowns must be represented (and distinguished from each other), such as in a system of equations word problem. Guided by principles from cognitive psychology and educational research, I formulate a framework for encouraging and supporting students' intuitive discovery of the concept of variable using purpose-driven contrast comparisons, active learning techniques such as constructive struggling with intuitive hints, and contextual facilitation of students' natural problem solving for meaningful, concrete tasks. Through this process, variables representations are introduced progressively, first by using more interpretable word equations and later by abbreviating word phrases into letter symbols. I implemented this framework into novel multimedia educational materials, which were iteratively piloted and revised, and then experimentally tested with middle and high school students against a more traditionally structured control version of the materials and a baseline condition. The results from this experimental testing suggest that students who were encouraged to infer the purpose of a variable before its formal representation was introduced went on to provide more correct answers to analogous problems on a post-test given 1-3 weeks later.

The dissertation of Jeffrey Kramer Bye is approved.

Keith Holyoak

James W. Stigler

Megan Loef Franke

Patricia Cheng, Committee Chair

University of California, Los Angeles

2016

DEDICATION

This dissertation is dedicated to my parents, Linda and Roland Bye, for all of the sacrifices they have made for me and the love and support they have given me for the past three decades. It is also dedicated to all of the great teachers I have had throughout my education, especially my mother, who taught me to love math and inspired me to try to make a difference in math education; my undergraduate advisors at Pomona College, Dr. Jay D. Atlas and Dr. Robert Thornton, who introduced me to cognitive science, taught me how to think critically, and encouraged me to seek a doctorate; the UCLA Reasoning Lab and my dissertation committee, who provided excellent feedback during this project; and my advisor, Dr. Patricia Cheng, who gave countless hours of advice, insight, and encouragement throughout my graduate career, especially on this research. Finally, this dissertation is dedicated to Annie Collins, the most supportive and inspiring partner imaginable.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS USED

AL: Active Learning

CC: Contrast Comparisons

DPT: Delayed Post-Test

FF: Formalisms First

G-PF: Generalization-Based Progressive Formalization

GC: Generic Context

IPT: Immediate Post-Test

P-PF: Purpose-Driven Progressive Formalization

PF: Progressive Formalization

PL: Passive Learning

PRM: Plausibility, Relatability, & Motivation

ACKNOWLEDGMENTS

The illustrations used in the multimedia materials depicting the characters, restaurants, and lunch table were drawn specifically for this project by artist Airom Bleicher. Students' delayed post-tests were analyzed by my research assistants, Andy Sanchez and Priyanka Mehta. And of course, none of this would be possible without the help of several math teachers, afterschool program coordinators, school administrators, parents, and most of all, students willing to learn. Thanks are due to all of these people for the critical roles they played in this project.

VITA

Jeffrey Kramer Bye, M.A.

EDUCATION

University of California, Los Angeles

M.A., Cognitive Psychology; December 2011

Specialization in Computational Cognition

Advisor: Dr. Patricia Cheng

Pomona College

B.A., Cognitive Science, *cum laude*; May 2009

Subconcentration in Computer Science & Minor in Philosophy

Advisors: Dr. Jay D. Atlas & Dr. Robert Thornton

PUBLICATIONS

- Collins, A. L., Greenfield, V. Y., **Bye, J. K.**, Linker, K. E., Wang, A. S., & Wassum, K. M. (2016). Dynamic mesolimbic dopamine signaling during action sequence learning and expectation violation. *Scientific Reports, 6*, 20231.
- **Bye, J. K.**, Nguyen, B. D., Lu, H., & Johnson, S. P. (2014). Anticipating an effect from predictive visual sequences: Development of infants' causal inference from 9 to 18 months. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1976-1981). Austin, TX: Cognitive Science Society.

CONFERENCE PRESENTATIONS

- **Bye, J. K.**, Nguyen, B. D., Lu, H., & Johnson, S. P. (2015, March). Infants' learning of causal shape sequences in a predictive gaze shift paradigm. Poster presented at the biennial meeting of the Society for Research in Child Development, Philadelphia, PA.
- **Bye, J. K.**, Nguyen, B. D., Lu, H., & Johnson, S. P. (2014, July). Anticipating an effect from predictive visual sequences: Development of infants' causal inference from 9 to 18 months. Poster presented at the 36th Annual Conference of the Cognitive Science Society. Quebec City, Canada.
- Thornton, R., & **Bye, J.** (2008, September). Syntactic category ambiguity in sentence production. Poster presented at the Fourteenth Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP), Cambridge, UK.
- Thornton, R., & **Bye, J.** (2008, March). Syntactic category ambiguity and grammatical agreement. Poster presented at the Twenty-First Annual CUNY Conference on Human Sentence Processing, Chapel Hill, NC.

AWARDS

Dissertation Year Fellowship, UCLA; 2015-2016

UCLA Distinguished Teaching Award; 2015

Shepherd Ivory Franz Distinguished Teaching Assistant Award, UCLA Psychology, 2014

Edwin W. Pauley Fellowship, UCLA; 2013-2014

Graduate Summer Research Mentorship, UCLA; 2011, 2012

Distinguished University Fellowship, UCLA; 2010-2011

John Purvis Prize for Best Cognitive Science Thesis, Pomona College; May 2009

1. Introduction

1.1 Overview

The concept of an *algebraic variable* is both important in its own right and foundational for higher levels of mathematics, but many students struggle to understand the *meaning* and *purpose* of a variable. In this dissertation, I argue that common math education practices fail to support a substantial portion of students in making important insights about the interpretation of a variable, its mathematical purpose, and its relevance in solving real-world problems. The overwhelming emphasis on formal procedures and notation in math education allows for some of these conceptually impoverished students to still progress through math class for some time, so long as they can procedurally manipulate variable symbols in rote practice. However, without a conceptual foundation, the rules of manipulation seem disconnected from arithmetic; and without a sense of purpose, the exercise of solving for a variable seems arbitrary and irrelevant. Such conceptual impoverishment prevents these students from appreciating algebra and from building on its concepts in more advanced math.

Math educators often overestimate students' ability to interpret symbolic notation in the manner intended; they assume that the meaning of the abstract, formal symbols of math notation is unambiguous and readily accessible to their students (Nathan, Koedinger, & Alibali, 2001), and that the purpose for using such symbols is self-evident. Algebra textbooks and classes tend to begin by defining variables *formally* (as letters that represent numbers) and proceeding directly to rote procedural practice (syntactic rules of manipulating variables), without allowing or encouraging students to first gain an intuitive sense of the meaning or purpose of using this new symbolic representation. This problem is compounded by the fact that these early examples are universally cases of using a variable symbol to represent a single unknown in problems so mathematically simple that they are readily solvable in purely arithmetic form (e.g., the algebraic '$2 + x = 6$' is operationally equivalent to the arithmetic '$6 - 2 =$ '). Because no motivation is given for the algebraic form's representational advantage over the arithmetic, students are understandably confused about the meaning and purpose of this new formalism.

Algebra students demonstrate a variety of misconceptions about the interpretation of a variable and its relation to arithmetic (Booth, 1988; Filloy & Rojano, 1989; Herscovics & Linchevski, 1994; Knuth et al., 2005, 2006; Koedinger & Nathan, 2004; Lee & Wheeler, 1989). Davis (1975) provides an anecdote of Henry (a high-performing 12-year-old math student), who asked "How can we multiply by x when we don't know what x is?" This case study provides insight into the conceptual difficulties faced by students learning variables. These issues are particularly problematic given the fundamental role of algebraic variables in higher forms of math (e.g., calculus, statistics).

In order to improve struggling students' learning and utilization of algebra, we must identify and remedy pedagogical practices that contribute to these misconceptions. In this project, I have created and assessed new educational materials that encourage and support students' discovery of the meaning and purpose of a variable, guided by principles from educational psychology and the cognitive psychology of learning, and iteratively tested with pre-Algebra students and improved. The techniques implemented in these materials all stem naturally from an underlying focus on developing students' conceptual understanding of algebraic variables, in particular, their meaning and purpose.

In the rest of this chapter, I describe my guiding principle, motivation, and outline the key components of my approach. In Chapter 2, I summarize relevant conceptual roadblocks previously identified in algebra education and how they inform my approach. In Chapter 3, I describe the structure and content of the written version of my materials, the design principles that guided their development, and, when informative, how the materials were improved through iterative piloting and testing. In Chapter 4, I detail how the materials were adapted into interactive multimedia videos and the improvements to their structure and content. In Chapter 5, I summarize the experimental testing of the materials, compare the results against both a control and baseline condition, discuss the conclusions that can be drawn from the experimental results thus far, and outline a plan for future improvement and assessment of these materials.

1.2 Guiding Principle

Many students who struggle with school Algebra are nonetheless capable of *algebraic thinking*, by which I mean that they can mentally represent an unknown value and even manipulate it if the mathematical structure it is embedded in is simple enough to intuitively proceduralize. For example, if asked "What number must be added to 2 to total 6?", reasoners can not only cue the procedure of subtracting 2 from 6 to find 4 (or even guessing and checking potential addends), but they may also think about the unknown with some relatively simple, explicit mental term such as 'some number' (though not yet necessarily as a formal symbol like 'x'). These students are also capable of reasoning about general mathematical relations (another explicit feature of algebraic representation), such as the zero property of addition (any number added to 0 equals itself), even if such a relation is not expressed symbolically as formal algebraic notation (e.g., '$x + 0 = x$'; Carpenter, Franke, & Levi, 2003).

Students' struggles with formal algebra problems indicate a failure of relating their intuitive cognitive capabilities to the formal external representations (i.e., symbolic notation) of algebra, thereby depriving these students of the fully realized concept of a variable and the computational power of formal algebraic representations for problems that are more complex or abstract. This failure is largely attributable to how variables are introduced and taught in Algebra classes and textbooks, in a manner that fails to build on top of students' intuitive algebraic thinking or demonstrate the purpose of the new formal representation.

What confuses many students is that the formal symbols of algebra are semantically opaque, and the operations upon them are seemingly arbitrary and disconnected from the previously learned operations of arithmetic (Herscovics & Linchevski, 1994). Because many educators fail to recognize the notable cognitive demands of learning algebra, Herscovics and Linchevski argue that:

> "…many students do not have the time to construct a good *intuitive basis* for the ideas of algebra or to connect these with the pre-algebraic ideas they have developed in primary school; they fail to construct *meaning* for the new symbolism and are reduced to performing *meaningless operations on symbols they do not understand*" (Herscovics & Linchevski, 1994, p. 60, emphasis mine)

This quote hits the crux of students' problems with algebra: they can memorize operational rules, but neither the rules nor the symbols themselves have purpose or meaning. I would further add that students don't just lack time, but they lack an appropriate pedagogy to encourage the formation of their own intuitive basis.

These problems can be addressed by developing educational materials that engage students' natural problem solving skills, encourage them to contrast increasingly complex problems to intuit the purpose of new symbolic representation before it's introduced, and challenge them to reflect on the meaning and purpose of different representational forms, as well as their limitations.

1.3 Motivation

While setting such goals for educational materials may seem obvious or even unoriginal, formal reviews of textbooks and teacher beliefs have indicated that they have never been implemented (Nathan & Koedinger, 2000a, 2000b; Nathan, Long, & Alibali, 2002; Sherman, Walkington, & Howell, 2016). Moreover, while previous research has argued for some changes (Nathan, 2012; Romberg, 2001), these approaches do not implement many important aspects of this current project. My own personal experience in developing these materials has demonstrated a clear reason why: in addition to being abstract, the concept of a variable is multi-faceted, and so conveying its meaning and purpose requires decomposing the concept, encouraging students to make intuitive insights into each component's purpose, and using relatable contexts to make the concept more meaningful. Accomplishing these goals requires iterative cycles of principled design, experimental testing, and subsequent revision. Romberg's (2001) summary of the *Mathematics in Context* curriculum (discussed below) acknowledges that two teams working together took six years to develop a longer curriculum that includes some similar goals, although their materials appear to fail to focus on purpose.

Broadly, my goal in this project is to develop an educational intervention that helps students to develop an intuitive basis for the *concept* of a variable, its *meaning*, and its *purpose*. I take a cognitive science approach to the task of developing new educational materials. I use the historical development of the concept of algebraic variables as a blueprint for developing in the students an intuitive sense of the multiple needs for creating the concept of a variable and sequencing students' learning so that the concept becomes a solution to those needs. The field of cognitive psychology offers several important insights into this endeavor.

Previous attempts at achieving similar goals have focused primarily on schematic induction of symbolic meaning across varied concrete examples (e.g., Koedinger & Anderson, 1998; Nathan, 2012; Romberg, 2001). Instead, my approach focuses not just on abstracting meaning for variables but also on intuitively understanding their purpose, using contrast comparisons to enable the students to discover the purpose of each elemental component of the concept via their implicit causal reasoning. Each problem is presented as a story with a relatable, concrete context, and students are encouraged to use their intuition to constructively struggle through the problem before more formal representations and strategies are gradually introduced, as necessary, for students to map their intuitions onto. The techniques implemented in this approach work collectively to provide a guiding framework for students' progression through their own individual conceptual development of the concept of a variable, its meaning, and its purpose.

1.4 Approach

My approach to teaching the meaning and purpose of a variable weaves together many strands to serve a common goal. The backbone of the materials is a curated sequence of relatable, contextual math problems that structurally mirror the historical development of the concept of a variable. Students are encouraged throughout to attempt to solve problems on their own, in line with active learning and constructive struggling techniques. By asking students to contrast across different problem structures, comparing ones they know how to solve with ones they do not, students will be able to identify the conceptual roadblocks preventing them from solving new problems with their existing toolkit. Through this process, students may begin to intuit the purpose of a new concept that would allow them to represent unknown quantities across multiple mathematical steps. New representations for this concept are gradually introduced to students as they are needed. Thus, from the beginning of this approach, students' intuitive understanding of the concept of a variable is driven by purpose and grounded in a meaningful context.

The meaning and the purpose of a variable are interrelated, but both notions are distinctly important and worth clarifying. One way of marking their distinction is to say that meaning is concerned with *semantic* notions like interpretation, content, context, and relatability, while purpose is concerned with *teleological* notions like motivation, goal-relevance, and utility.

For a variable to be *meaningful* to a student, it should be clear *how* it is being used (what the letter-symbol *signifies*): not only the variable's referent (e.g., a specific unknown number in a corresponding word problem), but also its characteristics (e.g., 'x' in '$x + 4$' is a general quantity, but 'x' in '$x + 4 = 6$' is a specific unknown). Additionally, in the case of a word or story problem, the use of a variable should ideally be personally meaningful, representing a context that is *relatable* to the student.

For a variable to be *purposeful* to a student, it should be clear *why* it is being used (what the concept *enables*): how a variable is an essential concept to the goal of solving a given math problem (which we might call 'proximal academic purpose'), as well as related math problems in the future ('distal academic purpose'), and ideally, problems likely to arise in student's real-world experience ('real-world purpose'). Moreover, we can distinguish between aspects of purpose that are *conceptual* in nature (e.g., the purpose of having the concept of a variable in our cognitive toolkit) and aspects of purpose that are *practical* in nature (e.g., the practical considerations for why we use a letter like '*x*' to represent a variable).

1.4.1 "Algebrafication" Through Concrete Problem Solving Activities

Multiple researchers have divvied up the landscape of algebraic activity in different ways, such as Bell's (1995) three core modes (generalizing, forming and solving equations, and working with functions and formulae) and Kaput's (1998, 1999) five interrelated forms (generalization and formalization of patterns and constraints, syntactic manipulation of formalisms, the study of structures and systems, the study of functions and relations, and modeling). Previous research has implemented the 'algebrafication' of elementary school math instruction (i.e., significantly prior to traditional Algebra coursework), with the overwhelming emphasis on generalization of arithmetic and the study of relations (Carpenter, Franke, & Levi, 2003; Jacobs, Franke, Carpenter, Levi, & Battey, 2007); this focus is congruent with popular definitions of algebra as 'generalized arithmetic' (Herscovics & Linchevski, 1994; Koedinger & Anderson, 1998).

These previous efforts have focused on building atop students' arithmetic experiences to generalize mathematical knowledge in an algebraic way. In one line of research, elementary school students have been taught to use relational algebraic reasoning to efficiently deduce the value of an unknown (symbolized as a box: □) without carrying out the full arithmetic implied by the problem. For example, '37 + 54 = □ + 55' can be solved by reasoning that since 55 is 1 more than 54, the unknown must be 1 less than 37, in order to preserve the equality relation, without actually computing the sum of 37 and 54 (Jacobs et al., 2007). In another line of work, as students learn about arithmetic relations, they have been prompted to express general numerical rules—an algebraic way of thinking that enhances their mathematical learning. For example, Carpenter and Levi (2000) asked students to verify the truth of equations like '58 + 0 = 58' in order to prompt them to conjecture general mathematical rules (in this case, the zero property of addition).

While this work is important and effective in promoting relational thinking and teaching mathematical properties, these activities are not sufficient for fully demonstrating the concept of a variable and its purpose (nor are they intended to be). Box symbols are useful symbols for some of these tasks, but an important reason for using letters to more formally represent unknowns is that the alphabet offers a wide array of familiar symbols that are readily distinguished, and the same letter can be used multiple times to indicate a sameness of value within a problem, while other letters can stand for potentially different values. These representational advantages are much clearer when multiple unknowns must be represented. Additionally, mathematical generalities such as the zero property of addition can be implicitly understood without recourse to formalisms and may be expressed less formally (e.g., "any number added to zero equals that number"), as first- and second-grade students are able to do (Carpenter & Levi, 2000). There is considerably less representational advantage to formalizing general relations with letter variables, relative to verbal descriptions.

Rather than formalizing variables through relational thinking, I posit that the clearest way to demonstrate the purpose of symbolic variables is through students' formulation and solution of concrete mathematical problems where multiple unknowns must be represented (and distinguished from each other). My approach thus builds on students' natural problem-solving ability and prior understanding of arithmetic procedures to demonstrate the need for a new concept: the algebraic variable. That this same concept can be used to formalize relational thinking via the same notation (e.g., '$x + 0 = x$' for the zero property of addition) allows for subsequent generalization to be formally encapsulated by symbols that have first been given meaning and purpose. Importantly, my approach is complementary to the research on "algebrafying" students' relational thinking, in that both approaches may be implemented in the same curriculum. Because concrete problem solving is an intuitive task and the purpose of the formal notation for variables is clearest with multiple unknowns, my materials encourage students to first ground the new symbols with meaning and purpose before then generalizing them to more abstract uses.

1.4.2 Historical Basis of the Problem Sequence

The history of mathematics demonstrates repeatedly that concepts and their representations were developed as a practical need for them arose; even as mathematics became more abstract and theoretical, it did so to serve practical purposes, such as with the unification of algebra and geometry (Klein, 1936/1968). Mathematical innovations are typically driven by a goal or purpose and achieved through gradual reconceptualization and extension of existing concepts and procedures (Kieran, 1992; Sfard, 1995). If the mathematics we teach is of any utility to students (as we claim much of it is, especially basic algebra), then they very well may benefit from following the goal-driven line of thought that drove the innovation in the first place.

Kieran (1992) and Sfard (1995) have both argued that the paradigmatic trajectories of the historical development of mathematical concepts parallel the trajectories of individual students' mathematical learning. Even if students' learning does not necessitate such a parallel, I take the historical purposes for the development of algebraic variables and their modern-day relevance as initial blueprints for my own approach to teaching their meaning and purpose. While historical trajectories need not be rigid guidelines, they should serve as practical starting points for developing materials. As such, sequences of problems should follow a historical and logical order allowing the student to intuitively and progressively develop the concept, rather than starting at the end (as many textbooks do by simply defining variables formally upfront).

My intuition, in line with active learning, is that students will learn from historical parallels best if they are led along a path of problems with conceptual roadblocks that mirror points of historical innovation and are encouraged to attempt to overcome each obstacle using their preexisting mathematical knowledge and intuitive reasoning. While I wouldn't expect most students to individually develop a system that took mathematicians over a thousand years within a classroom session or two, the important point is that the students struggle enough with their current mathematical toolkit's limitations in order to realize what the conceptual roadblocks *are* before being guided toward a solution. My hope is that students' struggle will increase both their understanding of and appreciation for variables as a conceptually and practically useful tool, not just for a given problem in class (proximal/distal academic purpose) but also for problems that are likely to occur in their own lives (real-world purpose). And the more of the concept that the student comes up with on their own, the more they will benefit from the experience (Chi et al., 1989; Pólya, 1957; Stigler & Hiebert, 1999) because information that is self-generated is better remembered than information that is provided externally (Bjork, 1994), and presumably, it is better tied into pre-existing conceptual knowledge (Hiebert & Grouws, 2007).

1.4.3 Mechanisms

1.4.3.1 Decomposition and Contrast Comparisons

The algebraic variable, its meaning, and its purpose, constitute complex concepts. To an appreciable extent, these concepts may be *decomposed* into constituent parts or aspects. Previous attempts at teaching this concept have focused on giving multiple concrete examples to help students to build schematic knowledge (Koedinger & Anderson, 1998; Nathan, 2012; Romberg, 2001). However, schematization is not the same as conceptual decomposition. Simply giving students multiple concrete examples to form abstract schemas from will likely not be sufficient for teaching most students the full meaning and purpose of the concept of a variable. It has been well established in analogy research that most reasoners do not spontaneously form meaningful connections between concrete examples (Gick & Holyoak, 1980, 1983); as such, students will benefit from using contexts that do not differ so drastically in surface details that their similar internal structure is obscured (Bassok, 1996; Kotovsky & Gentner, 1996). Yet while varied examples may demonstrate variables' formal abstractness and generalizability, schematic comparisons offer little in the way of explaining a variable's purpose.

To infer the purpose of a new concept, it is beneficial for most students to first make *contrast comparisons* between a case in which a goal is achievable and a case in which it is unachievable, or at least much more difficult to achieve (Walker, Cheng, & Stigler, 2014). For example, by contrasting a one-unknown problem with a two-unknown problem (that share as many contextual details as possible), students may infer that variables are conceptually useful in representing unknown values in problems that are too complex to model and solve by pure arithmetic (whereas variables have less obvious utility in solving a single-step, single-variable problem like $2 + x = 6$).

The contrast comparisons used in my materials are based on logic similar to that of Walker, Cheng, and Stigler's 'causal contrast' approach (2014). Both the causal contrast approach and my own recruit students' implicit faculty of causal learning (Cheng, 1997) to intuitively elucidate a goal-directed solution, rather than teaching the solution explicitly upfront. Like Walker et al. (2014), I aim not just to teach students how to solve a particular problem type (though that is a laudable goal in itself), but more crucially, the underlying concepts and the purpose of each procedural step in achieving the goal. My approach uses purpose-driven contrast comparisons to encourage students to discover the foundational concepts of algebra, such as the purpose of symbolically representing an unknown value. Thus, the contrasts in my materials aim to elucidate not only differences between problems' structure, but also the representational tradeoffs between different notational systems.

My approach assumes that students will learn the most by at first using concrete examples that are contextually similar to each other. By making the contexts as closely matched as possible, we eliminate the confounding differences across problems that distract or prevent effective contrasts. An effective contrast thus holds non-essential details as constant as possible while changing only the 'cause' of the difference between the two cases (e.g., what makes one problem solvable and another unsolvable), which allows the learner to infer the purpose of some new tool (e.g., a new representation or solution strategy that makes the previously-unsolvable problem now solvable). The series of problems used in my materials are as closely matched for context as possible (to allow for easier structural comparisons) while not seeming overly redundant (some surface details change in order to distinguish the problems). At each step, contrast comparisons are used to emphasize the changes in mathematical structure as the problems progress, how such changes reflect the meaning and purpose of a variable, and how formal representation may help make such problems easier to solve.

One particular sequence in my materials consists of multiple contrasts across different possible values and different states of knowledge. First, students encounter a one-equation, two-variable problem (presented as a story in a video) and are asked to find one variable's value (the price of one slice of cheese pizza). After entering an answer or correctly responding that it can't be found, they are specifically asked whether there is enough information to find the value, and then they are then asked a series of questions asking about whether specific values would be possible. One of the three values is actually impossible (under the reasonable assumption that the other item's price isn't a negative number), but both of the other two values are possible. This allows them to implicitly contrast two possible values to one impossible, so that they may realize there is not enough information. Then in the next video, they encounter a second, independent equation with both variables (as a continuation of the same story) and are asked whether they *now* have enough information to solve it. Then after solving the problems, students are asked to give an explanation of why a value could be found for when they knew both equations' worth of information, but it could not be found when they only had one of those equations (with both variables present). This final contrast between the two phases of the problem (one equation only and then both equations) encourages students to understand the differences in each phases' structure and that a two-variable problem requires at least two independent equations' worth of mathematical information.

Other contrasts focus on decomposing mathematical representations. For example, word equations are introduced in order to represent a story problem's mathematical structure; students are given both the initial story's sentences and their corresponding word equations and are asked to contrast the two to see which is more useful for helping them to find the answer to their question. Just prior to this contrast, students' intuitions for the representation of unknowns were likely evoked as they attempted to solve the problem on their own; even though the problem can technically be solved arithmetically, such a strategy would be contextually-specific, informal, and likely difficult to execute without error. The contrast is thus students' first real step toward formalizing those intuitions into their eventual representation, which will provide a more systematic and general method. And later on in their learning, by contrasting word equation representations of variables with letter equations, students intuit the practical reason for 'abbreviating' variables as letters. Taken together, these two representational contrasts decompose the purpose of using letters as variables: first, variables allow a reasoner to represent currently-unknown values in a mathematical structure, and second, letters are simply more compact ways to represent unknowns than words or phrases.

Eventually, once the purpose of a variable as a symbol for representing an unspecified number (or generalized number) is inferred from the decomposition throughout contrasts in the learning process, the problem contexts can broaden in order to highlight the generalizability of the concept across more disparate contexts. If such contextual broadening happened too early, this understanding of purpose would be just as lacking across all contexts. Due to the practical limitations of this current research, I have focused thus far in the project on developing materials that encourage students to learn as much as possible the conceptual elements of algebraic variables and their purpose in both math and the real world. Planned future work will add further lessons (to be given afterward) that use more varied contexts for more schematic representations of solution strategies, followed by a greater emphasis on applying the formalisms and procedures to abstract, symbolic domains.

1.4.3.2 Contextual Facilitation of Problem Solving

Many educators (see Nathan & Koedinger, 2000a) and researchers (e.g., Cummins et al., 1988) believe that word problems are *necessarily* more difficult than isomorphic symbolic problems, though this may be slowly changing (Chazan, Sela, & Herbst, 2012). While some empirical evidence for this has been found with simple arithmetic problems (Cummins et al., 1988; Kitsch & Greeno, 1985) and direct translations from verbal descriptions to algebraic form (Clement, 1982), the reality is much more nuanced (cf., Baranes, Perry, & Stigler, 1989; Carraher, Carraher, & Schliemann, 1987; Chazan, Sela, & Herbst, 2012; Koedinger & Nathan, 2004). The attitude that word problems are inherently more difficult relies on the assumption that such problems must first be converted from verbal form to math notation before any solving can be attempted. To give a simple arithmetic example, it is assumed that a problem like 'how much does one apple cost if three apples cost $6?' must necessarily be converted into the isomorphic mathematical notation ('$3x = 6$') before solving, so it must be at least as difficult as just solving '$3x = 6$'. However, this assumption does not always hold.

Koedinger and Nathan (2004) empirically demonstrate two important reasons why word problems can sometimes be *easier* to solve than their symbolic isomorphs. First, contextual situations may cue reasoners to use domain-specific knowledge and solution strategies (e.g., it may be more readily apparent that a dollar divided in four is a quarter (25¢) than that 100 divided by 4 is 25, Baranes, Perry, & Stigler, 1989). Such informal strategies may be domain-specific, limiting their generalizability, but they may also reduce cognitive load, which is especially beneficial during early learning, when additional cognitive resources are necessary to extract critical information (Sweller, 1988). Second, students have a greater amount and variety of experience with natural language than they do with mathematical notation, and so their fluency in understanding problem structures presented verbally may in fact be higher. So, despite the fact that real-world contexts add mathematically unnecessary surface details on top of a problem's internal mathematical structure, if such details are well-chosen, they may enable students to both more accurately represent the problem structure and to leverage their real-world knowledge in finding the solution.

There are many other associated factors that may contribute to such advantages to word problem contexts. Relatable stories with interesting concrete details (not just generic word problems, Gerofsky, 1996, 1999; Nesher, 1980) may be more interesting and enjoyable for students to think about (e.g., pizza is more fun than abstract x's and y's), which could give students more buy-in. Similarly, concrete entities may actually be easier to attend to, thank about, and remember (requiring less cognitive resources) due to their imaginability and familiarity. Certain types of concrete entities may be less error-prone in context, such as better decimal alignment with money due to the contextual distinction between dollars and cents (Koedinger & Nathan, 2004). And in many ways, concrete entities may be easier to 'track' throughout a problem solution, especially which objects' quantities or values are known or unknown and how they relate to each other, which in turn would reduce cognitive load. Since examples given early on in learning often overload cognitive resources, preventing more generalizable learning and inferences into an emerging concept's purpose (Sweller, 1988), any reduction in cognitive load may benefit learners.

Let's consider a simple example. A word problem about purchasing food items could use appropriately realistic quantities and prices to 'bias' students toward more plausible value ranges. Because it's easier to imagine a person ordering two slices of cheese pizza and one bottle of water than two bottles of water and one slice of cheese pizza (purely based on real-world experience), the sensibleness of the items' quantities makes them more memorable, which reduces the student's working memory load and eases their mathematical thinking. Additionally, if the items' prices are sensible for a real-world case, it would be easier for students to substitute one item's newly found price into an equation to find the other item's price. But using plausible values in and of itself is not going to produce generalizable student learning if it merely allows them to guess answers more easily; instead, this facilitation must be implemented in such a way as to tie the problem into the purpose of using a variable.

Not surprisingly, this simple kind of contextual biasing with real-world quantities is actually commonly used, it is just not used in such a way as to actually help teach the purpose of a variable. In traditional educational methods, such a word problem would not appear until after instruction with abstract formalisms (Nathan, Long, & Alibali, 2002; Sherman, Walkington, & Howell, 2016), and when it does appear, the contextual facilitation is typically negated by encouraging students to strip away all contextual information by converting the problem into purely symbolic form prior to solving it (Gellert & Jablonka, 2009; Verschaffel et al., 1994, 1997). Contextually biasing problems should instead be used as just one element of many used in a sequence of problems structured to teach the concept of a variable. Moreover, biasing should be used in such a way that when students encounter a problem that is not immediately obvious how to solve, guessing alone is not sufficient, but they can use contextual information to scaffold their thinking about the problem so that it leads them to the purpose for a new concept, and then that discovered purpose is actually relevant to the students' real world contexts.

Of course, if students *only* ever saw contextually 'biased' or schematic examples that were never tied into purpose-driven materials, they would never be able to learn the abstract formalisms that provide general, *un*biased means of solving problems (e.g., students should eventually be able to solve any analogous problem, even if the answer is that a bottle of water costs an unrealistic $3,235,987.23). The nuance of domain-specific contextual advantages is that they may facilitate better learning early on in instruction (if implemented well) by essentially 'biasing' students toward contextual mathematical thought that can then be mapped onto abstract symbols that generalize such thought to novel contexts and the abstract domain, where such contextual advantages are unavailable or unreliable. Just as a schematic representation of problem solution strategies can aid solvers who have correctly categorized a new problem by its internal structure (Sweller, 1988), domain-specific numerical knowledge can aid solvers who are given appropriate contextual cues in a new word problem.

Importantly, while contextual facilitation may free up cognitive resources and improve task-specific performance, these benefits alone are not sufficient to produce generalizable knowledge (i.e., generalization does not follow automatically). Contextual facilitation must be embodied within a well-designed learning environment in order to encourage students to induce the purpose of new formal representations (e.g., variables) and to map their context-specific, intuitive thinking onto the more general system. In essence, we do not just want to reduce learners' cognitive load, but also to provide them with a valuable task in which they can capitalize on their freed-up cognitive resources.

My approach is to use word problem contexts that allow students to think about the problem's mathematical structure in contextual, not necessarily symbolic, terms. However, it is of course an essential property of algebraic variables that any symbol can be used to represent any number (or set of numbers), regardless of whether it is plausible in context or even has an associated context. So, while my current approach implements meaningful and purposeful contexts upfront to facilitate conceptual learning, planned extensions of these materials will eventually aid students in schematic induction of broader solution strategies and in abstraction of both the concepts and procedures to symbolic algebra. In short, I am not advocating for all learning to be situationally bound (context is no panacea, Anderson, Reder, & Simon, 1996) since abstraction is crucial, but rather, that an initial grounding in appropriately chosen contexts may provide a better foundation for subsequent abstraction.

1.4.3.3 Active Learning and Constructive Struggling

Active learning is traditionally characterized by student engagement in the learning process (as opposed to the passive learning of reading or being lectured at). As Prince summarizes, "active learning requires students to do meaningful learning activities and think about what they are doing" (2004, p. 223), which helps them to make connections between the current task and previous knowledge, and these connections in turn facilitate better memory for the material later on. There is broad support for active learning, but different methods lead to different effect sizes (Prince, 2004).

One way in which students can be engaged in the learning process is to allow them to try to work through a solution to a novel problem type before teaching them a way to do it. Allowing students to *constructively struggle* and provide their own explanations are hallmarks of Japanese education that are less commonly used in American education (Stigler & Hiebert, 1999). The idea of constructive struggling also ties into the notion of "desirable difficulties" in learning and memory (Bjork, 1994). In essence, students are allowed and encouraged to generate that knowledge or strategy on their own, which they will remember better than if they had just been told what the answer is or how to find it.

Constructive struggling also allows students to identify where the gaps are in their own understanding. For example, if you give a student a system of equations problem before telling them how to solve it, then as they think through the problem, they may realize that they don't know yet how to combine information across two different equations. Ideally, the student can then fill that gap by attempting their own solutions, and in doing so, they will remember it even better for having gone through the process themselves and feel more positively for having done so.

It is thus important to challenge students, but in the event that they struggle through a challenge but fail to produce a correct solution, we should be ready with appropriate hints to help guide them in the right direction. I take a strong influence from Pólya's (1957) classic view of teaching problem solving. Pólya, a mathematician and educator, lays out the case for giving learners increasingly more specific hints, as needed. Ideally, students will learn without any hints or from the most minimally specific hint possible, effectively making the learning more active than it would have been with a more specific hint. In my materials, hints and explanations are framed as much as possible *at the level of the context itself*, as opposed to relying on overly formal or abstract concepts that have not been learned yet (as is done in the schematization in Koedinger and Anderson's cognitive tutor, 1998). These hints are used to guide students toward an intuitive way of thinking about the problem that naturally leads to a solution procedure.

The benefits of active learning and constructive struggling are clear *per se*, but in the context of this approach, they are compounded in importance. Because students attempt to solve problems that are difficult without the use of formal algebra, they will be encouraged to infer the need for a new tool, and they may even come up with it on their own. And if they struggle to do so, the gradual hint system will lead them toward an intuitive solution. Finally, new means of representation will be progressively introduced, and because students struggled through the problem on their own and developed their own intuitions, they will more fully understand the purpose of variables as a representational tool and be able to integrate it with their previous thinking.

1.4.4 Purpose-Driven Progressive Formalization (P-PF)

All of the preceding considerations naturally converge into a framework in which students begin their learning trajectory with carefully selected contextual word problems, ordered in a sequence that highlights conceptual components with questions that encourage contrast comparisons, and allowing students to generate as much of their own insight into these problems and their concepts as possible, with minimal hints as needed. The broad strokes of this framework are similar to Romberg's (2001) and Nathan's (2012) progressive formalization views, although my own approach is highly motivated by encouraging students to intuit the purpose of a new concept. Thus, instead of employing generalization across concrete examples to abstract a schematic representation (Koedinger & Anderson, 1998; Nathan, 2012; Romberg, 2001), my approach utilizes contrast comparisons to elucidate the purpose of such a representation in the first place. While these approaches are complementary, not mutually exclusive, it is helpful to distinguish between their emphases. As such, I refer to Nathan's (2012) and Romberg's (2001) approaches as *Generalization-Based Progressive Formalization* (G-PF) and my own as *Purpose-Driven Progressive Formalization* (P-PF).

1.4.4.1 Context in P-PF: The Plausibility, Relatability, Motivation (PRM) Framework

As discussed above (1.4.3.2), the use of word problem contexts can facilitate mathematical problem solving (Baranes, Perry, & Stigler, 1989; Koedinger & Nathan, 2004), but it depends on the circumstances and the learning objectives (Anderson, Reder, & Simon, 1996). Moreover, as Ainley argues, "The danger lies in thinking that having used problems in which the world is used to make sense of mathematics, we have also achieved the aim of giving children opportunities to see how mathematics can be used to make sense of the world" (2012, p. 94). Her point is well-taken that the vast majority of word problems embed the math into impoverished contexts in order to give (internal) meaning to the math, but this does not mean that the math has been given external meaning (i.e., real-world relevance). Context plays an important role in my P-PF approach as a means for facilitating students' intuitive thinking during constructive struggling and demonstrating the real-world purpose of the concept as it is learned. However, a framework is needed to assess the kinds of contexts to be used for a given problem type.

If we want to write word problems that have more effective and/or "real world" contexts, how do we determine what contexts are the best? First, let's consider what clearly doesn't work. Some common system of equation problems I've found in algebra textbooks are what I classify as "puzzle problems" or "riddles" because of their emphasis on cleverness over realism. The following are some examples that I pulled from the internet (for ease of reference), but they are highly representative of these puzzle problems found in textbooks:

1. “The sum of the digits of a two-digit number is 7. When the digits are reversed, the number is increased by 27. Find the number.” (from http://www.purplemath.com/modules/systprob.htm)
2. “A woman is now 30 years older than her son. 15 years ago, she was twice as old. What are the present ages of the woman and her son?” (from http://www.themathpage.com/alg/word-problems3-2.htm)
3. “Joe counts 48 heads and 134 legs among the chickens and dogs on his farm. How many dogs and how many chickens does he have?” (from http://mathforum.org/library/drmath/view/57511.html)

These puzzle problems may be great entertainment and enrichment for the kind of student who loves math and puzzles (I personally found them fun when I was taking Algebra), but for students who struggle with math, especially those who do not see its real-world relevance, these puzzle problem contexts are counterproductive, as they fail to convey meaning or purpose.

The first issue with these puzzle problems is that their contexts characteristically lack *plausibility*—the basic question that comes to my mind when reading these is “When would we ever know the details that are given but not know those that are not?” For example, when would we ever be faced with a situation where we know how many heads and how many legs are on a farm, but not know how many of those heads and legs belong to chickens or dogs? (Who counts heads and legs but not animals?) Plausibility is important in order to demonstrate the practical relevance of math to the real world. Another flaw in these types of problems is that their contexts have no obvious *relatability* to the student (personal relevance): none of these puzzle problems’ contexts are likely to be encountered by a typical Algebra student. A third problem is that it is unclear why the character in the story (if there is one) or the student herself is *motivated* to find the answer. Why should we care about finding some two-digit number, the ages of the mother and son (don't they know their own ages?), or the numerosities of a farmer's chickens and dogs (shouldn't a farmer already know?)?

The implausibility, irrelevance, and lack of motivation in these kinds of problems render them poor devices for student learning, especially in my P-PF approach. For many students, puzzle problems simply encourage them to ignore word problem contexts and further their belief that math is not useful in the real world (Gellert & Jablonka, 2009; Verschaffel et al., 1994, 1997). In order to produce better student learning, we should seek to use problems that are plausible, relatable, and motivated:

- *Plausibility*: Is our knowledge base in the problem *plausible* in the real world?
    - Is it plausible (in the real world) that we *would* know what we know (in the problem)?
    - Is it plausible (in the real world) that we *wouldn't* know what we don't know (in the problem)?
        - E.g., "Why don't we already know the value of *x*?"
    - Are the numerical values we *do* know (in the problem) plausible (in the real world), and are the eventual numerical values of what we don't *yet* know (in the problem) plausible (in the real world)?
        - (This may help contextual facilitation, which can reduce working memory load by requiring fewer cognitive resources for tracking knowns and unknowns, and also demonstrates real-world purpose of the concept when it is learned.)
- *Relatability*: Is the problem's context *relatable* (personally relevant) to the student?

- Is it likely that the student has encountered this type of context previously?
- If not, could the student at least relate to the context?
  - (These considerations are important for demonstrating that the variable has real-world purpose and for helping students to map its use onto something easier to think about.)

• *Motivation*: Is our goal-state/end-product clear, and is there a *motivation* for the student (and the character in the story, if there is one) to find it?

- E.g., "Why do we care about the value of *x*?"
  - (Motivation is useful for making the problem's task clearer and more concrete to students, which in turn should encourage their intuitive thinking and willingness to constructively struggle, in addition to demonstrating concepts' task-specific purpose.)

I combine these three considerations into the *Plausibility, Relatability, and Motivation* (PRM) Framework for word problem contexts.

Of course, these three qualities may often overlap (for example, the motivation for finding a solution should itself be relatable in-context and it should be plausible in-context why the solution is not already known). However, it is possible for the qualities to not overlap (there could be a relatable context with implausible values and an unmotivated query), which would likely lead to poorer outcomes. For example, in the "A woman is now 30 years older than her son; 15 years ago, she was twice as old" problem, it would be completely ineffective to just add "Now the woman wants to know how old she is because she's filling out some paperwork." This 'motivation' is simply too implausible and unrelatable. In essence, such PRM context would be added 'window dressing' and is not going to fix the essential fact that it's simply too arbitrary to be a PRM-type problem. Ideally, a PRM word problem context is one in which the PRM qualities follow naturally from the context itself so that these details are believable to the student, and they overlap and reinforce each other.

These PRM qualities are particularly relevant to my P-PF approach because they are designed to achieve contextual facilitation in ways that satisfy the broader goals of P-PF, such as encouraging intuitive mathematical thinking and demonstrating real-world purpose. The PRM framework is not meant to be applied to any and all mathematical learning or practice item. For example, when students begin to abstract schemas across examples, such generalization should be carried out for non-PRM problems as well. In fact, schematic generalization may be best achieved across a sequence of problems that gradually *decrease* in PRM qualities, eventually reaching problems with no context at all (i.e., formal symbolic math). Thus, the PRM framework is formulated as a guide for effective P-PF instruction, not for all types of instruction or practice problems.

1.4.4.2 Summary

The previously discussed aspects of my P-PF approach dovetail naturally to provide students with the environment to develop better insights into the concept of a variable, its meaning, and its purpose. Because the concepts and tasks have been decomposed as much as possible across problems in the sequence and in their representations (while holding non-essential details as constant as possible), students' contrast comparisons allow them to infer the constituent components of each concept and intuit the purpose of new representational forms. Students should be engaged by PRM problem contexts, and being able to draw from contextual knowledge may reduce the tasks' cognitive load, freeing up students' attentional and memory resources so that they can make more informative insights into problem structures and solutions; moreover, the PRM details should help to encourage intuitive thinking and motivation through constructive struggling, as well as to demonstrate the proximal task-specific and real-world purposes of the learned concepts. Constructive struggling, combined with increasingly specific hints as necessary, provides students with as much opportunity as possible to make their own intuitive insights into problem structures and solution procedures, which can subsequently be mapped onto the concepts and representations introduced as the materials progress.

All of these various angles converge to lead students to gradually learn the concept of a variable in a contextually-grounded and purpose-driven manner as more formal representations are progressively introduced. Each of these primary mechanisms follows naturally from a pedagogical emphasis on encouraging students to intuit the purpose of a variable. Moreover, each mechanism reinforces the effects of the others, likely increasing their individual efficacies superadditively. These mechanisms are merged into a cohesive framework, P-PF, which provides the environment best suited to progressively introduce new formalisms in such a way that their purpose is intuitively understood by the learner.

2. Background Literature

2.1 Mathematical Formalisms

The pure entities of mathematics are *abstractions*—generalized from a variety of specific experiences, pedagogies, and mental simulations. Our means of expressing mathematical ideas by verbal descriptions, written symbols, or graphical visualizations are merely different choices of *representation* for an underlying abstract notion. Each form of representation comes with its own set of strengths and weaknesses, and each is necessarily characterized by certain arbitrary perceptual features that must be learned from experience (not unlike the arbitrary phonology and syntax of natural languages). For example, nothing about the visual features of the numeral '2' necessitates a connection to the orthographic symbol 'two', phonetic symbol /tuː/, or the concept TWO—the associations between these representations and to the underlying concept must be learned from experience. Formal representations, or *formalisms* (Nathan, 2012), are most apparent in mathematics as symbolic notation (e.g., Arabic numerals, mathematical operators) and the syntactic rules (procedures) for manipulating these symbols (e.g., the distributive property, the long division algorithm).

Nathan characterizes formalisms in a *narrow* sense as "specialized representational forms that use heavily regulated notational systems with *no inherent meaning* except those that are established by convention to convey concepts and relations with a high degree of specificity" (2012, p. 125, emphasis mine). In the case of algebra, there are many instances of such narrow formalisms, including the symbolic form of variables (predominantly italicized Latin letters, and in higher math, Greek letters), numerals, and arithmetic operators, as well as the syntactic rules of arithmetic and algebraic symbol manipulation. Additionally, there are *broader* senses of formalisms (Nathan, 2012), which in algebra may include the conceptual underpinning of variables, the generalizable nature of algebraic equations, and conventions such as y being dependent on x. These broad formalisms are important aspects of the meaning and purpose of variables, and how they are used.

Formalisms are an integral part of mathematics generally and algebra specifically. They can serve as organizing principles, increase computational efficiency, reduce ambiguity, provide parsimonious representations for complex ideas, allow for generalization across contexts to highlight commonalities, and serve as a conceptual bridge between different domains (Nathan, 2012). All of these characteristics are indispensable to math, especially so with algebra and higher levels. However, the reasoner must be fluent with a formalism's notational and rule systems in order to reap these benefits in solving a mathematical problem.

For example, a function can be represented formally as '$y = 3x + 2$'. From this notation, a fluent reasoner can glean that the function denotes an infinite set of [x, y] value-pairs that satisfy a general rule (each y value is always two more than the triple of a corresponding x value); that the relationship between x and y is linear, which in turn means the function can be represented by a straight line on a Cartesian coordinate plane; that each unique x value corresponds to a single unique y value; that the function would intersect with a second, independent linear function at exactly one point; that purely by convention, y is deemed dependent on x, but not necessarily vice-versa; etc. These conceptual interpretations of the formalisms (e.g., the denoted infinite set of value-pairs) and connections to other forms of representation (e.g., the geometric line) are in no way entailed by the perceptual character of the formalism itself. The parsimonious expressiveness, clarity, and generality of '$y = 3x + 2$' is only available through fluency with its representational form. In particular, the fact that the notated relationship between x and y is linear is immediately obvious to the highly fluent from the lack of powered terms, but there would be no indication of this to a novice. Without such fluency, reasoners are left holding tools they know not how to wield.

The very same features of formalisms that endow them with their desirable properties (e.g., arbitrary perceptual features, exact rules of manipulation, generalizability, etc.) also make their meaning and purpose opaque to many beginning students. In leveraging abstractness for complexity and generalizability, symbolic notation is semantically weak but syntactically strong (Wheeler, 1989), allowing for formal and operational systematicity across any imaginable mathematical context. The lack of intrinsic meaning in formalisms is not a bug but a feature of their design, imbuing them with representational diversity. For example, the letter x can stand for *any* numerical quantity (singleton), finite set of numbers, or even infinite range of numbers, and all of these numerical possibilities may in turn map onto any real-life or imaginary quantitative construct—the single letter x can represent the amount of money in Warren Buffet's checking account at the turn of the millennium, the set of all abscissa coordinates for the vertices of a drawn trapezoid, or the infinite range of possible centigrade readings that can be converted to Fahrenheit using the same exact linear function. The pure abstraction of 'x' instills in the symbol the flexibility to represent each of these cases equally well, and thus, the *meaning* of a given x depends entirely upon the mathematical structure and situational context in which it is instantiated. We should not expect that all students can understand such an abstract concept right away—the meaning and purpose of a variable must be taught.

2.1.1 How Formalisms Are Typically Taught: The Formalisms First (FF) View

Given formalisms' syntactic-semantic tradeoff, it should come as no surprise that students are often confused when introduced to formalisms: they may ask, for example, as a canonical, high-performing 12-year-old math student named Henry did, "How can we multiply by *x* when we don't know what *x* is?" (Davis, 1975). Students' difficulties with interpreting formal notations are exacerbated (if not largely caused) by the predominance of what Nathan (2012) terms the *formalisms first* (FF) view—a predominant, often implicit belief that students must first learn the formal notations used in a domain before they can later connect these abstract representations to practical and applied knowledge. In the case of algebra, instruction often begins by defining a variable formally (e.g., as "a letter that represents a number") and then proceeding to give worked symbolic examples (e.g., how to transform '$x + 2 = 5$' into '$x = 3$') followed by students' rote practice (Nathan, Long, & Alibali, 2002); only later do students *apply* the formalisms to contextual problems (usually, generic word problems, Gerofsky, 1996, 1999; Nesher, 1980). The FF framework endorses a view of learning and development that moves from the abstract and symbolic to the concrete and applied (Nathan & Koedinger, 2000b).

Curricula are often determined and organized by content experts, who, embracing FF (even if implicitly), assume that the most theoretically fundamental (central) aspects of a topic must take precedence (be 'fronted') in the learning trajectory of a student. This assumption may often be fallacious, especially with formalisms, which are abstract and difficult to comprehend without any prior context or experience in the domain. The FF view is widespread in algebra, implicitly and explicitly, as evidenced by textbook analyses and surveys of teacher attitudes (Nathan & Koedinger, 2000a, 2000b; Nathan, Long, & Alibali, 2002; Sherman, Walkington, & Howell, 2016).

Specifically, educators assume that mastery of algebraic symbols *precedes* algebraic applications to practical purposes. Nathan, Long, and Alibali's (2002) corpus analysis of 10 common algebra textbooks demonstrated a predominant pattern of introducing concepts in symbolic form before giving any isomorphic problems in word (story) problem form, reflecting what the authors term the "symbol precedence view" (an instantiation of FF that is specific to algebra). And according to Nathan (2012), while there have been subsequent efforts to prioritize application problems in newer reform-based textbooks, such examples are often relegated to supplements. Moreover, given teachers' (both preservice and in-service) and researchers' beliefs that symbolic forms provide a more natural introduction (Nathan & Koedinger, 2000a, 2000b), educators may choose to avoid or delay such applied problems even if they were moved earlier in a textbook. In fact, Sherman et al. (2016) found that while the student versions of newer conventional textbooks do not significantly demonstrate such symbol precedence, the teacher editions for the same textbooks skip or delay verbal/applied problems anyway.

The prominence of the FF and symbol precedence views among educators is likely encouraged by what Nathan, Koedinger, and Alibali (2001) term the "expert blind spot", which is analogous to the "curse of knowledge". Simply put, because educators are already experts in their domain who have high fluency with symbolic notation, it is difficult for them to model the lower fluency that their novice students have; this problem is compounded by the fact that many math educators likely did not struggle in their own math education the way that some of their students might. As such, teachers may be unaware of the symbols' opaqueness, and even if they are aware, it still may be difficult for them to unpack their own fluency. The expert blind spot is a common problem at all levels of education and educational research (and I am certainly not claiming personal exemption), not one unique to Algebra teachers—abstract, complex concepts are difficult to teach regardless of the field. However, this particular blind spot is especially problematic in Algebra due to the fact that variables' formal notation piggybacks on familiar symbols (letters) while using them in unfamiliar ways (to stand for numbers), and students clearly struggle with this notational change (Herscovics & Linchevski, 1994).

The symbol precedence view leads to students being introduced to variables as abstract symbols, and only later (if at all) connecting the symbols to their meaning and purpose. As researchers and educators dedicated to improving struggling students' mathematical skills, we must examine how we communicate mathematically with students, in terms of what kinds of formalisms we use, how we introduce them, and how they affect and are affected by students' prior knowledge or lack thereof. If we do not create points of connection between modes of representation and the student's intuition, then we are going to be left with students who, regardless of whether they can provide a correct answer, do not understand what they are doing, why they are doing it, or how they would do it in a real-world context. Importantly, these students will not be able to utilize the analytical framework provided by formal mathematics in order to solve practical problems, one of the supposed goals of formal math education (NCTM, 2000).

2.1.2 Why Formalisms Can Be Difficult: Duality of Representation

While the FF framework emphasizes formal symbols early on, the formalisms' underlying concepts are not always well taught—in transitioning from arithmetic to algebra, for example, it is largely assumed by educators that students unproblematically interpret the literal symbols of algebra from the get-go (Sfard, 1991), and yet, there is a demonstrated "cognitive gap" between arithmetic and algebra (Hersovics & Linchevski, 1994). This problem is compounded by the aforementioned "expert blind spot" (Nathan, Koedinger, & Alibali, 2001).

A major contributing factor to students' confusion about formal notation stems from the fact that mathematical expressions can be interpreted in two interrelated ways: as the *process* of carrying out the designated operation(s), and as the resulting *quantity* of the operation(s). Davis (1975) identified the 'name-process dilemma', in which an expression such as '6*4' can indicate both a process of multiplying 6 by 4 (by means of a specific addition procedure or algorithm) and the mathematical object denoted (named) by the product of 6 and 4 (which is equivalent to 24). Sfard and Linchesvki (1993) renamed this as the 'process-product dilemma', and it's reflected in both Sfard's (1991) and Gray and Tall's (1994) frameworks. This dual representation may seem obvious or even trivial to proficient students and teachers, who easily alternate between the two interpretations, but many struggling students do not treat an expression of an operation as a mathematical object in its own right (Sfard, 1991). Specifically, to these students, '6*4' has always been a cue to carry out an arithmetic operation, and they haven't needed to develop an object conception of it (i.e., there hasn't been purpose to treat '6*4' as a quantity).

The importance and power of this duality is amplified as arithmetic leads into algebra. An algebraic expression such as '6*x' cannot be operationalized to the same extent as the arithmetic '6*4' (due to the inability to combine unlike terms such as known and unknown quantities). Rather, '6*x' is more usefully conceived of as the quantity produced by the multiplication of 6 by an unspecified number denoted by 'x'; that resulting quantity is a quantity in its own right, which can serve as an input to yet another operation (e.g., addition, in $6*x + 2$) or just be considered by itself. If a student understands why a variable like x is useful, then they would understand what '6*x' means. But because many students do not understand variables', '6*x' is opaque in meaning: it cannot be operationalized like the arithmetic '6*4', and it would not be sensible or useful as an object either, assuming students can even understand it that way.

The object conception of a variable also underpins the procedure of dividing 6**x* by 6 in order to isolate *x*; without such a conceptual understanding, students merely learn to perform 'opposite' operations (e.g., if a variable is being multiplied by 6, then divide by 6) without understanding *why* they are doing so. Students who conceive of arithmetic expressions only operationally will struggle to interpret algebraic expressions, for which the operational conception is less accessible. In my materials, variables are progressively introduced as they become useful for solving more complex problems, and contrast comparisons encourage students to understand the purpose and meaning of variables as they are used; thus, when students understand why an *x* is being used and what it represents, they will understand that an expression like '6**x*' is not taken to cue them to carry out multiplication, but rather, to represent 6 times an unknown quantity (as an object in its own right).

Sfard (1991) defines mathematical concepts (notions) as formal theoretical constructs (existing only in the abstract, formal realm of pure theory), which can be subjectively conceived in two qualitatively different ways: as objects (*structural* conceptions) and as processes (*operational* conceptions). Her distinction between structures and operations is at the psychological level (internal representation). Sfard's account implies that many mathematical concepts are not fully understood until both operational and structural conceptions are learned. Mathematical structures, by allowing the reasoner to simplify the procedural details of an operation into a static whole, reduce the cognitive demands on working memory capacity and allow for the kind of higher-order relations and broad generalization that characterizes mathematical thought. For example, it may be helpful to conceive of some ugly mathematical expression like 482.5*13 + 2 as being 2 greater than the product of 482.5*13, whatever that product may be (a generalization of this conception would allow us to say that $x + 2$ is always exactly 2 greater than x).

Using the same surface notation to represent interrelated meanings (conceptions) is a powerful feature of formal notation, in that it visually encapsulates the relationship between a process and its result, and makes mathematical equations more expressive by allowing for the reasoner to talk about the product of 6 and 4 without actually carrying out the operation of multiplication. But, as with all properties of formalisms, this expressive power is only useful insofar as it is understood. Being able to alternate between both procedural (process) and conceptual (object/name) interpretations is what Gray and Tall (1994) call *proceptual thinking* (a portmanteau of 'process' and 'conceptual'), which is characterized by an understanding (often implicit) of the connection between the two modes of interpretation. It seems likely that an attempt at explicitly teaching proceptual thinking would be confusing for students unless they first understood the purpose of the representation and had an intuitive sense of how the conceptual interpretation would be useful.

To a student who struggles to conceive of mathematical expressions as objects, they may instead turn to an operational conception of rote symbolic manipulation. Even those students who have learned how to manipulate algebraic variables will still be stuck with only an operational conception ($6*x$ is a cue to divide by 6 to isolate x) if they fail to develop proceptual thinking. This procedural interpretation of algebraic formalisms obscures the interpretation of a variable as an unknown or unspecified quantity. Ideally, we want to teach students to parse an algebraic expression such as $6*x$ as "the product of 6 and the unknown quantity, x" and connect the resulting (unknown) quantity to a corresponding (hypothetical) procedure, even if the procedure cannot be fully carried out until first determining the unknown *x*'s numerical value.

Research has shown that many students fail to interpret algebraic expressions as mathematical objects, and instead treat variables in algebraic expressions in a highly procedural manner that is qualitatively different from the procedures of arithmetic—a variable is to be moved by itself to the left side of the equation in order to reveal its value on the righthand side (Mayer, 1982). Given this procedural focus, many operations in algebra are actually inverted; for example, to isolate *x* from 6**x*, you use division instead of multiplication. Combining the insight that arithmetic operations are apparently inverted in algebra with the observation that variables are expected to be procedurally isolated from constants, it seems unsurprising that many students treat variables as somehow distinct from the world of arithmetic. This may lead to students' feeling that algebra is *fundamentally different* from, rather than an *extension of*, arithmetic. And as long as students are able to treat algebraic equations in such a different procedural manner while still producing answers acceptable to the teacher, they have no real motivation to interpret 6**x* as a mathematical object. But it is this interpretation that is necessary to fully link algebra to the established world of arithmetic. In fact, in the recently adopted Common Core standards, 6th grade students are expected to be able to "view one or more parts of an expression as a single entity" (CCSS.MATH.CONTENT.6.EE.A.2.A).

In my materials, to encourage students toward a conceptual interpretation of variables and variable expressions as numbers in their own right, I first give them a problem (system of equations type) that is difficult to solve in purely algebraic terms without having the concept of a variable at least as a placeholder; then, after they intuit the need for a new approach and representation, I introduce word-variables (e.g., 'the price of 1 donut') as representational bridges to letter-variables (e.g,. '*d*'), and consistently refer to both types of variables as representing *numbers*. Throughout the materials, students are asked several thought-provoking questions about the usefulness of various representations. While not explicitly teaching proceptual thinking, it is my hope that these purpose-driven lines of thought lead them to a better understanding of ways of interpreting the same notation.

2.1.3 Why Algebra Seems Distinct from Arithmetic: The Cognitive Gap

Herscovics and Linchevski (1994) refer to the disconnect between arithmetic and algebra as a *cognitive gap*—for many students, symbols and notation that are common across both arithmetic and algebra are thought to take on different meanings in each context. Because algebra is an extension of (not a separate domain from) arithmetic, any perceived gap hinders students' ability to leverage their previous understanding of arithmetic (even if incomplete) in new algebraic contexts. This gap underlies struggling students' lack of understanding of the meaning and purpose of algebraic variables and how they relate to the numbers of arithmetic. Herscovics and Linchevski argue that traditional Pre-Algebra classes and texts have not helped to close this gap because they simply introduce algebra earlier and more slowly, rather than providing a conceptual foundation.

Herscovics and Linchevski (1994) offer the term *operational generalized number* to describe the desired end-state of the concept of a variable, such that *x* may possess the properties of any number or number set, but it is also capable of being operationally manipulated even without its value(s) being specified. They add: "Although the letter in an equation or an algebraic expression may have a numerical referent in the pupil's mind, this does not necessarily render it *operational*. *Meaning* for these operations with literal symbols still has to be constructed” (Herscovics & Linchevski, 1994, p. 63, emphasis mine). As such, it is not enough to know that a letter may represent a number, or even that it represents numbers in general, but the student must also be able to perform operations with and on that general number, while understanding that its quantity remains unchanged throughout. Clearly, Herscovics and Linchevski are concerned with students’ conceptual understanding of the meaning of variables, but it is also important that students understand the purpose of such a concept in the first place. In my approach, the purpose of a variable is intuitively demonstrated to students before introducing the formal representation, so that variable formalisms are already meaningful when introduced.

Herscovics and Linchevski conducted a study with a classroom of 12-year-olds who had not yet taken Algebra, and concluded that the cognitive gap can be identified as “the student's inability to operate with or on the unknown” (1994, p. 75). This relates closely to what Collis (1975) identifies as students failure to exhibit “Acceptance of Lack of Closure” (ALC). Borrowing the term from Piaget, Collis notes that students developing formal operational thought must learn to perform operations on quantities that include an unknown even while accepting that it *is* unknown. Students’ lack of ALC reinforces the heavily procedural conception of variables at play, in that variables are conceived as mere tokens to be manipulated until isolated on one side of the equation; in duality terms, students fail to comprehend variables’ structural meaning. Yet again, it is easy to see that if students do not understand the purpose of using a variable to represent a number, they will not have reason to accept lack of closure, because the procedure of moving a variable around in an equation using mathematical operations seems arbitrary and disconnected.

Insofar as students build on arithmetic knowledge in algebra, some misconceptions in algebra indicate that the symbols and operations of arithmetic were not so well understood (Booth, 1988). For example, the use of the equals sign in school arithmetic is overwhelmingly as the final (i.e., rightmost) symbol in a sequence of numerals and operators (e.g., "2 + 4 - 3 = ") and is typically read as simply an operational prompt to carry out the computations to its left (Behr, Erlwanger, & Nichols, 1976); as such, students have difficulty when the left and right sides of the equals sign in such a use are flipped (decomposition, e.g., 6 = 2 + 4), or when operations appear on both sides (e.g., 2 + 4 = 1 + 5). For students who fail to learn that the equals sign represents equality, its use in algebra can get confusing. While the simple $2 + x = 6$ case may be easy enough to adjust to without reconceptualizing the equals sign, instances like $2x + 3x = x + 12$ are much more difficult to understand because x appears on both sides of the equals sign, so all instances of x must be moved to one side, combined, and then isolated in order to find x's value. Students may be able to memorize the requisite procedures of performing identical operations on each side of the equals sign, but in order to understand the meaning and purpose of this method, students must first understand that the equals sign expresses equivalence symmetrically and that such procedures then preserve the equality statement.

Perhaps the best way to prevent students' cognitive gap from ever forming is to first allow them to experience a contextually meaningful problem that is difficult to solve arithmetically without having some sort of a placeholder for the unknown. This encourages students to intuit the purpose of using a variable as essential to representing an unknown in more complex problems. Then as formal representations are progressively introduced, students have already laid the conceptual foundation on which these abstractions are built. In my framework, just as in the historical development of algebra, this foundation is ultimately driven by purpose, so that variables can be seen to grow out of arithmetic naturally, at first as a practical placeholder for a desired unknown, and then later as a means for expressing a general quantity that stands in some specific relation to another.

2.1.4 How Variables Came to Be: Historical Development of Formalisms

Given the goal-driven history of variables' development to solve practical problems (Kieran, 1992; Sfard, 1995), the historical trajectory is a useful model for the decomposition of the concept needed to lead students to intuitively learn the meaning and purpose of variables. Additionally, the history of algebraic variables is worth considering to better understand the nature of the reconceptualization: how variables provide a procedural and conceptual addition to our mathematical toolkit by allowing us to represent unknowns explicitly so that we may perform operations with them to find their value. This ability is useful for solving problems that have multiple unknowns to keep track of, or are otherwise complex in the series of steps that must be carried out to find the answer.

According to Klein (1936/1968), it took approximately 1300 years for mathematicians to move from using a single letter standing for a specific unknown quantity (Diophantus, circa 260 CE) to adding a second letter to represent a number whose value can be assumed to be given without actually being specified (Vieta, 1500s CE), which is roughly what we would today call an independent variable. Whereas the Diophantine variable is a placeholder for a quantity whose value we wish to discover, Vieta's innovation allows for the expression of general relations between quantities (e.g., as with '$y = 3x + 2$'). Both usages have their purposes, but the latter is much more conceptually complex.

Because the formal representation of algebraic variables is so powerful and essential to algebraic thought, experts operating in the FF frame of mind begin instruction by defining them formally (e.g., as ‘letters representing numbers’) without motivating the practice or leading students through this development. But if we examine historical developmental trajectories, formalisms are paradigmatically not an entry-point but an *end-result*—a capstone of abstraction atop a contextually-grounded foundation.

To teachers and students who excel in algebra, the interpretations and utilities of variables may seem uncontroversial and obvious. But the fact that 1300 years of mathematical geniuses came and went without making Vieta’s development should tell us something—the purpose and meaning of algebraic notation is probably not going to spontaneously occur to most students. And while we shouldn’t expect every student to ‘reinvent’ algebra variables in their own learning, encouraging them to follow similar lines of thinking to those that led to the development of variables will help give them an intuitive understanding of how variables are useful. In line with active learning research, the more of these conclusions students reach through their own intuition, the more they will understand the concept of a variable as it is introduced.

It seems reasonable to hypothesize that omitting the progression of algebra’s development and the inspiration for each innovation may leave students confused and personally unmotivated by algebra exercises. By emphasizing formalisms upfront, the predominant FF view essentially begins at the *end* of each historical development, not only depriving students of self-discovery in their own individual learning trajectory, but also potentially denying them the chance to ever fully learn the meaning or purpose of the formalism at all. In my approach, students are led through a sequence of problems designed to decompose the concept and its purpose, and they are encouraged to develop as much of their own intuitive understanding as they can; if necessary, hints and walk-throughs will help them to structure their thinking. Word equations are used as a representational bridge to eventual letter variables. Importantly, contrast questions are asked throughout this sequence, asking students to compare differences between problem structures and representations where a goal is reachable or not and where a goal is more or less reachable. These manipulations working in concert will partially replicate, on a personal scale, the historical development of algebraic variables, so that students can learn through experience how variables are useful.

2.2 Word Problem Contexts

In thinking about how we teach variables, we must also consider the role of context, imaginary or realistic, especially as used in word problems. An overarching goal of mathematics is to describe the real world in ways that are universal; however, practical applications receive considerably less emphasis than abstract forms of representation. A prominent method of introducing supposed examples of practical applications in mathematics is through the use of word problems (also called story problems). However, the current state of word problems, in terms of how they are written and how they are used in curricula, falls woefully short of supplying the kind of meaningful context that engages students or demonstrates the usefulness of math.

Mathematical word problems have been discovered in some of the earliest known examples of written language (Swetz, 2009), and when repeated over multiple content domains with common solutions, provided the primary evidence for the generalizability of mathematical principles before the development of algebra (Gerofsky, 2009). After algebra was brought to Europe during the Renaissance, word problems were repurposed as a means to demonstrate the practical benefits of mathematics for business and science (Swetz, 2009). In the modern day, word problems have become a fixture of math education (although more so in some topics than others, Stigler et al., 1986), even as they have become stale and stereotyped for both writers and solvers (Nesher, 1980). Commonly stated reasons for using word problems in math education include facilitating learning by making real-world connections and engaging students' interest, and assessing the transfer of arithmetic operations and algebraic representations to solving novel problem situations; in essence, word problems are thought to both enhance learning and also serve as difficult transfer questions (Ainley, 2012).

2.2.1 How Word Problems Are Typically Used: Generic Contexts

Students today struggle with word problems for many reasons; they are characterized by their formulaic generic structure (Gerofsky, 1996, 1999; Nesher, 1980), unrealistic models of the world (Ainley, 2012; Gellert & Jablonka, 2009; Verschaffel et al., 1994), ambiguous language (Cummins et al., 1988), and purposelessness (Ainley, 2012). To make matters much worse, educators' attitudes about word problems and their solution procedures indicate a stark de-emphasis of any consideration of problem content or relevant real-world knowledge (Verschaffel et al., 1997), consistent with the FF view.

Both students and teachers routinely ignore the content and context of word problems (Gellert & Jablonka, 2009; Verschaffel et al., 1994, 1997), and instead focus on routine, algorithmic translation strategies to convert them into arithmetic/algebraic equations which can then be solved, under the assumption that symbolic forms are solved more easily (Nathan &

Koedinger, 2000a, 2000b). Students are taught to ignore the content of a word problem and instead search the text for keywords that can be syntactically parsed into mathematical symbols (e.g., "if you see the word 'more', write the '+' sign"). These strategies are problematic, not only because they defeat the stated purpose of word problems in the first place (to enhance student learning and demonstrate math's applicability to real-world problems), but also because they're not even particularly effective for many students, due to difficulties translating words to equations (Clement, 1982) and, relatedly, incomplete understanding of arithmetic and algebraic forms (Knuth et al., 2005, 2006; Koedinger & Nathan, 2004).

Experience with the generic writing of word problems and teachers' instruction to ignore their content leads to students' "suspension of sense-making", whereby solution procedures are devoid of any connection to the supposedly relevant real-world knowledge (Gellert & Jablonka, 2009; Verschaffel et al., 1994), and the prevailing attitude is that word problems are pointless. This current approach to word problems also fails to connect to the goal-directedness of students' natural learning processes.

While it may seem that mathematical word problems are merely a secondary extension of formal symbolic mathematics, the consideration of real-world applications actually occupies a central role in math education and understanding. Although mathematics is essentially of an abstract, domain-general nature, its foundations are grounded in practical, domain-specific tasks, and the conceptual development of mathematical formalisms is paradigmatically goal-driven (Kieran, 1992; Sfard, 1995). Moreover, without its continued utility to our ability to understand and navigate the world, mathematics would be purely of academic interest (if it would have been so fully developed at all). In spite of all this, textbooks and teachers routinely relegate word problems to being supplementary to symbolic mathematics, tossing a few word problems in after repeated rote practice with symbols (Nathan & Koedinger, 2000a, 2000b; Nathan, Long, & Alibali, 2002; Sherman, Walkington, & Howell, 2016).

2.2.2 How Word Problem Contexts Can Actually Help Learning

In line with the FF view, many educators assume that solving a word problem involves converting words into an isomorphic arithmetic-algebraic symbolic representation, and then following the algorithms of computation to arrive at an answer within that representation (Koedinger & Nathan, 2004). The expectation is that the answer, though calculated within the formal language of symbolic mathematics, can be converted back into words, so that we only then (if at all) make contextual sense of the answer to the originally posed word problem. However, there is good empirical evidence to cast doubt on whether either of these mapping transformations is typically performed, or even whether such mapping is always beneficial for word problem solution.

There are multiple situations in which real-world scenarios actually improve problem solutions relative to isomorphic equations, such as operations like 100 divided by 4 in the context of exchanging money (Baranes, Perry, & Stigler, 1989). Koedinger and Nathan (2004) have shown evidence that for early algebra students, verbal problems can cue informal solution strategies that may be more reliable than their still-developing symbol comprehension and manipulation skills. This highlights both that the understanding of formal symbols takes time to develop and also that natural language and real-world scenarios can improve students' understanding. After all, early algebra students have had vastly more experience with spoken and written language than formal mathematical notation (Koedinger & Nathan, 2004). Moreover, their experience with language is varied and rich, whereas math formalisms are mostly encountered in specific, often impoverished, settings. By moving away from generic, unrealistic forms of word problems and reimagining them, we can leverage students' real-world knowledge and interest while highlighting the meaning and purpose of algebra.

Broad support for the role of semantic information in normatively syntactic domains also comes from cognitive psychology research on reasoning. Cheng and Holyoak’s (1985) classic studies on *pragmatic reasoning schemas* demonstrated that semantic content, in particular that which is not arbitrary but purposeful, evokes syntactic reasoning processes. When reasoners were asked to enforce a conditional rule couched in concrete details, subjects who had personal experience with a contextually similar rule reasoned more logically (following normative deductive reasoning with conditionals) than those who had no such experience. However, if reasoners were given the rationale (real-world purpose) for enforcement of the rule, subjects reasoned at the same highly-logical level regardless of their real-world experience. Importantly, subjects reasoned more normatively with abstract (meaningless) rules that were purposefully motivated (as a permission rule that must be enforced) than they did with arbitrary concrete rules without such motivation, although they reasoned *most* normatively with contexts that had both concrete details and purpose. Thus, it is not merely the incidental presence of concrete details *per se*, but rather, meaningful and purpose-driven concrete details, that best facilitate subjects’ logical reasoning by invoking relevant schemas (e.g., permission, obligation) that they have generalized across experiences with analogous contexts.

It’s clear from both cognitive psychology and prior work on math education that semantic information can affect the normatively syntactic processes of deductive reasoning. While semantic influences on syntactic processes may seem to some teachers like a reason *not* to use concrete examples early in teaching, such influences would actually facilitate learning if the semantic information was chosen such that it *aids* students in their thinking. In a sense, semantic information could be used early on to invoke helpful schemas. In my approach, such biasing is used in the examples to allow students to reason as intuitively as possible, so that those intuitions can then be formalized as they are mapped onto symbolic representations with syntactic transformations that parallel the students' contextual thought.

2.3 Problem Solving: Means-Ends Analysis, Schemas, & Contrasts

In our consideration of the factors influencing students' learning how to solve certain types of problems, it is also important to review research from the problem solving literature. At their core, the vast majority of mathematical problems require the solver to transform an initial state to a goal state using only the legal rules for the domain (Greeno, 1978). For example, let's consider a typical algebra word problem. The solver is given some mathematical information about known quantities, unknown(s), and how they are mathematically related; this information (the initial state) may be provided in symbolic, verbal, and/or pictorial representations. There is also a specific unknown quantity whose value is to be found (the goal state). The solver moves from the initial state to the goal state using whatever means of transformation (operators) that are valid in the domain, in this case, arithmetic and algebraic procedures. But how are the appropriate operators chosen?

One way to achieve a goal would be to use means-ends analysis (Newell, Shaw, & Simon, 1958), a strategy in which a solver identifies the highest priority difference between the current state and the goal, and then selects as the next operator the rule that most reduces that difference; if such an operator is inapplicable to the current state, a new subgoal state is generated wherein that operator would be applicable, and the solver works toward the subgoal. Means-ends analysis is a *general* strategy because it does not depend use any domain-specific information to guide its behavior. Essentially, means-ends analysis involves working backward from the goal to set subgoals, and then proceeding forward toward the next subgoal from the current state (Simon & Simon, 1978), which requires a lot of cognitive resources to carry out (Sweller, 1988).

An alternative way to achieve goal-based problem solving is to utilize prior experience with similar problem types, especially by use of schemas. A schematic representation of a problem's structure can guide the solver to use a sequence of operators generalized from the previous similar problems (Sweller, 1988). Schematic representations are inaccessible to early learners, who have had minimal prior exposure and are more easily distracted by problems' irrelevant surface details. And simply being previously exposed to similar problems is not sufficient for schematic reasoning—schemas are only useful when they are activated by matching a current problem's internal structure to previous problems with possibly distinct surface features (Cheng & Holyoak, 1985; Gick & Holyoak, 1980, 1983). Experts in a domain have extracted schematic representation across exposures to many different exemplars, which allows them to recognize problem types more quickly and encode them into memory more efficiently (Cooper & Sweller, 1987), reducing their working memory load and improving their performance.

For example, novice physics students categorize physics problems based on surface features (e.g., the presence of an inclined plane), whereas experts group them based on their internal structure, readily leading to more efficient solution strategies (Chi, Glaser, & Rees, 1982). In another line of work, Simon and Simon (1978) demonstrated that novice physics students used means-ends analysis by first working backward from the goal state to define subgoals along the way, while experts skipped the backward phase altogether and proceeded forward from the initial state directly to the goal.

So while experienced algebra students may largely agree on the categorization of algebra word problems (Hinsley, Hayes, & Simon, 1977), such schematic representations are unavailable to early learners. Sweller (1988) argues that because means-ends analysis drains cognitive resources, this can prevent early learners from extracting the internal structure to build schematic representations for future use.

2.3.1 Helping Students Build Schemas

Much research in educational psychology has focused on the role of learning-by-doing and generalizing across concrete examples to form abstract ideas. It seems intuitive that the meaning of a mathematical concept is more easily learned across varied practical examples. It is a lot easier, at least at first, to think about a variable when it represents a concrete quantity (e.g., the amount of money a friend owes me), compared to an abstract symbol. As Nathan puts it, "Concrete entities are meaningful to learners early on and so provide accessible entry points" (2012, p. 139). Romberg (2001) also stresses giving repeated, varied, concrete examples, to allow students to generalize knowledge from the patterns they observe across domains.

Sfard (1991) presents a very detailed theory of how conceptual knowledge is built in math learning. She offers the mantra "operational before structural" as a prescription for teaching mathematical concepts (using her duality terms). She argues that the precedence of the operational (procedures) over structural (abstract objects) is a Piagetian invariant, a constant observation regardless of the specifics of each example; simply put, abstract knowledge is generalized from procedural experiences. Piaget (1970) himself conjectured that mathematical abstraction is borne from generalizing actions, not the objects upon which the actions are carried out.

Sfard's (1991) recommendation is to embrace the primacy of the operational by engaging students with more meaningful activities that encourage their exploration and allow them to abstract conceptual knowledge. Students will induce schemas for how to solve similar types of problems. However, she argues that schematic representations built from operational experience alone are 'unstructured', merely a sequence of procedures to carry out without any connection to concepts. In her view, students must then 'reify' the operational into an 'object' conception, which imposes the structure on schematic procedure sequences just as a roadmap imposes structure on driving routes.

My own approach is to encourage students to constructively struggle through new problem types (e.g., system of equations) in a meaningful context, and with contrast comparisons, lead them to intuit the purpose of a new formalism. Students are then shown word equations (a 'bridge' representation to letter variable equations) and how they can be used to solve the system via the elimination method. But rather than just simply then giving students rote procedural practice, they are given a new type of system of equations problem where elimination won't readily work because the second equation is relational; after students figure out or are shown how variables can also help them in this second problem by substitution, they have seen word variables used in two distinct ways, and in the case of substitution, it should begin to be clear that substituting one variable for its equivalent in terms of the other requires that the word variables be numbers in their own right to be equal. This sequence hopefully encourages students to reify what begins as operational into an object in its own right.

2.3.2 Causal Contrasts

As discussed above (1.4.3.1), the contrast comparisons used in my approach follow logic similar to that of the recent 'causal contrast' research (Walker, Cheng, & Stigler, 2014). In Walker et al.'s research, causal contrasts were shown to encourage students to implicitly represent their *failure* to solve a given problem type (e.g., '$x^2 - 13x - 30 = 0$') as an *effect* that they are motivated to prevent. The presence of this effect for the given problem is first contrasted with subsequent problems that are perceptually similar but differ crucially in their mathematical structure (e.g., '$x - 30 = 0$' and '$x^2 - 30 = 0$'); these latter problems are implicitly represented as cases in which the failure effect is absent. By contrasting these two types of cases, students invoke their intuitive causal reasoning (Cheng, 1997) to induce that the presence of both 'x^2' and 'x' terms in a single equation *causes* their failure effect by preventing the familiar procedural isolation of x to find x's value. Thus, the removal of this cause naturally becomes students' new subgoal, which through additional causal contrasts, they come to learn is achieved by factoring.

Next, the factored form of the initial problem, $(x - 15)(x + 2) = 0$, is contrasted with equations that are perceptually different but structurally similar (e.g., '$ab = 0$'), which the zero-product property is more apparent (a and/or b must be equal to 0). From these contrasts, students are able to infer what causes them to fail to recognize the product in the factored form and thus be unable to see the relevance of the zero product property. Students in Walker et al.'s (2014) study were able to learn the purposes of each step in the solution procedure, identify perceptual difficulties blocking them from recognizing different mathematical representations of the same mathematical concept, as well as the goal-relevant relationships between its constituent concepts: factoring and the zero-product property, with the former enabling the latter. My approach uses contrast comparisons to evoke the same kind of implicit causal learning in Walker et al. (2014) to teach the fundamental concept of a variable and its purpose.

2.4 How Formalisms Have Been Taught Differently: Generalization-Based Progressive Formalization (G-PF)

Both Romberg (2001) and Nathan (2012) argue against FF in favor of what they term 'progressive formalization', a pedagogy by which students first experience concepts in more grounded, familiar contexts, so that as abstract symbols are introduced into the learning process, they are founded atop more meaningful knowledge. Their views, which embody a similar framework but differ somewhat from each other in areas of emphasis and implementation, each inform my own approach to change the way that students are introduced to the formalisms of algebraic variables. However, because they each employ generalization as the primary mechanism, I refer to their approaches as generalization-based progressive formalization (G-PF), in contrast to my purpose-driven progressive formalization (P-PF).

G-PF allows educators to place an early emphasis on the contextual meaning of a concept, and through varied concrete examples, a sense of generality. As learners progress, formalisms such as symbolic notation can be gradually introduced in a way that aligns with the learner's goals, and they can be mapped onto the learner's meaningful understanding of the domain. This not only makes the abstractness of formalisms more palatable to the learner, but it also highlights the meaning and the purpose of the formalisms themselves: as domain-general, computationally effective tools for quantitative modeling.

Nathan (2008, 2012) endorses an 'embodied cognition' approach that strongly emphasizes symbol-grounding to gradually construct meaning for new formalisms: "It is through *grounded* relationships that connect to our direct physical and perceptual *experiences* (or through chains of relations that connect to things that connect to our experiences) that [mathematical] formal entities attain their *meaning*" (Nathan, 2012, p. 139, emphasis mine). Although acknowledging that overly rich levels of concrete details may sometimes distract from abstract transfer (DeLoache, 1995; Kaminsky, Sloutsky, & Heckler, 2008, 2009; Uttal, Liu, & DeLoache, 1999; Uttal, Scudder, & DeLoache, 1997), he argues that grounding early experiences in relevant contexts is essential to giving formalisms meaning, attracting students' interests, and facilitating transfer to novel concrete contexts. Nathan's approach posits that well-chosen concrete examples early on in instruction, followed by the gradual introduction of formalisms, will lead to better student learning outcomes—in essence, concrete problems are better entry points, and abstraction should grow from there. Advantages in learning and reasoning with concrete examples can be found throughout cognitive psychology and educational research (Cheng & Holyoak, 1985; Clark & Paivio, 1991; Goldstone & Son, 2005; Koedinger & Nathan, 2004), although the relationship between concrete and abstract is nuanced (Anderson et al., 1996; Kaminsky et al., 2008, 2009).

Nathan asserts that it is critical that "learners connect new knowledge to old and *meaningfully* relate specialized notational systems…to the objects and events in the world that they are intended to represent" (2012, p. 138, emphasis mine). He worries that while some students may indeed connect new knowledge to old, if they do so only within the confines of symbolic representations, their resulting knowledge base will be ungrounded and shallow (Harnad, 1990). He cites Searle's (1980) Chinese Room thought experiment as an analogue to the student mapping symbols and rules only onto other symbols and rules, never understanding their true meaning (similar to Herscovics & Linchevski, 1994).

Romberg's account of G-PF (2001) does discuss the notion that students should intuitively discover the purpose that drives the use of formal symbols. He articulates this framework as part of a summary of a 5th-8th grade mathematical curriculum (*Mathematics in Context*, or *MiC*) based in part on both mathematizing realistic contexts and gradually moving from simple notations to more abstract reasoning. The curriculum was developed by a collaboration between the National Center for Research in Mathematical Sciences Education (NCRSME) at University of Wisconsin–Madison and the Freudenthal Institute at the University of Utrecht. Following Freudenthal's views (1987), *MiC* grounds students' early activities in solving contextual problems, emphasizing the use of symbolic notation to 'mathematize' contexts as a way to make quantitative sense out of them. Additionally, and more uniquely, *MiC* ordered the sequence of activities so that students could play a role in creating and using informal symbol systems, gradually refining them toward their formal end-results.

Romberg states that the rationale for this approach is that "…rather than starting with the presentation of formal terms, signs, symbols, and rules and expecting students to use these to solve problems (too commonly done in mathematics classes), activities should *lead students to the need for the formal semiotics* of mathematics" (Romberg, 2001, p. 3, emphasis mine). His language indicates a clear desire to allow students to develop a sense of symbols' *purpose*, and by use of the term "semiotics", places an emphasis on symbols' *meaning* as well. And *MiC* does appear to be a valiant effort toward gradually teaching meaning and purpose through a G-PF framework. However, given its emphasis on generalization, Romberg's (2001) approach does not appear to support students' inferences about purpose.

2.4.1 Why We Should Use Contrast Comparisons, Not Just Schematic

Romberg's description (2001) of *MiC*'s approach to gradual generalizing highlights comparisons across different contexts meant to allow students to induce *schemas* across problem domains (and the contexts vary quite a bit in relatability). While schematization is of course important, *MiC* does not appear to also implement *contrast* comparisons as a means to allow students to intuit the purpose of formalisms. A similar pattern can be observed in Nathan's (2012) discussion of previous research.

Much of the research on problem solving, analogical transfer, and math learning has been focused on schema induction across problems with different contexts (i.e., abstraction, generalization, transfer), and it's not surprising that both Nathan (2012) and Romberg (2001) prioritize schematization from experiential learning with concrete examples. Indeed, schematization is a critical part of learning a formalism's generalizability to novel contexts and purely symbolic problems. However, schematization requires an abstraction of the internal structure common to each example, and a disregard for the irrelevant surface details that vary across contexts. Novice reasoners' ability to separate the internal from the external is notably weak (Chi et al., 1982; Gick & Holyoak, 1980, 1983). So, if students learn a richer representation of the initial problem's internal structure, we would naturally expect that their subsequent schematization would be more effective.

The distinction between schema comparisons (abstracting across similar cases) and contrast comparisons (distinguishing between states) is reflected more broadly in theories of learning. Early account of animal learning theorized that the number of CS-US pairings determined the strength of association (Hull, 1943). By controlling for CS-US association while varying the probability of the US occurring in the presence vs. absence of the CS, Rescorla

(1968) showed that this ‘contingency’ was the determining factor in producing learned behaviors. While he did not use the term, this implicit process is clearly a *contrast* between two probabilities. Likewise, modern accounts of causal reasoning (Cheng, 1997) posit an implicit process that contrasts estimates of the conditional probabilities of an effect given the candidate cause’s presence or absence. Associative and causal contrasts are essential to animal and human learning, and as such, we may expect that contrast comparisons are an evolutionarily old, implicit process (i.e., what may be called ‘System 1’ according to a dual-system model, Evans, 2003; Stanovich & West, 2000). Schematization, on the other hand, is an analogical process requiring greater working memory and cognitive effort (Sweller, 1988), and it may be a distinctly human, more explicit mode of thought (i.e., ‘System 2’).

There is good reason to think that contrast comparisons will not only improve students’ representation of problem structures, but even more importantly, they will allow students to better understand the purpose of using formalisms in explicit representations of the problem structures; this has the additional benefit of making clearer the mapping between mental models of problem structures and the formal symbols of written mathematical notation and procedures. First, contrasts between types of problems highlight their structural differences, which may lead to finer-grained representations. Second, contrasts between forms of representation encourage students to think about formalisms as goal-directed tools with their own benefits and drawbacks, which may lead to a deeper understanding of why letter variables in particular, and formalisms more broadly, are useful.

My own approach emphasizes not only meaningful contexts, but also contrast comparisons across problem structures and forms of representation, in order to allow students to intuit a sense of purpose for the eventual formalization and to induce a fuller schema of the problem structure. This process is done on a much smaller time-scale than the *MiC* curriculum (Romberg, 2001), which is not merely a choice of convenience but rather an important condition for students' contrasts and decomposition to be more effective. By progressing through multiple problem structures in a single lesson, students get a chance to see the usefulness of the formalism before it's introduced, and they're better able to compare the problem structures because they are temporally contiguous.

2.4.2 Research on Teaching Formalisms: FF and G-PF Examples

There are several instances of previous research worthy of discussion that can be characterized as promoting a somewhat more FF or G-PF view. It is worth discussing a few of these in turn. Like many supposed dichotomies, one can detect in the literature that over time the pendulum swings back and forth from emphases on concreteness to the abstract, and that the reality is likely to be more nuanced than either end of the continuum. A clear failure of an overly abstract FF approach, discussed by Nathan (2012), is the "New Math" movement, which promoted formal set theory descriptions even at lower levels of basic mathematics, leading to confused and unmotivated students and teachers alike (Kline, 1973).

However, the use of concrete materials is not without issues either. Uttal, Scudder, and DeLoache (1997) argue that *manipulatives* used in mathematics education (e.g., Dienes blocks) are nonetheless symbols in their own right, and proper learning from manipulatives requires the student's understanding that they bear a certain relationship to the mathematical formalisms they are meant to map onto. They posit that students sometimes struggle to transfer knowledge from manipulatives or other concrete models because they require "dual representation": they are concrete objects *and* they represent something else. However, a similar point could be made of abstract symbols that can be used to stand for real or imagined objects or quantities—in the case of an algebraic variable, x, the letter itself is a concrete object (though considerably less perceptually rich than a picture or physical object) *and* it represents something else (a number or set of numbers, which themselves may correspond to a quantity in a word problem or in real life).

Researchers promoting manipulatives or concrete problems may have overlooked the difficulty of mapping to the abstract, but researchers endorsing FF may likewise have overlooked the difficulty of interpreting formal symbols. We may conclude that the difficulty with both concrete and abstract representations is that they must be able to map onto each other; indeed, that is their purpose. What matters for designing educational materials is when and how each form should be used: which should come *first*, and how do we *map* it onto the other?

An important early example of a G-PF-type approach cited by Nathan (2012) is Koedinger & Anderson's (1998) cognitive tutor utilizing what they termed *inductive support*. In their program, students in the inductive support group were first asked to solve two different arithmetic variants of the same base word problem. For example, they were told that a plumbing service costs a service fee of \$35 plus an hourly rate of \$42 per hour. They were first asked how much would 3 hours of work cost, and then asked how much 4.5 hours would cost. These questions were termed "result-unknown" because the result of the arithmetic was the unknown, not a constituent element of the arithmetic (as is common in algebra). They were then asked to "Create a variable for the number of hours the company works. Then, write an expression for the number of dollars you must pay them." (Koedinger & Anderson, 1998, p. 165). By asking students to first answer concrete arithmetic questions and then generalize from their equations to the use of a variable, Koedinger and Anderson hoped to encourage students to see algebra as a *generalization* of arithmetic.

This inductive support method was compared to the more traditional (FF-like) approach, which they termed *normative-deductive*, in which students are encouraged to *first* translate the verbal description of the plumbing company's fee system into an expression with a variable for the number of hours, and *then* use the expression to solve each result-unknown query. The two strategies were compared by examining students' ability to write a correct algebraic expression, as well as answer a final "start-unknown" because it asked how many hours the company worked if the final bill was $140 (in other words, the result is known but the variable's value is unspecified). Indeed, students in the G-PF-like inductive support condition outperformed those in the FF-like normative-deductive group.

Importantly, the subjects in Koedinger and Anderson's (1998) study were high school students who had *already* completed an Algebra course. In other words, these students had already learned a lot about variables, presumably through an FF-like paradigm. This was not a flaw in their design, because Koedinger and Anderson's research team was invested in cognitive *tutors*, not cognitive *teachers*. In a sense, Koedinger and Anderson's cognitive tutor acts as a 'debugger', identifying students' misconceptions based on their incorrect answers, and offering appropriate follow-up problems to aid students' learning from their mistakes. This is a worthwhile approach for tutoring, but it makes Koedinger and Anderson's (1998) manipulation a relatively weak test of what Nathan (2012) prescribes for researching G-PF approaches.

The biggest problem with Koedinger & Anderson's (1998) cognitive tutor is that they fail to encourage any sense of purpose for using a variable. Like *MiC* (Romberg, 2001), the cognitive tutor emphasizes schematic abstraction across examples. By asking students in the inductive support group to first solve two result-unknown problems and then write an algebraic expression to represent the company's rates, they are prompting students to generalize a schema for how the company's rates are determined and to represent this with algebraic notation. This highlights the role of algebra as generalized arithmetic, which fits their stated goal.

While schematic abstraction is important, the concrete example used by the cognitive tutor does not seem to convey the real purpose of the formalism: it's unclear from the students' perspective what such schematic abstraction is actually useful for in this example. Students may wonder, for example, why they should care whether a plumber's rates can be represented as '$42h + 35$', since in practice, if they were paying a plumber, there would be a specific number of hours the plumbers had worked, so arithmetic alone would be sufficient to multiply that number by 42 and add 35; it's simply not necessary to create a new symbol to stand for an unknown that goes through two simple arithmetic steps. Inductive support, in moving from concrete examples to an abstract representation, is 'progressive'. However, the schematic abstraction by itself does not convey why a variable is useful; such abstraction must be preceded and supported by intuitive reasoning that helps students to understand the concept and its purpose. In my materials, variable representations are not used for such one-unknown problems because they aren't strictly necessary; instead, students move from one-unknown problems to two-unknown systems of equations, at which point having a way to represent both unknowns becomes essential.

Another example of a G-PF-type approach is Goldstone and Son's (2005) "concreteness fading", in which the researchers split subjects into four groups: those who received idealized (abstract) examples throughout, those who received concrete examples throughout, those who moved from the abstract to concrete ("concreteness introduction"), and those who moved from the concrete to the abstract ("concreteness fading"). The task was for the students to learn competitive specialization, a concept in which individual agents self-organize in a way that maximizes efficiency across the group (e.g., primary visual cortex neurons' spatial orientation, flies' allocation across a resource supply). Subjects were allowed to explore graphical computer simulation of ants moving around a 2D environment to access food 'painted' on screen by the subject. Subjects could also manipulate parameters with slider bars and observe the resulting changes. The idealized version used small black dots to represent the ants and green patches to represent food, all on a white background; the concrete version showed somewhat simplistic line drawings of ants and apples for the food.

Their results showed that subjects made fewer errors overall when they saw both idealized and concrete visualizations, indicating that the variability was helpful; importantly, the best performance was found in subjects who moved from concrete to abstract (concreteness fading). They concluded that this was likely due to the fact that the concrete representations early on allowed learners to connect the visualization better to what it is meant to represent, but then the subsequent idealization helped to promote transfer by abstracting away the concept's structure from the distracting details. So, just as with PF more broadly, concreteness facilitates early learning but it must then be mapped onto more abstract representations. Of course, Goldstone and Son's materials are different from those traditionally used to teach math problems (although there is certainly math underlying the visualized systems), but their results are perhaps even more relevant to my own approach than they are to many other math learning studies because my multimedia materials use a lot of visualization.

In a broadly analogous approach, Kotovsky and Gentner's (1996) research on "progressive alignment" has demonstrated that children's analogical transfer of a relation across a source and target that differ in multiple dimensions is aided by gradually decreasing the salient similarities between the analogues. For example, if children are first given a relational analogy task in which they must choose between two probes that differ from the source only along one dimension (e.g., color) and then the next two probes differ from the source in two dimensions (e.g., color and shape), children will subsequently be much more likely to pick the correct probe on a final task in which the probes differ from the source in three dimensions (e.g., color, shape, and size). Thus, if salient features are only varied gradually across examples, the resulting schematic representations will allow for farther transfer to more disparate contexts.

Looking across all of these prior research studies, it seems that the simple addition of concrete examples early in learning is not sufficiently effective, nor is introducing representations progressively if they are not motivated in the first place. However, many students clearly struggle to understand abstract concepts and formal representations when they are taught in an FF manner.

For my own approach, the use of relatable concrete domains for early example problems, complete with contextual item images and illustrated characters that interact with each other, is expected to facilitate student interest early on. As students progress through the materials, techniques like color-coding and animations are used to convey which contextual items map onto which representations. For example, in one story, the narrator's friend purchases a donut and a cupcake, which appear red and blue, respectively, from their frosting. As word equations are introduced to represent their prices, money-specific words are colored green while the food item names are colored to match their images. Of course, the words are less concrete than the images, but the color match encourages students to correctly map between them. Then as students later learn to abbreviate word equations using letter variables, the connection between the letter and the item's price it represents becomes even less direct. In a sense, this progressive change in representation is its own form of concreteness fading.

But my materials differ significantly from these previous approaches in that they actively encourage students to intuit the purpose of developing more abstract representations before they are introduced. To achieve this, students are allowed to constructively struggle through solving difficult new problems so that form an intuitive understanding of how a new type of formal representation would be useful. By asking them to make contrast comparisons across cases where they can and cannot solve the problem and across different types of representation, they are able to infer each constituent piece of the concept of a variable that has been decomposed in the materials' design; for example, a variable is a useful placeholder for an unknown when there are many steps to carry out to find its value, and letters are an efficient way to represent variables formally because they can arbitrarily stand for any quantity and they are considerably more compact than using words or images. And as representations are progressively introduced and refined, students can better map their contextual intuitions onto these new abstract forms.

3. Written Materials: Design, Testing, and Lessons Learned

The materials for this project began in written form. This medium was initially chosen over digital formats so that any school, regardless of their technological capabilities, could participate in the research and/or receive the eventual finished product. These materials were extensively piloted one-on-one and in small groups, but in scaling it to the first classroom-level experiment, it became clear that the format of the materials needed to be changed to be multimedia (see Chapter 4). The development and testing of these materials, however, significantly factors into their multimedia adaption. Moreover, this process illustrates an important lesson about the creation of educational materials using cognitive psychology principles: that the realities of classroom experience and educational environments vary greatly, and in order to create effective educational materials, it is important to learn from these experiences and adapt materials accordingly.

3.1 The Initial Design

Development of the written materials began with the basic idea of having pre-Algebra students attempt to solve a series of PRM word problems, logically sequenced according to the historical development of algebra, with conditional hints and advanced hints as needed, and open-response prompts to contrast aspects of each problem. The words ‘algebra’ or ‘variable’ would never be used, students would learn after each problem how to solve it using *intuitive* and *contextual* thinking, they would be gradually introduced to “word equation” form, and they would end with a lesson on how to use “letter equations” (i.e., algebra).

I used as a starting point, a word problem from Kraft (2013), which already contained many meaningful and purpose-driven details:

> "I went to Dunkin' Donuts to get some of my colleagues some doughnuts and muffins. I bought five doughnuts and two muffins for $4.95. While this made some of my colleagues really happy, I noticed that a couple of them gave me dirty looks in the hallway as they didn't get anything. So the next day I had to buy seven doughnuts and two muffins for $6.25. Later, one of the teachers insists that he pay me for the muffin that I got him. I realize that I had no idea how much a muffin cost. Can you figure it out?" (Kraft, 2013)

In this word problem, Kraft provided several details that fit nicely into my PRM framework; these PRM details facilitate students' contextual thinking, which in turn may reduce their working memory and other cognitive resources' loads, allowing them to think more deeply about the problem's structure.

- *Plausibility*: It's reasonable to assume that someone would buy different donuts and muffins on consecutive days, and that they would remember how many of each that they bought as well as the total amount, without knowing how much each individual item cost. There's a plausible reason for the change in quantities between the two days (colleagues' unhappiness at their exclusion).
- *Relatability*: Students should have a decent amount of experience with buying food items, and probably some experience with leaving out a friend (accidentally or not) from some snack or activity. They probably even have some experience with having a friend pay them back for something.

- *Motivation*: It's very clear why the narrator in the story (the teacher) needs to know how much a muffin cost because another teacher wants to pay him back. Leaving out just that single sentence, "Later, one of the teachers insists that he pay me for the muffin that I got him," would remove this motivation (why would the narrator really care how much a muffin cost?). It's important to note that the motivating sentence is entirely unnecessary from a mathematical point of view—it adds nothing to the problem structure, and its omission does not change the solvability of the problem at all. But its inclusion does add a very human element: purpose.

This was an excellent PRM problem to start with. I then worked through a couple improvements according to my PRM framework: in terms of plausibility, I clarified why the narrator remembers the total cost but not the per-item cost (the student lost the receipt but remembers the total cost, or still has the receipt but it was not itemized); for relatability, I made the narrator and other characters students (rather than teachers). Working both forwards and backwards, I adapted Kraft's general problem context to fit the sequence of four problems I developed.

Notably, Kraft (2013) presented this problem to his students before introducing any algebraic formalisms into the lesson. He then challenged them to solve it on their own (in line with constructive struggling). After a while, he demonstrated to them that the first day's transaction alone was insufficient to determine the cost of the muffin. Then, Kraft recounts, "I asked them to think about how the situation changed from one day to the next. It was at that point that the light-bulbs began to turn on. Students started to see that the price went up by $1.30 and I only added two more doughnuts" (Kraft, 2013). Notice how the hint's focus on making a contrast between the two days invoked their causal inference—purchasing two more doughnuts for the unhappy colleagues caused the price to increase $1.30; this is the only plausible explanation because all other factors in the price were held constant (given the very reasonable assumption the prices haven't changed overnight). From this, it's easy to see that two doughnuts must cost \$1.30, so one must cost \$0.65, and this amount can be applied to one of the transactions to subtract out the cost of the doughnuts in order to be left with only the cost of the muffins. Again, these features of Kraft's design fit well with my own thinking and goals. Kraft (2013, personal communication) wanted to introduce the elimination method in a new way, so he followed his own intuition in switching the order of teaching to use a word problem before teaching the formal procedure with algebraic equations. In essence, students could intuitively perform the elimination method by focusing on the *change* between the two orders without necessarily converting the problem to formal symbols first.

3.2 Materials

The written materials were developed to have three experimental factors that could be manipulated between conditions: Progressive Formalization (PF) vs. Formalisms First (FF), Active Learning (AL) vs. Passive Learning (PL), and PRM Context (PRM) vs. Generic Context (GC). The initial development for these materials focused primarily on an Experimental condition with PF, AL, and PRM and a "Traditional Materials" Control condition with FF, PL, and GC. The Control condition is termed "Traditional Materials" because it is an attempt to represent how algebra is typically taught, using FF (Nathan, 2012), GC (Gerofsky, 1996, 1999; Nesher, 1980), and PL (Prince, 2004); in the latter case, some AL is used more often in education, but the specific types in these materials' AL manipulations (discussed below) are much less common. Each factor can also be manipulated separately to allow for other conditions, such as Context Control (PF, AL, and GC) and Formalisms Control (FF, AL, and PRM), which would allow for factors to be assessed independently.

For all conditions, the materials consist of two separate packs of worksheets, each designed to be given on a separate day. Broadly speaking, the first worksheet contains the bulk of the learning and the second worksheet contains the bulk of the assessment (after a delay). Below, the Experimental condition and "Traditional Materials" Control condition are the primary conditions discussed, unless otherwise noted. All conditions begin with an identical Instructions page, which informs the student that this is not a test and they are not being graded, that they will read some stories about a student who buys food and then be asked to answer some questions, that the questions may be difficult and they are not expected to know how to solve every problem, and to please try their best.

3.2.1 Worksheet

The first session's worksheet consists primarily of Stories (word problems) for subjects to solve and Lessons for how to solve them; the main difference between the Experimental and Control conditions is *when* and *how* these materials appear. In the Experimental condition, students begin each section with a contextual Story (PRM) and are encouraged to solve it on their own, even though they are unlikely to have encountered some of the problems (in particular, Stories 2-4, which are system of equations problems). By allowing students to think through the problem on their own first, this supports constructive struggling (AL) and provides students the opportunity to form their own intuition for the formalisms to be introduced later (in line with PF). Additionally, there is conditional branching for some of these Stories: students who get the Story incorrect at first are given Hints (AL) and asked to try again, those who get it correct initially but use a "Guess and Check" method are then given a more computationally complex version (making that method harder to use), and finally, all students get the Explanation page, which helps walk them through the problem using an intuitive, informal solution strategy (PF). At the end of the worksheet, students are given some Lesson pages which introduce variables and demonstrate how to use them to carry out the elimination method.

In the Control condition, each Story is preceded by a Lesson that demonstrates how to solve that type of a Story (but with different numbers) using variables and an appropriate algebraic solution strategy (FF). Then students are given the Story, which is mathematically identical to the corresponding Story in the Experimental condition but does not have the extra (mathematically unnecessary) contextual details (GC); students are then asked to solve the Story using the strategy they just learned (PL). There are no Hints for the Control condition (PL), but there is some conditional branching for the more computationally complex version (if they use "Guess and Check" the first time).

3.2.1.1 The Stories

There are four main Stories in the Experimental condition, although the second story has two parts (one unsolvable, one solvable), and the second and third stories also have numerically more complex variants (labeled as Story 2B and Story 3B). Each of the Story variants is explained below.

While the Stories do vary somewhat from each other (in terms of location, food items, story details, and motivation for finding the unknown), they follow a basic pattern of knowing how many of each food item were ordered and what the total cost was, and then wanting to know how much a single food item cost. I attempted to vary enough details across stories to prevent

Table 1

Written Materials' Story Structures

Story	Story Label	Problem Structure & Description	Food Item(s)	*Implied* Form	Desired Unknown
1[a]	One-Unknown	1 unknown, 1 equation	Tacos	$2t = \$4$	Price of 1 taco (t)
2.1[a]	Unsolvable Two-Unknown	2 unknowns, 1 equation (Unsolvable)	Pizza & water	$2p + w = \$5$	Price of 1 slice of pizza (p)
2.2	Change	2 unknowns, 2 equations (1 unknown changes quantities)	Pizza & water	$2p + w = \$5$ $4p + w = \$9.50$	1. Price of 1 slice of pizza (p); 2. Price of 1 bottle of water (w)
2.2B[b]	Change complex	2 unknowns, 2 equations (1 unknown changes quantities)	Pizza & water	$2p + w = \$4.79$ $4p + w = \$8.63$	1. Price of 1 slice of pizza (p); 2. Price of 1 bottle of water (w)
3	Relational	2 unknowns, 2 equations (1 equation represents relationship)	Gatorade & Coke	$c + g = \$4.50$ $c = 2g$	1. Price of 1 Gatorade (g); 2. Price of 1 Coke (c)
3B[b]	Relational complex	2 unknowns, 2 equations (1 equation represents relationship)	Gatorade & Coke	$c + g = \$5.38$ $c = 2g + \$0.10$	1. Price of 1 Gatorade (g); 2. Price of 1 Coke (c)
4	Double change	2 variables, 2 equations (2 unknowns change quantities)	Hot dog & juice	$2h + j = \$10$ $5h + 2j = \$23.75$	1. Price of 1 hot dog (h); 2. Price of 1 juice (j)

[a] Stories 1 and 2.1 are not given to subjects in the Control condition. [b] Given conditional on "Guess & Check" method.

redundancy, but kept much of the structure the same to highlight similarities. In the Experimental condition, motivations are also given for all changes in food items' quantities and for why the narrator wants to know an item's cost (PRM).

The mathematical structure of each problem type (elaborated in Table 1) was chosen based on a consideration of historical and logical progressions (see Sections 1.4.2 and 2.1.4 above). The specific values used in the Stories have been adjusted based on feedback from piloting.

Table 1 outlines the four main stories and their variants. Students in all conditions are required to complete (i.e., attempt an answer) the Change (2.2), Relational (3), and Double Change (4) problems. Students in the Experimental condition additionally receive the One-Unknown (1) and Unsolvable Two-Variable (2.1) problems.

Conditional branching is used in the Experimental condition for the Change, Relational, and Double Change problems. For these Stories, if the student fails to produce the correct answer initially, they are given a relatively vague suggestion on a subsequent Hint page (see 3.2.1.2), an AL manipulation. Then all students (regardless of whether they received the Hint) are given an Explanation page (see 3.2.1.3), which expands on the Hint more explicitly and, when helpful, introduces new representations (e.g., word equations in the Change problem).

Additional branching is also implemented in both Experimental and Control conditions based on students' solution strategy. For the Change (2.2) and Relational (3) problems, computationally more complex variants (Change complex, 2.2B, and Relational complex, 3B) are given conditionally to subjects who correctly solve the original Change or Relational problem by using a "Guess & Check" strategy (e.g., simply guessing possible answers and then checking the math to see if they work) or who do not show any work (they may have simply performed "Guess & Check" with mental math). This variant is given in order to encourage them to develop a more general solution strategy (because "Guess & Check" is inefficient with complex numbers), allowing them to see under which conditions "Guess & Check" does or doesn't work well, which helps clarify the purpose of using a more general strategy.

By not giving all students the more complex variant at the beginning, the branching structure both decreases the computational demands early on (which is especially important for students who struggle) and allows students to formulate and use whatever solution strategy they want. Thus, this conditional branching strikes the best balance between the two goals of allowing students the freedom to think through the problem on their own without computational complications and also encouraging students who successfully use "Guess & Check" to then find a more general strategy.

Importantly, the students are not told in the Experimental condition what type of solution strategy to use—rather, it is intended that they will learn on their own, through constructive struggling, that "Guess & Check" does not scale well to complex problems. Additionally, these complex variants are included in the Control condition in case students do not use the strategy they just learned in their Lesson. However, it is expected that these variants will be much more effective with students in the Experimental condition, because they encourage students away from "Guess & Check" and toward a more general solution strategy. And in thinking about a more systematic approach, students may begin to feel the need for a more general representation, which in turn will provide intuitive support for new formalisms when they are introduced (PF) in Explanations or the final Lesson.

3.2.1.2 The Hints (Experimental Only)

The Hint pages are designed to provide minimal, but effective hints to students who request them, fail to produce an answer, and/or produce an incorrect answer for a Story. Because they are meant to aid students who do not yet know how to solve a problem (AL, constructive struggling), Hint pages are only used in the Experimental condition. The hints themselves are written so as to be as vague as possible, while still providing a nudge toward a desired solution strategy (e.g., the subtraction method).

For example, the Hint page for the Change problem (2.2) suggests, "Think about what changed in what I bought between the first day and the next day, and what changed in how much money I paid." This focus on the change in the problem (which was a helpful hint for Kraft's class, 2013) tacitly suggests subtraction. If a student notices that the only change in the order from the first day to the next was an increase of 2 slices of cheese pizza, and the total cost subsequently went up $4.50, seeing that the increase in cost is due to the two extra slices is in essence an implicit subtraction of each element of Day 1's order from the respective element of Day 2's order ($4p - 2p = 2p$; $w - w = 0$; $9.50 – $5 = $4.50). By framing the hint in the real-world context (the change is conceived as happening to the order and the total cost, not to formal symbols), students may be better able to apply their intuitive reasoning.

3.2.1.3 The Explanations and Lessons

For the Experimental condition, the Change, Relational, and Double Change problems all have Explanation pages, which are given at the end of each section, either after a correct initial answer or the Hint page. These Explanations make suggestions that point to a possible solution strategy (elimination for Change, substitution for Relational and Double Change) by asking "thinking" questions (which are typically more fleshed-out versions of the same Story's hint, AL), then re-phrasing the problem in terms of their answers to the thinking questions.

For example, the previously mentioned Hint for the Change Story "Think about what changed in what I bought between the first day and the next day, and what changed in how much money I paid," is further broken down in the Change Story Explanation by asking the thinking questions: "What changed in the number of items that I ordered from Day 1 to Day 2?" and "How did the total cost change from Day 1 to Day 2?", and finally, "Based on your answers to the two thinking questions above, how much money did 1 slice of cheese pizza cost?" Because Explanations maintain a thinking question format, they are not merely passive lessons (like those in the Control), but instead invite students to actively think about the structure of the problem (AL). The Change Story Explanation also introduces word equations as a formalism to represent unknown quantities mathematically.

When students in the Experimental condition have finished all of the Story problems and their Explanations, they are given a final Lesson that introduces the use of letters to abbreviate the words in the word equations from the Change problem. The Lesson then proceeds to show how to solve the Change problem formally using letter equation elimination, and this is related back to the intuitive Hints and Explanation thinking questions that encouraged an informal elimination-style solution (by focusing on the change between the two days' orders).

For the Control condition, all three Stories they are given (Change, Relational, and Double Change) are preceded by Lessons that demonstrate from the outset how to formally use letter equations to solve that upcoming Story type (using an isomorph of the story in the Lesson itself as an example). As such, students are taught from the beginning what variables are in formal terms (FF) and are shown procedurally how to use them (PL). The material covered by each of the Explanations and the final Lesson in the Experimental condition is distributed across the three pre-Story Lessons in the Control condition (so there is no need for a final Lesson at the end of Control). Essentially, the information presented is highly similar across conditions, but it differs in when (PF vs. FF) and how (AL vs. PL) it is presented.

Lessons differ from Explanations in a few critical ways (a fact highlighted by the change in the labels):

- In the Control condition, Lessons always come before Stories, whereas in the Experimental condition, Explanations always come after Stories, and the only Lesson takes place at the very end of the worksheet. This change in ordering reflects the temporal distinction between FF and PF.
- The "thinking" questions (AL) in the Explanations are converted into statements in the Lessons (PL) so that content is equated as much as possible. For example, in the Relational Story Explanation (for the Experimental condition), the student is asked, "Would it help if you could express an order in terms of only 1 unknown number (one item's cost), so you can find that number first?", whereas in the Lesson prior to the Relational Story (for the Control condition), this idea is expressed by a declaration, "It would help if we could express the order in terms of only 1 variable, so that we can find that number first."
- Lessons in the Control condition emphasize formal notation and procedures (FF) over sense-making approaches that encourage intuitive thinking (PF). For example, in the Relational Story Explanation (Experimental), the student is asked, "How many Gatorades could we buy for the same price as 1 Coke (looking at the Relationship word equation)?" This cues them to imagine the relation between the items' prices in the context of

purchasing items. In the Lesson for the Relational Story (Control), this question is not only converted to a statement (as in point #2 above), but the referents are changed from the 'real-world units' of Gatorades and Cokes to the abstract symbols of algebraic variables: "Looking at the Relationship equation, notice that c is equal to $2g$." These problems convey very similar information, in that they both cue the reasoner toward substitution by calling attention to the equivalence of two values (a prerequisite for substitution). But in the Experimental case, this equivalence is queried (not stated as fact), and its reference to the purchase of items grounds the equivalence in the context; this potentially makes it more intuitive to think about substitution, since one could imagine buying 2 Gatorades instead of 1 Coke and the price remaining unchanged. By contrast, talking about replacing c with $2g$ is much more abstract and less accessible.

These differences are an attempt to operationally distinguish between FF and PF approaches as well as the AL and PL approaches.

3.2.1.4 Progressive Arcs Throughout the Experimental Materials

The Experimental condition, embodying both PF and AL approaches, draws students along a sequence of problems and questions that form two major arcs through which major steps in formalization are made. In the first arc, students move from using descriptive sentences to word equations as they intuit the need for symbolically representing unknowns in more complex mathematical structures; and in the second arc, students learn the practical purpose for abbreviating word equations as letter equations.

Progressive Arc 1: Intuitive Purpose of a Word Variable as a Placeholder

In the Experimental condition, there is an arc of questions beginning with the One-Unknown problem (Story 1) and ending with the Explanation for the Change problem (2.2), which lead students toward intuiting the purpose of a variable as a placeholder for an unknown to be found. Students move from the easily solvable One-Unknown problem (a single arithmetic step) to the *Unsolvable* Two-Variable problem, and the only structural difference between the two is that the latter has an extra unknown; thus, implicitly students can infer by contrasting the two that this second variable made the problem unsolvable (the only other differences between the two problems are superficial).

Then in moving from the Unsolvable Two-Variable problem to the Change problem, they can see that not all two-variable problems are unsolvable. In a sense, the Unsolvable problem serves as a conceptual bridge between One-Unknown and Change, as a challenge to students to really think about a problem's structure, and as a lesson that some problems do not have enough information to be solvable. Ideally, the students also learn that unsolvable problems aren't uninformative (i.e., "not enough information" does not entail "no information"), as indicated by the repetition of the Two-Variable information in the Change problem (since Change is a continuation).

Finally, as the student completes the Change problem's Explanation sheet, they are introduced to word equations as a means to represent the unknown quantities in a mathematically specific way. Whether or not they previously solved the Change problem, by virtue of the fact that they thought through it on their own beforehand, students may have formed an intuitive understanding that a new formalism would be useful as a 'placeholder' for the unknown to be found. Thus, when such a formalism is introduced, they can easily connect it to its purpose for keeping track of unknowns across solutions with many steps. In essence, the arc provides a bridge between students' intuitions and the explicit representation of word equations, which motivates their introduction and use. Importantly, effective word equations are also semantically meaningful to the student, as it's easy to easy to understand what 'the price of 1 slice of cheese pizza' represents and how it relates to the problem context. Thus, at the end of this arc, students have a representational form that is both meaningful and serves a clear purpose to them.

This arc is specific to the Experimental condition (not Control) because it is part of the PF flow of materials and uses AL techniques: formalisms are progressively introduced as they are needed, and contrast comparisons focus students' attention on how and why variables may be used. Neither the One-Unknown or Unsolvable Two-Variable problem are given to the Control condition because their inclusion in the materials is to provide contrasts with the Change problem (and such contrasts are not given in Control). Moreover, in piloting the Control materials with One-Unknown included, students were notably confused why they were being taught to use a variable representation when a single-step arithmetic solution is easily accessible. While this discovery in piloting itself presents a microcosm of difficulties with the FF approach, it was prudent to exclude this problem since it is inessential to the Control and was known to cause confusion to students; had it instead been left in, the simple presence of early confusion could potentially explain away any between-condition differences.

Progressive Arc 2: Practical Purpose of Letter Equations as Abbreviations

Word variables, while serving a conceptual purpose and being contextually meaningful, have a serious practical drawback. A word variable like 'the price of 1 slice of cheese pizza' is a rather long symbol, which makes it tedious to use repeatedly and harder to perceptually match with common instances (e.g., it is perceptually very similar to 'the price of 1 slice of chicken pizza', which is a different unknown that is quite likely to have a different value). The reason for using *letters* instead of words and phrases is precisely because they are more compact, which makes them quicker to write and easier to perceptually match. These are not conceptual or abstract considerations, but rather, merely a matter of practical purpose at the level of representation.

The second arc thus leads students to anticipate this practicality of representation by asking them at the beginning of the Lesson how the word equations could be made shorter so that they are not so tedious. But the foundations for this arc are set even before this, as students practice writing word equations through the Change, Relational, and Double Change problems, an experience through which they may notice on their own that they are weary of writing out full word equations each time. Thus, the focus on abbreviating representations in this Lesson capitalizes on students' own experiences. And across these two arcs, students progressively move from concrete descriptions to more abstract letter symbols, highlighting the purpose of having such a symbol as well as why letters in particular serve that purpose well.

3.2.2 Delayed Post-Test

After a delay of approximately a week, students are assessed on their near- and far-transfer to novel problems on a post-test. Importantly, all conditions receive identical post-tests so they can be directly compared. For all post-test problems, students may be assessed not just for the correctness for their response but also for their solution strategy (e.g., did they use variables, did they use subtraction or substitution, etc.) The post-test is labeled in the experiment as 'Worksheet 2' in order to mitigate testing anxiety and stereotype threat associated with math tests, but it is referred to simply as 'post-test' below.

The first problem in the post-test is the “Final Story”, which is an isomorph of the Change Story (2.2) from the (learning) Worksheet. The purpose of the Final Story is to assess students’ retention of the learned solution strategy (elimination) for Change-type problems and transfer it to a new (but similar) context with different numbers and item types.

Next, students are given 11 ‘conceptual’ items, which are intended to measure students’ understanding of conceptual aspects of arithmetic and algebraic formalisms, such as the interpretation of the equals sign, that a given variable maintains its same value throughout a problem, that variables can be manipulated without knowing their value, that some variables represent exact values and others general quantities, etc. Most of these items are discussed in greater detail in section 4.6.3.

At the end of the post-test is a survey that was developed in order to measure students’ attitudes and beliefs about math so that we can characterize the student population and also look for possible differences across conditions. The survey consists of 25 closed items that are rated on a four-choice Likert-scale (“strongly disagree”, “somewhat disagree”, “somewhat agree”, “strongly agree”), an open-ended question about possible uses for the math in the worksheet, and three Yes/No questions to assess students’ previous exposure to Algebra. Fourteen of the items were collected from various math self-efficacy and attitude scales (Fennema & Sherman, 1976; Shafer, Wagner, & Davis, 1997; Schoenfeld, 1989; Tapia, 1996), with some editing for consistency and clarity. In addition, nine new items were written specifically for this project. The first 21 items are intended to measure four constructs: the perceived usefulness of math, the student’s self-confidence in math, the belief in solution flexibility in math, and the importance of word problem context. The last four items are included merely to check students’ interest and engagement in the materials.

3.3 Piloting of Written Materials: One-on-One and Small Groups

The written materials were piloted in four different phases. All pilot subjects were given information about the research and had to provide signed parent permission and their own assent to participate in the study. They were all compensated with $5 in Target gift cards for each half hour they participated in (different pilot phases lasted different amounts of time).

The first phase of piloting (Pilot Phase 1) was performed at the same after-school program in Los Angeles, California. This program (referred to here as Pilot Site 1) serves over 4000 students from 1st through 12th grade, in a racially diverse community. Pilot Site 1 reports that 48% of attendees are from single-parent households and 56% qualify for free or reduced-price lunch at school.

During the first phase of piloting, 10 youths from 6th to 8th grade (age range: 11–15 years old) participated in small groups and one-on-one sessions. There were a total of two groups of three students at a time, one group of two students, and two one-on-one sessions. Six of these students completed the initial version of what was then called the 'Purposeful' condition, which embodied elements of PF, AL, and PRM, but did not contain any final Lessons. Four students completed the initial version of the 'Generic' condition, which only differed in context (GC). From Pilot Phase 1, it became clear that many students were able to successfully use a "Guess & Check" method to find the correct answer without engaging in the intended intuitive algebraic thought. Additionally, these students pointed out some confessions about the writing.

Nine of these ten Pilot Phase 1 students were revisited a week later with an updated version that instead used more computationally complex math, in order to see if this encouraged more algebraic thinking. Additionally, the decision was also made to make the items priced in Japanese Yen (e.g., ¥ 983) in order to prevent students from using a plausible real-life value as a starting point for "Guess & Check". While this worked for some students, too many were confused by the concept of a foreign currency, and some students had too much difficulty with the more complex math, which required longer arithmetic algorithms that are error-prone and cognitive draining. After this second visit in Pilot Phase 1, it became clear that a balance was needed between the easier and harder versions of the problems, and Yen would never be used again due to the obviously distracting confusion it created.

The second phase of piloting (Pilot Phase 2) was held again at Pilot Site 1, this time as five one-on-one sessions with new students (from the same age and grade range), using only the revamped Purposeful condition. Based on feedback from the two sessions in Pilot Phase 1, Pilot Phase 2's materials were heavily tweaked to create the conditional branching for a more computationally complex problem only if the student used a "Guess & Check" strategy.

The third phase of piloting (Pilot Phase 3) was held in a new location, Pilot Site 2, the private residence of a teacher in a considerably more affluent community in Newport Beach, California. Students were recruited through the teacher from a nearby school (the teacher's residence was used because it was summer break). The choice to pilot in a very different community allowed for the assessment of the materials working as well for students who likely attended a much more well-funded school that may have different educational practices. Pilot Phase 3 tested five more students, 6th to 7th grade (11–12 years old). This version tested out some new concepts that ultimately were abandoned, such as attempting to have a problem without specifying any numbers at all ("all I know is I bought two tacos for some amount of money that was listed on the receipt I lost"), and getting students to express their answer using verbal math ("what operation could I use to find the cost of one taco?"); something similar to this is planned for future extensions of the Multimedia materials (see end of Chapter 5), but it was discarded for the present simply because it is too abstract to use in the first session's materials.

The fourth and final phase of piloting (Pilot Phase 4) was held once again at Pilot Site 1, testing three more 5th and 6th grade students (10–11 years old), all using the Control condition. The pre-Story Lessons for the Control condition had been adapted from the Lesson at the end of the Experimental condition. However, the Lesson before the One Variable problem confused students because it introduced a variable to solve a single-step problem that was easily solved using straightforward arithmetic. Thus, the One Variable problem was dropped from the Control condition (although it was reworked into an improved example in the multimedia materials, see 4.3).

3.4 Testing of Written Materials at the Classroom Level

3.4.1 Test Site and Population

The written materials were tested at the classroom level at a public middle school (serving 6th to 8th grade) in Santa Ana, California. According to data from the National Center for Education Statistics, this school has approximately 1400 students total, with 98% of students reporting Hispanic as their ethnicity, 86% qualifying for free lunch, and 10% for reduced-price lunch. The study was approved by the district, principal, and teacher prior to student contact.

Two classes participated in the study. Each class was a supplementary math class called 'Support', which was taken electively by students struggling with their primary math class. Both support classes were taught by the same teacher, and one consisted of all 6th graders and the other all 7th graders. While there was some variability within each support class in terms of which primary math class they were enrolled in, all subjects had not yet taken Algebra.

All students in both classes were invited to participate, receiving a packet to take home with their family that included a recruitment letter, parent permission form, and student assent form. As per district policy, Spanish translations of the recruitment letter and parent permission form were provided for families with parents or guardians who primarily speak Spanish. Only students with signed parent permission forms and signed student assent forms were allowed to participate. Out of 20 students in the 6th grade support class, 15 submitted the appropriate documents to participate; out of 23 students in the 7th grade support class, 16 did the same. Participating students' identities were protected by randomly assigning them a numerical code, which was used to match Worksheets. Students who did not participate were given a class worksheet from the teacher to work on instead.

3.4.2 Design

Because some pilot subjects took longer than an hour to finish Worksheet 1 but the test school's class schedule provided slightly less than an hour, the decision was made prior to the testing to allow students to resume Worksheet 1 where they left off on the second day, and then complete Worksheet 2. To help make it more likely that students could finish Worksheet 1 in the first session, the Double Change problem was removed from the materials since it is the hardest Story and is less essential to the manipulations than the other three.

Additionally, it was decided to test only the Experimental condition in this population, since it was the first classroom-level study, and this would give more information about the Experimental condition's efficacy, while future studies could randomly assign students to different conditions to compare Experimental to the various controls. While not a perfect design, these decisions were made for practical reasons in order to learn as much as possible about the Experimental condition from this testing.

3.4.3 Administration

The test was administered entirely by myself across two sessions a week apart. Due to the conditional branching nature of the materials, Worksheet 1 was delivered to each student in batches (within a given branch). Before they began, students were told to raise their hand when they had finished the current batch, and I would bring them the next batch. I used a seating chart of the class to keep track of each student's numerical code by their seat so that when they raised their hand, I could grab the appropriate packet of pre-numbered paper to their seat, where I discerned which branch to give them next based on their answer (and work, if they were correct).

Each student started with a three-page packet consisting of the Instructions, the One-Unknown Story, and the Unsolvable Two-Unknown. Even though there is no branching between the Unsolvable Two-Unknown and the Change problem, I still only brought them the Change problem after they finished the Unsolvable to ensure that they fully thought it through before being given the new information in the Change problem. When they raised their hand to signal they were finished with the Change problem, I came by and looked at their answer. If they did not answer "$2.25", I gave them the 2.2 Hint; if they did answer "$2.25" but either showed no work or otherwise didn't show where the number came from, I gave them Change Complex (the more computationally complex variant); if they answered "$2.25" and did show where the number came from (even if it was just $4.50/2), I handed them the Change Explanation. Subjects finishing the Change Hint or Complex were automatically given the Change

Explanation next. This same general process was repeated for the Relational Story. After all Stories were completed, students were given one Lesson page at a time (out of three), and finally, Worksheet 2: the post-test items, and then the survey.

3.4.4 Results and Lessons Learned

It was very apparent when examining the students' results that written materials did not scale well to the classroom level. While the Hints and Explanation pages had worked well in piloting (especially after refinement), many students tested in these classes continued to write the same (incorrect) answer throughout the Hints and Explanations. In at least one case writing the number in large print with multiple exclamation points on the Explanation page, as if they were exasperated by having to answer the same question multiple times. Very few students got any questions correct other than the One-Unknown, in contrast to the piloting results (where students did fairly well on later Stories). The patterns were so clear that it was obvious that performing data analysis would not be worth the effort; the written materials simply did not work for the classrooms. However, I learned many important lessons from this experience.

First, the reading demands of the written materials were far too much. While this was not so apparent in piloting, it may be that in the classroom environment, students just do not expect to read that much on their own to learn math. Regardless of whether there was a 'classroom effect', however, the amount of reading is likely to draw too much attention and cognitive resources away from students' ability to focus on the math. Another aspect of the writing that clearly did not scale well was the fact that the Hints were predicated on students' previous answer being incorrect; somehow that fact was not clearly conveyed to the students who continued to give the same incorrect answer for Hints and Explanations, even though this was not an issue in piloting.

Second, running around a classroom for an hour to perform conditional branching myself for upwards of 16 students at a time is extremely impractical. These written materials simply went beyond the limits of conditional branching that can be done on that scale. While I knew this would of course be harder to do than a small group, the complexity did not seem to scale linearly but exponentially. Clearly, this was not an intervention that would be useful for teachers even with a relatively small class to use, which was a problem for my goals.

Third, while this only applied to a small portion of students, it became clear at this testing phase that the Explanations were too redundant for students who had already answered a question correctly. Since the same Explanation page was used for all students (regardless of whether they got the question correct before or after a Hint), it just didn't flow equally well for each individual case. In order to fix this issue, even more conditional branching would need to be introduced, which was clearly a bad idea for written materials like these.

There was some good news, though. Fourteen out of 31 students completed the survey, although this sample is admittedly biased, by virtue of the fact that students had to finish all other pages before getting the survey, so those who may have struggled did not get to respond. However, from this biased sample, all 14 students indicated they either 'somewhat agreed' or 'strongly agreed' with the statement that they enjoyed the stories in the worksheet ($M = 3.6$, *Median* = 4; Likert scale ranging 1-4). And in a few cases, students really did appear to learn from the Hints and Explanations, arriving at the correct answer after some struggle and then thoughtful answering of the AL prompts.

Still, the message from testing these written materials was clear. Overall, they were clearly too reading-intensive, but there are simply too many essential ideas in these materials to cut down on much of the content. Therefore, reading had to be converted into something less cognitively draining. Moreover, this new form also had to make conditional branching easier to individualize even if its complexity increases. While I had initially been opposed to computer-based materials because their technological requirements make them less accessible, it became obvious that adapting the materials to be multimedia videos was the only way to achieve my goals. By using visualizations and animations, students would not only be more engaged (relative to over-reading) but this format would also free up students' cognitive resources to allow them to understand the fundamental concepts being taught. The PRM contextual details in the written format were just extra text, but in a multimedia video, they would be conveyed visually with illustrated characters and aurally with narration and spoken dialogue. The AL prompts could be more clearly separated and linearly ordered (not all squished together on a single piece of paper that could be read or answered in any order). And the PF concepts could be visualized by animations that demonstrate how different representational forms relate to each other; for example, letter variables can be animated to show them expanding into word variables. In short, multimedia video lessons provided a solution to every problem discovered in testing the written materials in the classroom.

4. Multimedia Materials

The written materials (see Chapter 3) were adapted into online multimedia materials in order to alleviate the issues discovered in testing. Video animations were created in Keynote and exported to iMovie, where they were edited together with recorded narrations. Exported video files were then uploaded to Zaption (Stigler, Geller, & Givvin, 2015), a web platform for educational videos, where all three conditions of the experiment are housed. Zaption provides an interface for playing the videos, recording students' answers to open response and multiple choice questions, and even performing conditional branching on multiple choice responses.

The materials for the multimedia project consist of a learning intervention and immediate post-test (Sesson 1, all online on Zaption), as well as a delayed post-test (Session 2, on paper). There are three different manipulations that have been designed for these materials: Contrast Comparisons (CC), Active Learning (AL), and PRM Context (PRM), all of which combine to form the overall Purpose-Driven Progressive Formalization (P-PF) approach. These manipulations will be discussed in more detail throughout this chapter. Given the practical limitations of time and resources for the present phase of this project, all three manipulations are currently implemented in a single condition (called 'Experimental'), which is meant to embody the entire P-PF framework discussed in Chapters 1 and 2. All three manipulations are currently absent in another condition (called 'Control'), which is meant to provide a lesson much closer to the status quo of teaching variables. In future phases of this project, the manipulations will be varied systematically in a factorial design, in order to assess their relative effect sizes and the presence of any interactions between them. For now though, 'Experimental' embodies all three manipulations. The manipulations are referred to individually when possible to delineate which aspects of the Experimental condition can be separately varied in future work. Additionally, there is a 'Baseline' condition, in which students receive no material on algebra except for the immediate and delayed post-tests; this condition is included only to measure the populations' baseline knowledge of algebra in order to demonstrate the efficacy of both the Experimental and Control conditions.

These materials were originally developed in written format, iteratively piloted and revised, and tested at the classroom level (see Chapter 3). After testing, it was clear that the written format was too reading-intensive (3.4.4), so the materials were adapted to multimedia video format, iteratively piloted and revised, and finally tested again at the classroom level. This process was heavily iterative throughout (not just during piloting), with constant revisions and adjustments being made during development as well. While it is not possible to document and discuss every such change here, I will focus on the current version of the materials, and when helpful, the major design decisions and changes that led to it; I will also discuss how my own framework was implemented, the failure of the initial written materials and reasons for adapting them, and how the framework was implemented even more effectively in the multimedia format while alleviating problems from the written format.

4.1 Session 1: The Core Mathematical Problems

Both the Experimental and Control conditions of the materials are structured around three *core* mathematical problems. In order to reduce possible math anxiety, the term 'Problem' is not used with the students. In the Experimental condition, because the students receive additional, story-like PRM details (Context manipulation), each problem is labeled as a 'Story' that has a 'Question'. In the less story-like Control condition, the problems are instead labeled 'Practice'.

Despite the different labels, these core problems are identical across conditions. Below, 'Story', 'Practice', and 'Problem' are used interchangeably unless otherwise noted. Likewise, the entirety of this section (4.1) applies to both Experimental and Control conditions, unless otherwise noted.

4.1.1 Problems' Mathematical Structures and Sequence

The mathematical structures for the three *core problems* (see Table 2) are 1) a single-unknown, single-equation problem (One-Unknown); 2) a two-unknown, two-equation problem where the to-be-found-first unknown's coefficient increases across the two equations (Change); and 3) a two-unknown, two-equation problem where the second equation states a multiplicative relation between the two unknowns (Relational). The schematic for the mathematical structure of the

Table 2

Multimedia Materials' Story Structures

Story	Story Label	Problem Structure & Description	*Implied* Form	Desired Unknown
1	One-Unknown	1 unknown, 1 equation	$2t = \$4$	Price of 1 taco (t)
2.1[a]	Unsolvable Two-Unknown	2 unknowns, 1 equation (Unsolvable)	$2p + w = \$5$	Price of 1 slice of pizza (p)
2.2	Change	2 unknowns, 2 equations (1 unknown changes quantities)	$2p + w = \$5$ $4p + w = \$9.50$	1. Price of 1 slice of pizza (p); 2. Price of 1 bottle of water (w)
3	Relational	2 unknowns, 2 equations (1 equation represents relation)	$c + d = \$4.50$ $c = 2d$	1. Price of 1 donut (d); 2. Price of 1 cupcake (c)

[a] Story 2.1 is not given in the Control condition, as it is used only as part of the Contrast Comparison and Active Learning manipulations.

problem sequence is described in Table 2 in algebraic terms (for reasons of compactness and precision, which of course reflects some of the strengths of algebraic representation). However, it bears repeating that in practice they are always presented as ‘word problems’ in video form (in the PRM manipulation there are added contextual details, and in all of the Experimental condition there is never any explicit mention of ‘algebra’ or ‘variables’). The problems are sequenced logically according to their structural and intuitive complexity, in the order in which they were historically developed. They were revised as necessary throughout piloting.

The One-Unknown problem (1) is primarily used as a warm-up and to provide a contrast with the Change problem (2.2), especially for the Experimental condition, in which students are asked to try to solve the the intermediate Unsolvable Two-Unknown problem (2.1), which contains the first part (equation) of the Change problem (2.2), in between the two solvable problems (1 and 2.2, see section 4.2 below). The Change problem is expected to be solved using an elimination technique, because focusing on the unknown being solved for (the price of 1 slice of cheese pizza) would more naturally lead to an elimination-style insight, since the number of slices of pizza ordered (thus, the coefficient of the price) *changes* while the number of bottles of water does not. Then solving for the second unknown (the price of 1 bottle of water) in the Change problem encourages substitution-style thinking (substituting a specific value for the previously unknown quantity, analogous to ‘back substitution’ in linear algebra). The Relational problem (3) is less easily solved by elimination (an important practical lesson), and is expected to be more naturally solved by the substitution method; this substitution is much more abstract than finding the second unknown in Change, because it requires substituting one unknown for another, as opposed to a specific value.

The precedence of the Change problem before the Relational problem is predicated on the former's intuitiveness. First, because it is more natural to think implicitly about changing quantities than replacing one type of quantity with another. And second, because the use of more concrete back-substitution after elimination in the Change problem forms a progressive bridge to the more abstract unknown-substitution in the Relational.

The succession of these problems is designed to facilitate *contrast* comparisons across each adjacent pair. Although subjects in the Control condition may intuitively make such contrasts on their own, they are made explicit in the contrast prompts in the Experimental condition (CC manipulation). For example, in contrasting the one-unknown first problem with the two-unknown Change problem, students may notice that the addition of another unknown increases the complexity of the problem structure, and as such, having a concept that represents unknown quantities across multiple operations becomes essential to the representation and solution of such a problem. The *conceptual purpose* of a variable is in part to organize information in a way that elucidates the sequence of procedures that must be carried out to find the unknown. And if students can be led to at least have an intuition that they need a new form of representation, then when such representation is subsequently introduced (using progressive formalization), its purpose is clear.

Importantly, in line with decomposition, the primary change between each adjacent pair of stories is the single component of the concept they are meant to learn. Minor contextual details change from problem to problem only to differentiate the new problem's quantities, reducing interference with the previous story and making it a less redundant task. Students' comparisons between problems in this logical/historical order can help them realize which elements of the problems determine their complexity and how that relates to a more effective solution strategy.

As students progress to the harder problems, their problem solving strategy and concept of a variable may become more generalized and abstract. And because historical conceptual development is driven by purpose, the purpose of using variables becomes clearer in these more complex problems. The same problem sequence is used in both Experimental and Control conditions to allow for Control subjects to still spontaneously make the same inferences as those in Experimental.

In the initial written version of these experimental materials (see 3.2.1.1), there was a fourth problem (Double Change: two-unknown, two-equation where both unknowns' coefficients changed in opposite directions); this problem type was meant to consolidate the relational nature of a substitution method (as in Relational) with a mathematical structure more similar to the Change problem (though it could not readily be solved using elimination), and lead to an even more abstract type of substitution (substituting a whole expression for an unknown). This problem was only omitted from the multimedia format because of time constraints (to keep the entire first session under an hour and a half), although it is planned for future versions of the experiment that have multiple learning sessions.

4.1.2 Word Problem Contexts

Each of the three core problem's contexts consists of a student (the narrator) or his friend purchasing some food items, knowing the total cost, and then wondering what an individual item cost. This general context was largely determined by the PRM framework and inspired by an example from Kraft (2013). His "muffins and donuts" problem provided a good basis for my own Change problem, though I added more PRM details by clarifying how the narrator's knowledge state was plausible and making the narrator more relatable as a student; I also switched which quantity was being solved for first so that the changing quantity becomes a focal point and the goal structure is clearer. Working both forwards and backwards with the Change problem as the central focus, I adapted Kraft's (2013) general problem context to fit each problem structure I had designed. These core problems' structures are outlined in Table 2.

Students are likely to have experienced purchasing food and drink items, so the context is *relatable* (much more so than the contexts of the puzzle problems discussed in Section 1.4.4.1). The food items themselves (tacos, pizza, bottle of water, cupcakes, and donuts) are also relatable. These aspects of the context's relatability are consistent across all conditions. I attempted to vary enough details across stories to prevent redundancy, but kept much of the bulk of the context as similar as possible to highlight similarities and differences across problems. In the PRM Context manipulation of the Experimental condition, many additional PRM details are included in extra Context videos (while not adding anything to the mathematical structure itself). The full text of each Context video (and all other videos) is available in Appendix A, and some PRM details are summarized in section 4.3.

4.1.3 The Numbers: Coefficients, Constants, and Unknowns

The values of the problems' coefficients, constants, and to-be-found unknowns were determined by both design principles and feedback from piloting. Following the PRM framework, during the initial phase of learning, all values had to be *plausible* given the word problem context. As such, the food items' prices had to be reasonable in-context: each taco in the One-Unknown problem cost \$2, each slice of cheese pizza in Change cost \$2.25 and the bottle of water cost \$0.50, the donut in Relational cost \$1.50 and the cupcake cost \$3.00. Moreover, within the

Change and Relational stories, the relative costs of each item had to be sensible as well: pizza should cost more than water, and a cupcake should cost more than a donut.

All items' costs were chosen to be multiples of $0.25 so that the math wouldn't be overly cumbersome, which would have drained students' cognitive resources by forcing them to carry out arithmetic procedures and prevented them from making the conceptual insights desired. However, the costs were specifically chosen to not be whole dollar amounts (except the One-Unknown problem, which was intended partially as a warm-up) so as to minimize "guess and check" strategies and to add a bit more plausibility.

The cost of one slice of cheese pizza in the Change problem ($2.25) deserves its own discussion. Importantly, in the Experimental condition, the information is presented in two separate phases. In the first part, Unsolvable Two-Unknowns, only the first equation's worth of information is presented (narration: "I went to Joe's Pizzeria, and I bought two slices of cheese pizza and one bottle of water for five dollars.") Subjects are then asked to find the price of one slice of cheese pizza, despite the fact that there is not enough information. Because students have just solved One-Unknown (where the price of one taco was $2), and because the total for two slices and one water is $5, they may be tempted to think that the price of one slice of cheese pizza is $2 (and the water is $1). Immediately afterward, they are then asked whether there is enough information to know the price for sure. This question sequence is meant to push them to contrast an intuitive guess (if they thought $2) with the realization that actually, there are multiple possible prices. This contrast is meant to evoke an intuition that the presence of more unknowns than equations prevents there from being a unique solution, as well as an intuitive sense that a variable can actually represent multiple possible values. Each of these intuitions is then contrasted with the next part of the story (the Change problem), in which the same two unknowns can now be found, and are restricted to a single solution set. And in the Change problem's multiple choice options, $2 is used as an enticing lure.

4.1.4 Story Videos: Illustration, Animation, Text, & Narration

The Story (Problem) videos show slightly-animated characters interacting, with narration and on-screen text as necessary to convey the contextual and mathematical information. These videos were created as a way to greatly reduce the reading demands of the materials (especially relative to the initial written format) and to facilitate students' attention and interest. Examples of the Story, Context, and Question videos (for the One-Unknown problem) are available at http://zapt.io/t3grxnh4.

Each Story video shows an illustrated character (the narrator in One-Unknown and Change Stories, or his friend in Relational) standing in front of a counter that has a cash register on top

Figure 1. Screenshot example of a Story video at the end of animation (One Unknown Story).

and worker behind it (Figure 1). A sign on the panel above the worker displays the name of the restaurant ("Gloria's Tacos", "Joe's Pizzeria", or "Patty's Bakery", each in a font reminiscent of the type of establishment). The audio narration is synchronized with the onscreen text (which consists of the most critical words from the narration, following Mayer & Moreno, 2003).

Throughout all of these Story videos, while the narration talks about ordering, the character's arm moves up and down to indicate they are placing an order. When the narration starts to talk about the total cost, the character's arm stops moving in the "up" position, and the amount of money (in green text) appears just above the character's hand, as if they are handing the cashier the money. Food items are also depicted in the Story videos with their illustrated images appearing on the counter (in the appropriate quantity), and they fade in when they are mentioned in the narration.

While the visual scenes in the Story videos are kept relatively sparse, there are still a lot of visual details. Thus, colors and formatting are used to help students identify the most important aspects. Numbers are always bolded, as are food item words (e.g., "slice of cheese pizza") and dollar signs. Food item words are always colored either red or blue (whichever represents the food/drink better and contrasts with the other item if there are two), and money is always colored green. As much as possible, the food item images also reflect the color of their corresponding text (the pizza has a lot of red sauce, the water is blue; the donut has red glaze, and the cupcake has blue frosting). The rest of the Story scenes are in black and white, but more important aspects (e.g., the narrator) appear in greyscale (multiple shades) while less important aspects (e.g., the cashier) appear only as white with a thin black outline. All of these visual cues are expected to facilitate students' attention most to the colors, then to greyscale, and least to black and white.

4.2 Session 1: Experimental-Specific Manipulations

While the Experimental and Control conditions share the same core problem sequence, the majority of the remainder of Session 1 (except the Immediate Post-Test, discussed in section 4.5) is manipulated differently across the two conditions. As much as possible, the Experimental and Control conditions receive the same *information*, only presented differently or in a different order. In this section, I will discuss the aspects of the materials that are specific to the Experimental condition. When helpful, these aspects will be contrasted with their counterparts in the Control condition, although see section 4.3 for a fuller discussion of the Control.

There are three manipulations present in the Experimental condition that are absent, negated, or reversed in the Control. First, students are asked to make Contrast Comparisons (CC) between different problem structures and representational forms. Second, Active Learning (AL) is implemented throughout the materials by allowing students to constructively struggle through a problem and providing increasingly specific hints as necessary. And third, additional Context (PRM) is implemented in the form of supplementary videos that incorporate PRM details (without any relevant mathematical information), such as why the characters in the story want to find the unknown's value and why they don't already know it.

Throughout the Experimental materials, Progressive Formalization (PF) is implemented by introducing formal representations only after the student has intuitively understood their conceptual and practical purpose. PF is achieved through (and partially dependent upon) the three manipulations (CC, AL, PRM), so it is not independently manipulable from them. Instead, the manipulations are what makes the overall PF design of the Experimental condition *purpose-driven* (P-PF), as opposed to generalization-based (G-PF; Nathan, 2012; Romberg, 2001).

In the Experimental condition, the core problems are given in succession, but in the Change and Relational stories, there is significant conditional branching to allow for follow-up Hints in the case of an incorrect answer (AL). And at the end, there is a Walk-Through (adapted from the Explanations in the written materials, see 3.2.1.3) for each of these problems, which is somewhat streamlined if the student got the questions correct; but if the student got either question incorrect, the Walk-Through videos first give a more advanced version of the Hint (AL), allow them to attempt an answer for a third time, and then give them an appropriately customized Walk-Through. At the end of all three Stories (and Walk-Throughs), the students go through a few Lesson videos, which formally encapsulate the ideas they progress through in the Stories and Walk-Throughs by introducing letter variable representation and demonstrating a symbolic solution for the Change problem.

The Experimental manipulations are described in more detail in the sections below, but the the scripts for all of the narration and its relation to onscreen text and animations is extensively documented for each video in Appendix A (Control scripts documented in Appendix B), and the Experimental flowchart in Appendix C helps to visualize the conditional branching. An example of the Experimental condition, hosted on Zaption.com, is available at http://zapt.io/thkwjr8x. Because these materials are highly interactive and contain a lot of conditional branching, it is recommended that readers explore the materials online, skipping backward to change answers to questions to see how this affects the progression through the materials (by conditional branching).

4.2.1 Progressive Formalization

Progressive Formalization (PF) is implemented throughout the materials in several ways, and forms the core organizing principle of the materials; because all three independent manipulations (AL, CC, and PRM, discussed below) are present in the Experimental condition, PF is implemented in a purpose-driven manner (P-PF). Importantly, students are asked to attempt to solve each problem before any instruction is given. While this is an Active Learning (AL) technique (constructive struggling, Stigler & Hiebert, 1999), it sets important groundwork for PF by letting students develop their own intuitions about the problem and how to solve it (thus, PF cannot be independently manipulated from AL). Then, when new representations and solution strategies are progressively introduced, students can map their intuitions onto those formalisms. Similarly to the written materials (3.2.1.4), PF in the multimedia version can be roughly separated into two different 'progressive arcs' (from sentences to word equations, and from word equations to letter equations), each of which takes place across multiple points in time (and using both AL and CC manipulations, discussed more fully below).

4.2.1.1 Progressive Arc 1: From Sentences to Word Equations (Walk-Throughs)

In the materials' PF, 'word equations' (e.g., "2 * the price of 1 slice of cheese pizza + the price of 1 bottle of water = $5") play an important intermediate role in progressing to the use of letters as variables. The purpose of having a formal representation for an unknown or variable quantity becomes considerably clearer when students come to two-unknown problems. Word equations, just like 'letter equations', could be introduced in a single-step, single-unknown equation (e.g., '2 * the price of 1 taco = $4'), but these representations seem overkill for such simple problems because they are easily solved by arithmetic (the position of the unknown is easily retained in working memory, while arithmetic may be written down explicitly as separate steps). In order to make a single-unknown equation warrant such lengthy notation as a word variable, we'd have to construct a more complex, multistep equation (e.g., the necessity of a placeholder like 'x' in '$6 * 3 + 5(x - 2)$') that would be difficult to match to an easily relatable real-world context. As such, the most appropriate situation in which to introduce word equations is in solving a two-variable system of equations, where the use of the new notation allows for the solution strategy to be better visualized when the unknowns would otherwise be difficult to keep track of mentally.

Importantly, these materials build toward this concept from the beginning, inviting contrasts both implicit (due the sequencing of problems) and explicit (with CC prompts) between problem structures with one unknown in one equation (1, One-Unknown), two unknowns in one equation (2.1, Unsolvable Two-Unknown), and finally, two unknowns in two equations (2.2, Change). These contrasts are aimed at distinguishing between the problem structures, and in turn, how the structures determine solution strategies and solvability. Since the One-Unknown problem is a single-step procedure, it is easily solved with arithmetic, and there is no need for students to explicitly represent the unknown. Then, upon being given the Unsolvable Two-Unknown problem, they may attempt an arithmetic solution or just offer a plausible guess. But through the subsequent CC prompts, they are encouraged to discover that multiple values are possible for the unknowns (i.e., they are underdetermined). Thus, the student may conclude that arithmetic is not sufficient to guarantee a solution, and sometimes a quantity may not be knowable.

These intuitions then pave the way for the clear contrast between the Change problem (which includes the previously unsolvable equation with a second, independent equation) and the two previous problems. As students attempt to find a solution to the Change problem, arithmetic or otherwise, they may begin to form intuitions about the purpose of a new representation that *extends* arithmetic, which connects algebra to their previous arithmetic knowledge as opposed to letting a gap form between them (Herscovics & Linchevski, 1994). After the intuitive purpose of the concept of a variable has been encouraged to form, word equations are introduced in the Walk-Through for the Change problem, which is presented after a student has either correctly solved the problem or failed to solve it even after the Hint videos (AL).

Word equations are more *mathematically precise* than sentence descriptions: they specify the exact quantitative relationship between known and unknown values, according to exact rules of mathematical symbolization and compositional syntax. While there is some allowable notational diversity in math (e.g., ‘6*4’, ‘4 x 6’, ‘6(4)’, etc., all refer to the same quantitative relation), it pales in comparison to the combinatorial power of natural language. And because the formal symbols of math are unambiguous (to those who are fluent in them, of course), word equations, if properly understood, impose a much more precise and rigorous structure than sentences. However, word equations are also more *abstract* than sentences. Like all mathematical formalisms (Nathan, 2012), word equations trade some concrete semantics for more systematic syntax. And while students may think intuitively with more concrete contexts (e.g., PRM details), the formal rigor of mathematical expressions is more generalizable. As such, depending on the student’s task, each representational form has its own benefits.

In order to encourage students to think about these tradeoffs between representational forms, after students are shown in the Change Story Walk-Throughs how the story information is converted into mathematical form using word-variable terms such as “the price of 1 bottle of water”, they are then asked to contrast the two representations (CC). Specifically, the word equation representations (e.g., Day 1: “The price of 2 slices of cheese pizza + the price of 1 bottle of water = $5.00") are contrasted with the initial sentences from the story (e.g., "I bought 2 slices of cheese pizza and 1 bottle of water for $5.00") by asking students to indicate which representational form is more *useful* for finding the price of 1 slice of cheese pizza (Figure 2). This contrast drives students' attention to focus on a *purpose* of the representation—both the mere fact that one representational form can be more useful than another for a given task, and also that it is useful *per se* to have a form that mathematically models an unknown number.

Later on in the Walk-Through videos (after students who have not yet correctly answered the problem receive advanced hints and a final attempt), all students see a Walk-Through of how to solve the Change problem using word equations, with some details abridged for those who already got it correct. This video demonstrates how to use an intuitive elimination method on the word equations. First, the part of the word equation expressing the price of the pizza each day is highlighted as the narration says "the difference between Day 1 and Day 2's order was two extra

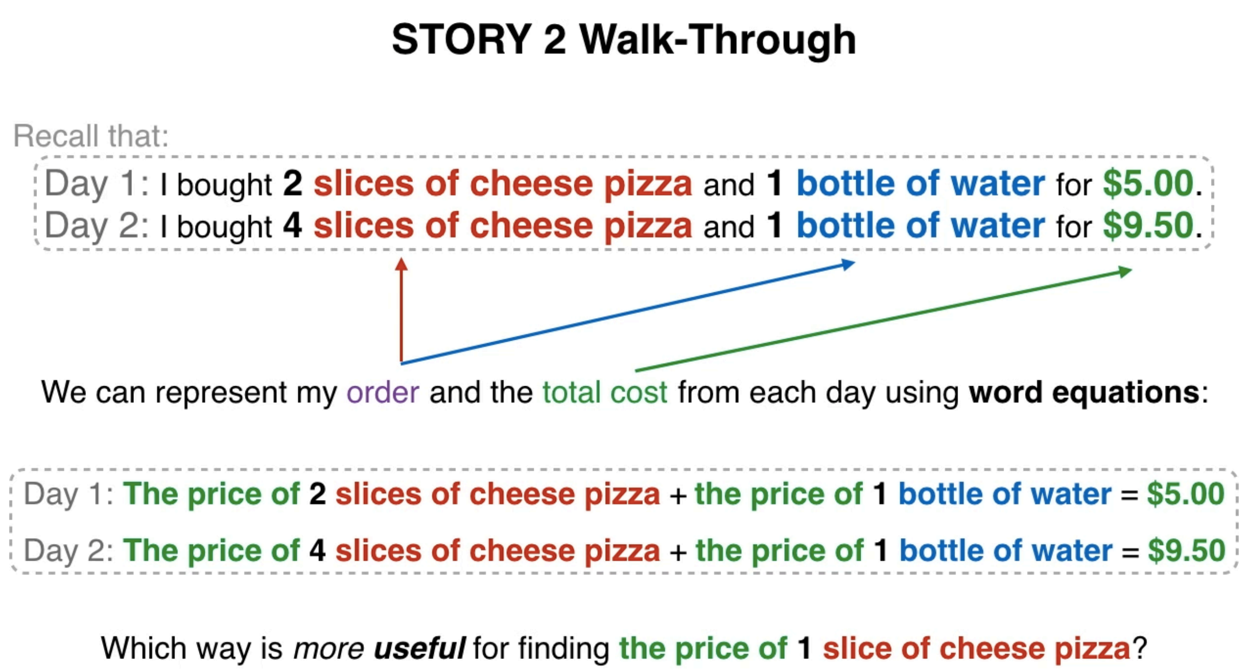


Figure 2. Screenshot of contrast comparison between sentence descriptions and word equations (Change Story Walk-Through).

slices of cheese pizza," and the text "the price of 2 extra slices of cheese pizza" appears below. Then, the total cost in each equation is similarly highlighted as the difference between days' total cost is stated to be \$4.50, and "\$4.50" appears below on the same line. Students who did not get Q1 correct are then told "The increase in the cost was due to the extra pizza ordered" (this is elided for others to streamline the video), which uses implicitly causal language to facilitate the cause-effect connection between the extra pizza and the increase in price; subjects' implicit causal reasoning will help them intuitively 'eliminate' the constancies across the equations to focus on the remaining differences. Lastly, for all subjects, the narration states that the price of the two extra slices and \$4.50 must be equal, at which point the equals sign appears between them so that they now form a new word equation expressing the differences. From there, it is an easy step to see that one slice's price must be \$2.25. Thus, without using any formal subtraction, students can intuitively understand the elimination method, a progressive step on the way to understanding formal elimination, which is later encapsulated with letter variables in the Lessons.

Similarly, the Walk-Through of how to solve the second question in the problem (the price of 1 bottle of water) gives an intuitive, visual demonstration of substituting the now-known value of the price of 1 slice of cheese pizza for the *words* that represent that same quantity. The price is animated to move in such a way that it appears to actually replace the words where they once were. Animations like these, used throughout the videos when relevant, are not merely decorative—they actually intuitively represent the concept being performed. And seeing the substitution of a number for words is a progressive step toward the substitution of words for other words, which comes in the Relational story.

4.2.1.2 Progressive Arc 2: From Word Equations to Letter Equations (Lessons)

Word equations play an important role as a 'bridge' to letter equations (traditional symbolic algebra with letter variables). Both a word variable and letter variable represent the same concept and may be instantiated to have the same referent quantity. The move from using word equations to letter equations is made at the level of *representation*, not along a continuum of conceptual abstraction: letter variables are merely *abbreviated* word variables, and variables are conceptually abstract, independent of their symbolic form.

Letter variables are both *more compact* and *less specific* than word variables. Their lack of specificity may make letter variables seem more abstract than their word counterparts, but this is because their reference is more opaque. Even in the case of purely non-contextual, symbolic algebra with x's and y's that have no referents in a story or the real world, each letter variable is essentially an abbreviation of more specific phrases like "some number we call x" and "some *other* number we call y, which may have a different value than x", because that is exactly how the letters are intended to be interpreted.

The predominance of letter variables over word variables is entirely due to their compactness, which is a purely representational advantage. In shifting our mode of representation from words to letters, we are making a *practical* decision. This is a critical insight to make because it informs the contrast we must encourage students to make between word equations and letter equations: they differ neither in referents nor abstraction, only in compactness and specificity.

After students first learn about word equations in the the first Walk-Through video for the Change Story, they are instructed to practice writing down the word equations on their scratch paper. Then, for Question 1 in the Relational Story, they are encouraged by onscreen text to first write down word equations for the story and then use them to solve it. This provides students with experience both in recognizing the internal mathematical structure for a story and also in formalizing that structure into word equation representations. It also helps them see how the representation may be useful in solving a new problem type (the Change and Relational Stories differ in structure). Since students now have experienced writing down word equations for two different problems (four equations total), they have some idea what it is like to use word equations. From a practical standpoint, they may have noticed that word equations take a while to write and take up much of the page, which may distract students, make it harder to identify common terms, and see the overall mathematical structures.

After completing the Relational Story and its Walk-Throughs, students begin the Lessons. First, they are reminded that word phrases like 'the price of 1 slice of cheese pizza' represent a number, and it's noted that word equations may become tediously long. Students are then asked

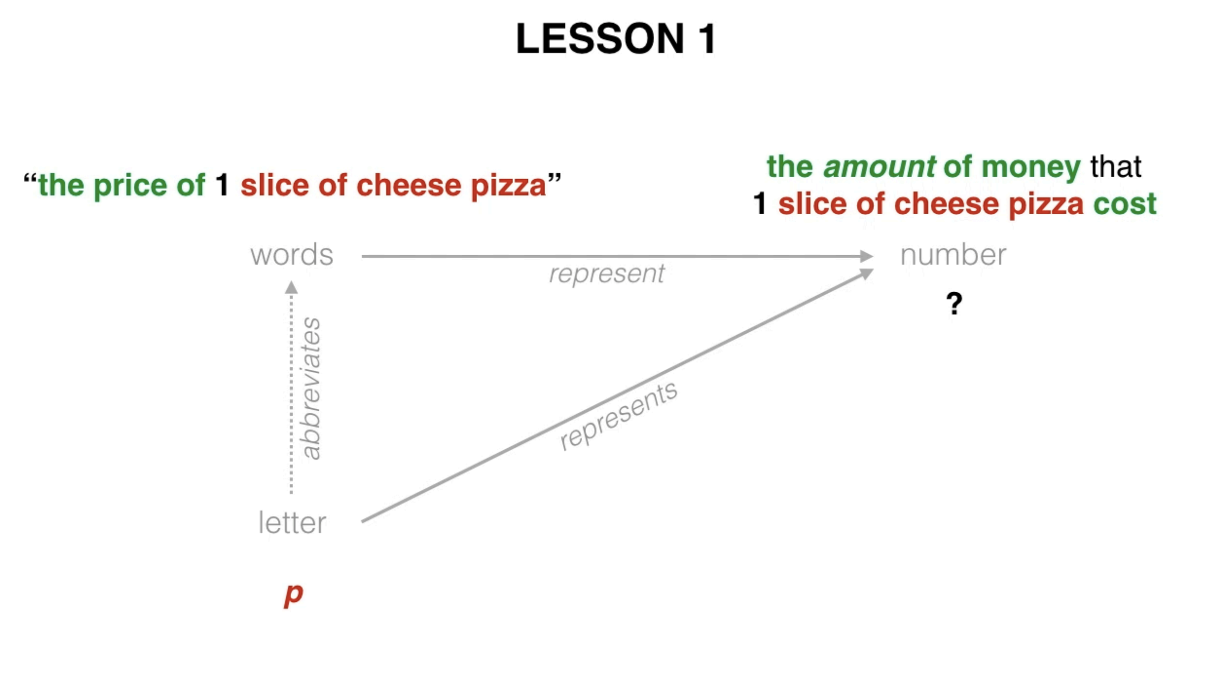


Figure 3. Screenshot of schematic relating letter variables, word variables, and the contextual number they represent (Lesson 1 Lesson Video).

how they could make word equations shorter (in an open response question). After being allowed to come up with their own solution (AL), the subsequent video demonstrates that the letter '*p*' can be used to abbreviate the word phrase, and '*p*' itself also represents the same number (Figure 3). Thus, letter variables are simply introduced as abbreviations of word variables that have the same referent. Then in the next videos, students are shown how to abbreviate each entire word equation into an equation with letters.

Students are told in Lesson 3 that "Many people find letter equations to be the most useful way to represent information because they clearly show the math, and are shorter than word equations, while still representing the same numbers." (Animations are used to highlight what the narration is stating.) This consolidates their gradual learning of letter variables by the two primary means of progression: first, that equations clearly show mathematical structure, and second, that letters are simply shorter than words. Subsequently, a series of animations help students visualize that the words are 'contained' inside the letter that abbreviates them.

In the final Lesson, the elimination method for the Change Story is demonstrated formally as subtraction of letter variables. This video, as best as possible, visually demonstrates the concepts that have been progressively developed from students' initial constructive struggling, through intuitive elimination with word equations in the Walk-Throughs, and into formal subtraction of letter variables in the final Lesson videos.

It is worth noting that nowhere in any of the videos for the Experimental condition is the word 'variable' used; instead, students are just asked to use letters to represent numbers whose values aren't known yet. Since many students are likely to have heard the word 'variable' used before, it better fits the goals of PF to let the students learn as much on their own as possible without leaning on a term they may already have preconceived notions about. In the Control condition, by contrast, ‘variable’ is used and defined in the very first video, as would be commonly done in traditional instruction.

4.2.2 Contrast Comparisons and Active Learning

As stated previously, students are encouraged to solve each problem before being given any sort of instruction; this constructive struggling (Stigler & Hiebert, 1999) is itself a type of Active Learning (AL). Beyond that, however, AL is also implemented through increasingly specific Hints, which are given as necessary. Independently of the AL manipulation, Contrast Comparisons (CC) are also given throughout the materials. Because AL and CC prompts are often interwoven within the same parts of the materials (though they can still be independently manipulated), these two separate manipulations are both discussed in this section. The general schema for the Change and Relational Stories (which have most of the AL and branching) is depicted in Figure 4 and described below.

4.2.2.1 Change Story

As briefly discussed above in Decomposition (1.4.3.1), CC prompts are used throughout the Unsolvable Two-Unknown and Change Stories to help students discover the structural and interpretational differences between a one-equation, one-unknown problem (Story 1); a one-equation, two-unknown problem (Unsolvable, 2.1); and a two-equation, two-unknown problem (Change, 2.2).

After completing the One-Unknown problem, students are shown the Unsolvable Story and are asked to find the price of one slice of cheese pizza (one of the two variables) in an open-response question (constructive struggling, AL). Of course, the price is underdetermined given the current state of knowledge, so there are a range of possible values. After giving their answer, students are then asked whether there is enough information to know for sure what the price is.

Next, students are led down a series of follow-up CC questions to get them to think more critically about the amount of information known. The order of these three questions is meant to demonstrate first that there is at least one value it can't have, and then to *contrast* that with *two* other possible values. First, they are asked whether it is *possible* for one slice to cost $2.75 (which it can't be under the reasonable assumptions that the bottle of water doesn't have a negative number for a price and that each slice costs the same amount). Second, they are asked whether it is possible it cost $1.75 (which seems reasonable), and third, that it cost $2.10 (also reasonable). Because the first example is not possible, this emphasizes to student that some information is known, but since the latter two are both possible, there clearly isn't enough information to know a value for sure. Therefore, at least two values are possible, but at least one value is not; this inference combined with the problem context makes as concrete as possible the insight that the unknown may be intuitively thought of as a set of values, which may even generalize to other values beyond the two possible ones thus far. As feedback for the third and final example, they are given an explanation clarifying that although $2.75 is impossible, both $2.10 and $1.75 are possible prices (along with others), so we do not have enough information to know for sure.

The Unsolvable Two-Unknown problem itself and these CC prompts serve as a conceptual bridge between the One-Unknown and Change problems, as a challenge to students to really think about a problem's structure, and as a lesson that not only are some problems unsolvable, but at least one type of unsolvable problem (2 variables, 1 equation) is easily recognizable (though further examples would help clarify this). Ideally, the students also learn that unsolvable problems aren't uninformative (i.e., "not enough information" does not entail "no information"), as indicated both by the fact that $2.75 could still be ruled out as a possible value and that the information in the Unsolvable problem is repeated in Change, which is solvable (so there must be information there).

After seeing the Change problem, students are asked whether there is *now* enough information to know the price for sure (CC). This contrast between two different states of knowledge (one equation vs. two equations) actively pushes them toward understanding that an unknown can go from unknowable (variable) to knowable (fixed) with more information. Thus, the interpretation (meaning) of an unknown can differ depending on what else is known.

Finally, students are asked to find the price now that it is knowable, which is where the conditional branching (part of AL) fully begins (Figure 4). The overall goals of the branching are as follows: first, students who answer incorrectly should get increasingly more specific hints, as needed, to lead them toward one possible solution strategy that is intuitive in context; second, because students who answer correctly may have used a "guess and check" method or some other strategy that is different from the one given by hints, they should be walked through the intuitive hint strategy as well so that they can still learn, although it can be a bit more streamlined since they already got it correct. All students get to attempt to answer a two-unknown problem's Question 1 up to three times (initially, with a Hint, and after the Walk-Through) and Question 2 up to twice (initially *or* with a Hint, and after the Walk-Through). Broadly, whenever a student provides an incorrect answer, they are given a nudge toward an intuitive solution and asked to try again.

For example, if a student gets the Change Question 1 (the price of the pizza) incorrect (the topmost red-arrow branch in Figure 4), they receive a series of two Hint videos (with AL ‘thinking’ questions) before being asked to try the problem again. The Change Hint 1 video tells students to think about what the narrator ordered and how much it cost. Then students are asked how the order *changed* between the two days (open response question). After answering, they

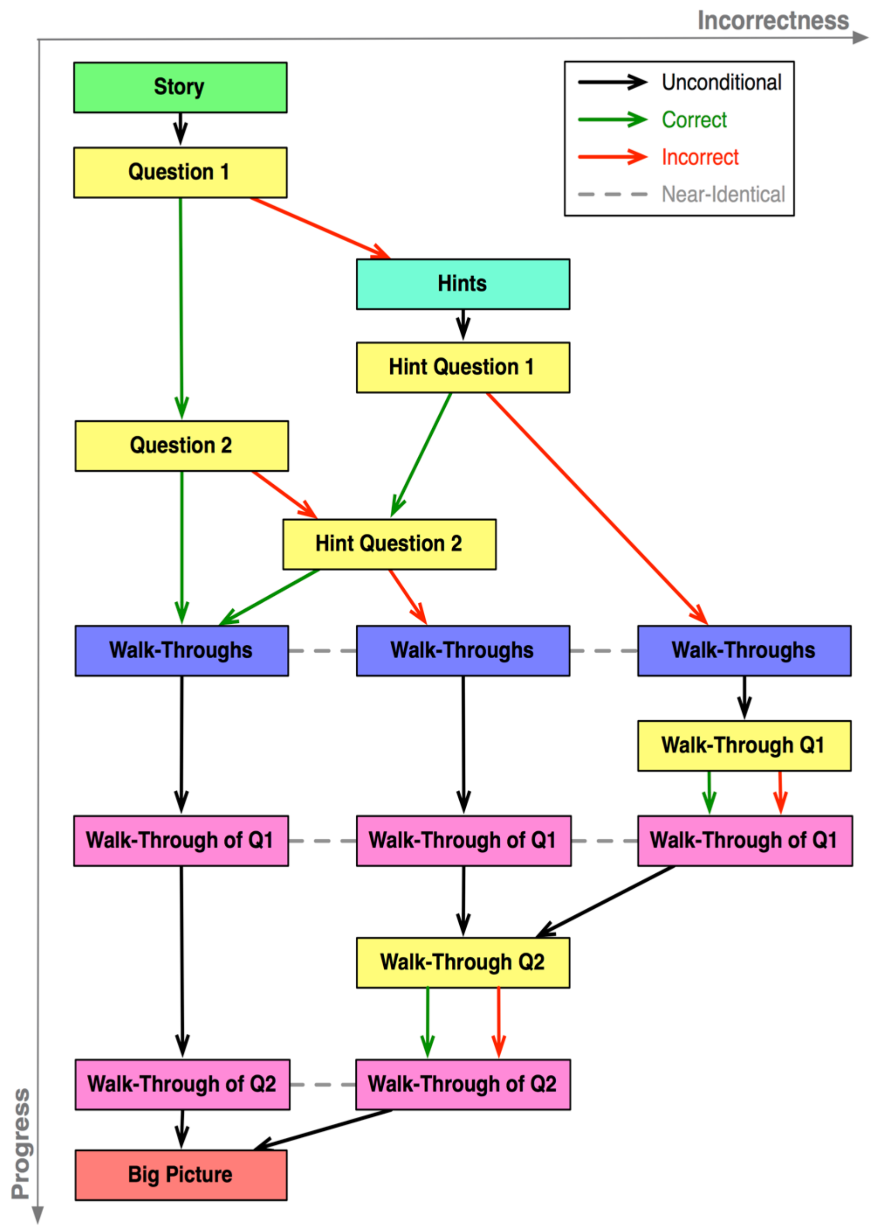


Figure 4. Simplified flowchart of conditional branching (Hints and Walk-Throughs, AL) for Change and Relational Stories. Students progress from the top of the flowchart to the bottom, branching rightward if they answer incorrectly. At each yellow node, the student is asked one of the Story’s Questions, turquoise represents Hints, lilac is advanced hints / PF, and pink is an animated explanation.

are then asked in the Change Hint 2 video how the total cost changed (open response). These questions tap into students' intuitive thinking about the problem implicitly in terms of cause and effect: the narrator ordered two more slices of cheese pizza, which increased the price by $4.50. Thus, it naturally follows that one slice cost $2.25. This is an intuitive, contextual way of performing what is essentially the *elimination method* (eliminating the other variable, the price of the water, to isolate the first). Importantly, all of the Hint videos' content discusses the problem at the level of the *context*, not in any way at the representational or symbolic level, which may give students a more intuitive understanding of the eventual solution.

Students in this Hint track then get a second opportunity to answer Question 1, and if they are incorrect again, they are introduced to word equation representations (PF), given more advanced versions of the hints (AL), and then given a third and final chance to answer Q1 before moving onto Q2. All students are shown demonstrations in the Walk-Through videos of how to solve each Question from word equation representations: in the case of the Change Story's Question 1, students are shown how to intuitively carry out the elimination method using word equations, and for Question 2, how to use value substitution to find the second unknown. Additionally, when word equations are first introduced in the initial Walk-Through video, students are asked to contrast word equation and sentence descriptions (CC) in terms of which is more useful for finding the answer to the problem.

The coda for the arc beginning with the One-Unknown Story and moving through the Unsolvable Two-Unknown and Change Stories is the "Big Picture" video (CC). In this video, students are asked why it is that they couldn't find the price of one slice of cheese pizza after Day 1, but they could find it after both Day 1 and Day 2 (open response). The purpose of this CC question is to encourage students to consolidate all of the contrasts between the three prior problems to arrive at the conclusion that two-unknown problems require at least two equations' worth of information. After students provide an open response answer, they are asked a final multiple choice question of whether it is possible, now that both Days' orders are known, that one slice of cheese pizza could cost $1.75. This exact number was queried in the CC prompts from the Unsolvable Two-Unknown Story as a possible value, so this invites one final contrast to clarify that at least some previously possible values are no longer possible when the new information constrains the solution set. The feedback for this multiple choice item clarifies this point, with more detailed explanation if they incorrectly answered 'Yes'.

4.2.2.2 Relational Story

The AL manipulations for the Relational Story are structurally identical to those for the Change Story, differing only in the specifics of the Hints and Walk-Throughs; CC prompts are not used in this Story. The schema for Hints and Walk-Throughs and their conditional branching is almost identical. As such, only the distinctions in the Hints' and Walk-Throughs' content are discussed here.

The Hint (AL) video for the Relational Story first tells students to make use of the relationship between the price of the donut and the price of the cupcake and then asks the open-response question, "What would the total cost have been if my friend got 1 donut for me and 2 donuts for herself *instead* of the cupcake?" This questions taps into students' intuitive thinking about the problem implicitly in terms of a counterfactual: what if the friend purchased herself two donuts instead of one cupcake? This is an intuitive, contextual way of performing what is essentially the *substitution method* (substituting one variable, the price of one cupcake, for its equivalent in terms of the other variable, the price of two donuts). Importantly, this Hint is at the level of the *context*, not the symbolic, which should encourage students' intuitive reasoning.

The advanced hints (AL) in the Walk-Through videos expand on Hint prompt. First, students are asked how many donuts could be purchased for the same price as one cupcake; again, this is a contextual question, but it is a relatively easy step in the right direction toward intuitive substitution. Second, they're asked directly how many donuts could be bought for $4.50 total, which essentially requires them to perform cupcake-for-donut substitution in order to answer. These AL prompts in succession should lead students toward the insight necessary to isolate the price of a single donut.

4.2.2.3 Lessons

Before each Lesson video's demonstration (see 4.2.1.2 for details), one or more CC and/or AL questions are asked of the student to encourage them to intuit each new part of the lesson on their own. At the very beginning of the Lesson stage, students are asked to think again about word equations as representation (AL): "Writing all of these words every time makes the equations very long, and we might get tired of doing so much writing. How might we make these word equations shorter?" Only after answer this open response question are they introduced to the concept of abbreviating the words into letters (PF). Thus, after giving students the opportunity to actively intuit the practical purpose of using a letter as a variable, they are introduced to the concept with that justification, in case they did not come up with it on their own.

After being shown that 'the price of 1 slice of cheese pizza' can be abbreviated as just '*p*', students are also asked to come up with a way to represent 'the price of 2 slices of cheese pizza', and then a way to represent 'the price of 1 bottle of water'. These AL questions encourage students to practice thinking about letter representations and whether the chosen letter matters. Then they're asked to try converting the entire word equation to a letter equation.

After students have attempted and/or seen how both word equations can be converted into letter equations, they are shown all three representational forms thus far (sentences, word equations, letter equations) and asked which of the representations is the most useful for finding the price of 1 slice of pizza (see Figure 5). This CC question mirrors the one used when word equations were introduced (4.2.1.1 and 4.2.2.1) and leads them to once again contrast between multiple representations to see which has the most practical utility.

Lastly, students are reminded of the fact that in the Change Story, they isolated the change in item quantities ordered and in the total cost to find those items' prices; now they are asked which mathematical operation (given the 4 choices of addition, subtraction, multiplication, and

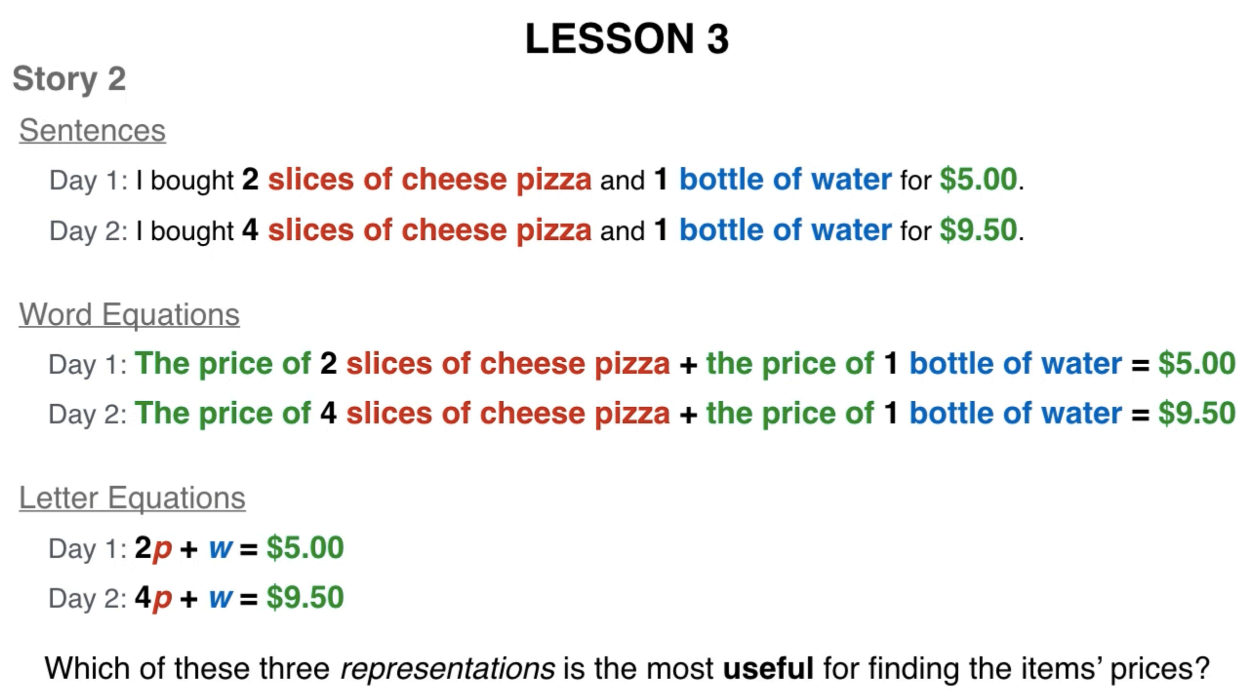


Figure 5. Screenshot of contrast comparison between sentence descriptions, word equations, and letter equations (Lesson 3 Question).

division) would help them to formally find the change between numbers (AL). If they answer subtraction, they are given a streamlined follow-up video showing formal subtraction with letter variables; if addition, a slightly less streamlined; and if multiplication or division, a longer, more explicit demonstration.

4.2.3 Context Videos

The extra Context videos only serve to increase the PRM details in the Context manipulation of the Experimental condition (example available at http://zapt.io/t3grxnh4, second video). The subjects in the Control condition do not see these Context videos, however, they still have all of the mathematical information necessary to solve the problem. The added PRM details are very similar to those implemented in the written versions (see Chapter 3), while also greatly reducing students' reading load. Since the additional context is supplied in a multimedia form, it was natural to adapt the descriptions from the written materials into spoken dialogue between illustrated characters. Characters' dialogue is spoken out loud, and shown visually by speech bubbles that fade in when they start to talk, and partially fade out when the next character begins to talk. The partial fade-out clarifies who is currently talking (whichever bubble has not faded) while still allowing each bubble to still be readable after the character finishes talking (lessening the cognitive load required). Additionally, the appropriate food item images are shown on the table, and when multiple people have different food items or amounts, that is reflected in the food in front of them at the table.

In each Context video, the narrator (the same character illustrated in the Story videos) appears at a table with one (One-Unknown and Relational) or two (Unsolvable Two-Unknown and Change) friends. The number of friends is determined by what is necessary to convey the motivation for finding the desired quantity. In the One-Unknown Story, the narrator buys two tacos, and his friend wants to know how much one costs so she can buy one the next day. In the Unsolvable Two-Unknown, the narrator buys two slices of cheese pizza and one bottle of water, and then in Change, he buys four slices of cheese pizza and one bottle of water. In order to motivate this, in the Unsolvable part (2.1), two of his friends want to know how much one costs because it looks good, and then in a Context video just before the Change Story part (2.2), he decides to surprise them and buy them each a slice the next time he goes (and order the same as last time for himself). Then in the Change Story's Context video, they want to pay him back for the pizza, so they need to know how much each of them owes him. This allows for the price of only one slice to be motivated in both stages of the problem (2.1 and 2.2). In the Context prior to Change Question 2, the narrator's friend tells him that the school cafeteria sells the same bottle of water he bought for 75¢, which may be cheaper than the pizzeria he bought it at. This motivates finding the second quantity as well (to see if it's cheaper). In the Relational Story, the narrator's friend buys him a donut, so he wants to pay her back. And then for Question 2, he's curious about her cupcake and wants to know how much it cost.

4.3 Session 1: Control-Specific Manipulations

As much as possible, the Control condition contains all of the same instructional *information* as the Experimental condition, but it is presented in a different manner and/or order. The Control condition begins each section with an initial Lesson that gives a formal overview of the concepts that will be taught in the current section (Formalisms First, FF). This is followed by a Story video, which is itself immediately followed by Lesson videos. Unlike the Lessons at the end of the Experimental condition, these Lesson videos demonstrate how to find the unknown quantities in the Story without ever asking students to attempt to do so on their own (i.e., there are no Question videos in this part, therefore there is no constructive struggling, so Control students receive Passive Learning, PL). The Lessons are followed by practice problems (more Stories) for the students to implement what they just learned from the Lesson. In order to allow for comparisons between the Experimental condition students' first attempts at answering each Story, it is necessary to make the first practice problem in the Control condition an exact match for One-Unknown, Change, and Relational Stories in the Experimental condition. As such, the example problems whose solutions are demonstrated throughout the Lessons before those Stories are isomorphic to the Experimental Stories but with different numbers (and names of the restaurants).

Because the constructive struggling and extensive conditional branching of the Experimental condition inherently takes more time, the Control condition has additional practice for each problem type in order to equate time on task for the intervention. Repeated practice is in fact a common feature in traditional instruction anyway, so this is a sensible adjustment. And since exposure to multiple examples increases schematic learning of problem structures, it makes the Control a very strong comparison for the Experimental.

The Control condition is described in more detail in the sections below, but the scripts for all of the narration and its relation to onscreen text and animations is extensively documented for each video in Appendix B (Experimental scripts documented in Appendix A). An example of the Control condition, hosted on Zaption's website, is available at http://zapt.io/thcevgs8. Because these materials are highly interactive and contain a lot of conditional branching, it is recommended that readers explore the materials online, skipping backward to change answers to questions to see how this affects the progression through the materials (by conditional branching).

4.3.1 Control Condition Lessons

At the beginning of each section, there is an overview Lesson, which states in formal terms what is to be learned in the remainder of the section (FF). Then an example Story video is shown and the subsequent Lesson videos demonstrate how to find the unknown(s) in the example Story (PL). Each Story and its Lessons follow this same pattern, so I will just use the first Story as the primary example.

Control Story 1 (the One-Unknown problem to be solved, matched to Experimental) is prefaced by an overview Lesson that defines a variable using language common to math textbooks (FF). The narrator, reinforced by bullet points onscreen, makes statements such as "a variable is a symbol that represents a number", "the symbol is usually a letter like *x* or *y*", and "variables are helpful because they can represent a number we don't know yet". In other words, variables are formally defined upfront and it is stated why they are useful (rather than allowing students to discover this on their own, FF and PL). Then students are shown the example Story, in which the narrator buys three tacos for $9 at Maria's Restaurant (analogous to the matched One-Unknown).

The first Lesson video sets the goal to find the price of one taco at Maria's, stating that "we can use the letter *t* as a variable that represents the price of one taco." An animation is shown that is essentially a partial replication of how letter variables are introduced in the Experimental condition's final Lesson videos, but it is just done at the very beginning of the session

(Formalisms First) and with a one-unknown, one-equation example. The video then walks through how to use *t* to represent 'the price of 1 taco' and to use that to find that $t = \$3$.

The Lessons for Control Story 2 (Change) and Control Story 3 (Relational) are very similar in structure and format, but they show how to use two different letter variables and find their values using formal elimination (for Change) and substitution (Relational) strategies. These videos, like those for Control Story 1, contain very similar information and animations as the Walk-Throughs and Lessons from the Experimental condition, except that the Control videos use letter variables instead of word variables, define terms formally upfront (FF), and demonstrate solution strategies before allowing students to try to solve any problems on their own (PL).

4.3.2 Control Condition Problems and Repeated Practice

After the Lesson video, students are given two practice problems for that Control Story type, the first of which is identical to the ones used in the Experimental condition. This correspondence across conditions is done so that Experimental students' first attempts to solve a problem through unaided constructive struggling (before Hints or Walk-Throughs) can be compared directly to Control students' first attempts to enact the formal procedure they just learned with a very similar problem (that merely had different numbers). Because students have already seen an example isomorph in the Lesson, the Story video begins by clarifying that the prices at this second location are different than the previous place.

The Story videos' contexts are matched to Experimental in the sense that they show the narrator (or friend) purchasing the items at the restaurant, which is still likely to carry some contextual benefits for Control students. However, the Control condition does not use any of the additional Context videos implemented in the Experimental condition to demonstrate PRM details (Generic Context, GC).

There is minimal conditional branching in the Control condition: when a student gets a practice problem incorrect, they are shown an Explanation video which shows them how to solve it formally with variables. These Explanation videos are essentially streamlined versions of the Lesson videos that came before, in order to job the student's memory without going through every little detail. While certainly not all traditional educational materials have similar types of feedback, it is not uncommon for teachers to provide such reminders; as such, this is a fair feature to implement into the Control condition and it only makes it stronger.

The second practice problem of each problem type is yet another isomorph with different numbers (and a different restaurant name). In order to improve students' learning in the Control condition (and again, make it a stronger control), these second practice problems are scheduled in an *interleaved* manner (Bjork, 1994) so that another problem type's story appears in between each pair of same-type problems. The second practice problem for Control Story 1 (One-Unknown) appears just after the first problem for Control Story 2 (Change), which itself comes after the intervening Lesson for Control Story 2. Likewise, the second practice problem for Control Story 2 appears in between the two practice problems for Control Story 3 (Relational). This scheduling maximizes the time between the first and second practice for each problem type, a desirable difficulty (Bjork, 1994).

4.4 Session 1: Baseline Condition

The Baseline condition provides a way to assess whether the Experimental and Control conditions are statistically better than receiving no relevant training. Students in the Baseline condition thus watch videos from Khan Academy on how to find the least common multiple (LCM) and the greatest common factor (GCF) of sets of numbers. These videos were chosen because they have no direct connection to algebra or variables, but they are still part of the Common Core math standards for 6th grade (CCSS.MATH.CONTENT.6.NS.B.4), just as early understanding of algebraic expressions (CCSS.MATH.CONTENT.6.EE.A.2). Luckily, the three available LCM videos, two GCF videos, and one LCF and GCF word problem videos all combined to an appropriate amount of time. By pointing to each video through Zaption, I was able to add 2 practice questions for each of the 6 videos. Altogether, these videos and their practice questions equate well to the time on task for Experimental and Control conditions. The Baseline materials are available at http://zapt.io/teyhpyx2.

4.5 Session 1: Immediate Post-Test

Subjects in all three conditions (Experimental, Control, and Baseline) are given the same two immediate post-test problems (each with two questions). The first problem, Immediate Post-Test 1 (IPT1) is an analogue to the Change problem, but with the numbers, food items, and restaurant name all changed. The second problem, Immediate Post-Test 2 (IPT2) is an analogue to the Relational problem, also with all of the details changed.

Importantly, IPT1 in the Experimental condition contains some branching. If a student answers Question 1 incorrectly, they are then given a Hint, which just reminds them to think about how letter equations were used to find the price of one slice of cheese pizza in the Change Story Lesson. If the student gets Question 1 incorrect again, they are given a Walk-Through, which converts the two sentences into letter equations with Day 2's on top and the subtraction symbol to the left of Day 1, as if to cue them to perform subtraction for the elimination method. The purpose of this Hint and Walk-Through is to provide one last opportunity for consolidation in the Experimental condition. There are no other Hints or Walk-Throughs for IPT1's Question 2 or either Question for IPT2.

However, given this difference between the Experimental compared to the Control and Baseline conditions, it is only appropriate to make cross-condition comparisons for IPT1-Q1 answers prior to any Hints or Walk-Throughs, since Control and Baseline subjects do not receive such branching; moreover, it is also not appropriate to directly compare IPT1-Q2 across conditions because the extra assistance on IPT1-Q1 for Experimental students who answered Q1 incorrect initially would increase their likelihood of answering Q2 correctly. Direct comparisons across conditions for IPT2-Q1 and IPT2-Q2 are both permissible.

4.6 Session 2: Delayed Post-Test

The Delayed Post-Test (DPT; Appendix D) is given during a second session after at least a week has passed since the first. This post-test is administered on paper for convenience's sake, and consists of two near-transfer items, three far-transfer items, and nine conceptual items, as well as a brief survey about the first session.

4.6.1 Near-Transfer Items

There are two near-transfer items on the delayed post-test (DPT-N1 and DPT-N2; #s 1–2 on post-test). The first is an analogue of both the Change and IPT-1 problems. The second is an analogue of both the Relational and IPT-2 problems. These items are meant to see whether the knowledge from Session 1 was robust enough to transfer to novel examples after a delay. These items are the most critical measure of each condition's efficacy.

4.6.2 Far-Transfer Items

There are three far-transfer items on the delayed post-test (DPT-F1, 2, and 3; #s 3–5 on the post-test). The first (DPT-F1) is an analogue of the Change, IPT-1, and DPT-N1 problems, but with the roles of the variables and their coefficients reversed—that is, instead of knowing how many items were ordered each day and wanting to know how much they cost, in this problem, we know what each item type cost at each of two stores but not how many of each were purchased at each store (the same quantities were purchased both places, but the prices changed). This is meant to measure whether students' knowledge about the elimination method would extend to an example in the same general context (buying items at a store) where the mathematical information seems different, even though structurally, it is isomorphic.

The second far-transfer item (DPT-F2) is an analogue of the Relational, IPT-2, and DPT-N2 problems, but in purely symbolic algebra form (with x's and y's and no contextual details at all). This item is intended to measure whether the knowledge learned from contextual word problems transferred to symbolic forms.

The third and final far-transfer item (DPT-F3) is another analogue of the Change, IPT-1, and DPT-N1 problems, but it is a "puzzle problem" type (see 1.4.4.1) where the contextual details do not fit the PRM framework at all. Specifically, this is an adaptation of the farmer problem; typically, the farmer problem states the number of animals' heads and legs on a farm, where we know the heads represents the total amount of animals and two types of animals differ in the number of their legs (two vs. four). The problem content had to be changed away from legs in order to allow for the problem to fit the elimination-style pattern: you can't intuitively eliminate either animal type because the coefficients for heads are 1 and 1, and for legs are 2 and 4; in other words, both quantities change, so the elimination method won't work. Camels were chosen so that one-humped camels could stand in a 1:1 ratio with the number of their heads, allowing for them to be 'eliminated'.

4.6.3 Conceptual Items

There are nine conceptual items that were adapted from nine arithmetic and algebraic item scales and experimental stimuli. They are not expected to be directly influenced by the different conditions in the learning phase, but students who learn new conceptual information may still perform better on these items.

The first, third, and fifth items (DPT-C1, 3, and 5; #6, 8, 10 on post-test) are adapted from Falkner, Levi, & Carpenter (1999) and tests students' understanding that the equals sign represents equality of both sides (symmetrical), not a cue to carry out arithmetic (asymmetrical); DPT-C5 also tests students' willingness/ability to compare two variables' values without explicitly knowing them (relative judgments). The second item (DPT-C2; #7) is adapted from Linchevski & Herscovics (1996) and tests students' understanding that a letter variable must carry the same value in all instances within the same problem. The fourth item (DPT-C4; #9) is based on Collis' "Acceptance of Lack of Closure" (1975) and tests students' ability to isolate and solve for a variable that begins on both sides of an equation, which requires their ability to manipulate the variable across the equation while accepting they don't know its value. The sixth item (DPT-C6) is adapted from Linchevski & Herscovics (1996) and tests students' understanding of combining like terms as well as the 'generalizable number' conception of a variable. The seventh item (DPT-C7) is adapted from Weinberg et al. (2004) and tests students' understanding that variable symbols in an expression stand for numbers (not the objects themselves) and the result of an operation on them is also a number. The eighth item (DPT-C8) is adapted from MacGregor & Stacey (1997) and tests students' understanding of the object conception of an algebraic expression. The ninth item (DPT-C9) is adapted from Weinberg et al. (2004) and tests students' ability to relatively compare unknowns when no relational information is given, and to do so given a contextual problem (not symbolic).

4.6.4 Survey Items

The last page of the delayed post-test consists of six rating items and an open response for comments. The first four rating items allow four levels ("strongly disagree", "somewhat disagree", "somewhat agree", and "strongly agree") in order to discourage middling; the items ask students whether the videos were enjoyable (#1), relatable (#2), more helpful than other materials (#3), and that the website was easy to use (#4). They are then asked to indicate whether the video speed was "too slow", "just right", or "too fast", and whether the material covered was "too little", "just right", or "too much". Finally, they are given the chance to write down any additional thoughts.

5. Multimedia Experiment

5.1 Test Sites, Populations, and Administrations

The interactive multimedia video materials have been tested at four different school sites thus far, with more scheduled. At each site, all eligible students were invited to participate, receiving a packet to take home to their family that included a recruitment letter, parent permission form, and student assent form. Only students with signed parent permission forms and signed student assent forms were allowed to participate (or in the case of Test Site 2, emancipated youths and students aged 18 or older submitted their own consent forms). Participating students' identities were always protected by randomly assigning them a numerical code, which was used to match their answers from the online materials hosted on Zaption to their scratch paper and Delayed Post-Test.

5.1.1 Test Site 1

Test Site 1 (TS1) was a charter school located in Vacaville, California, that was in its third academic year of operation when tested. According to the school's most recent School Accountability Report Card (SARC) report, there are 468 total students enrolled from Kindergarten through 8th grade, with 58 students in the targeted 6th grade class. School-wide, 61.3% of students are white, 15.2% Hispanic or Latino, 9.4% two or more races, 4.5% Filipino, 2.8% African-American, 2.8% Asian, and 1.1% American Indian. Additionally, 9% of students are classified as socioeconomically disadvantaged and 5.6% of students have disabilities. The study was approved by the principal and middle school math teacher prior to student contact (the school is not part of a district).

Since all students in the middle school have the same math teacher (who had approved of the project), all 6th graders were invited to participate. In order to increase student involvement (as suggested by the teacher), students were compensated with $10 in Target gift cards for their participation in both sessions. Out of 58 students in the 6th grade, 25 submitted the appropriate documents to participate, but one student later decided not to participate (reason not given). In order to not take up class time, the teacher decided to run the experiment after school. Because of the impractical distance between the school and myself, both sessions were administered remotely by the students' math teacher.

The Delayed Post-Test (Session 2) was held either three or four weeks after the online portion (Session 1). Since the sessions were held after school, 6 of the 24 students ended up having a conflict with the three-week delay and so were tested exactly one week later. Of these six students who had a four-week delay, three were in the Experimental condition, two in Control, and one in Baseline. Of course, the condition assignments were randomized before Session 1, so this was purely due to chance.

A few small technical difficulties occurred during the computer portion of the study (Session 1), as reported by the teacher. First, the school network's firewall blocked the very last video (Immediate Post-Test 2, Question 2) for the Experimental condition. Despite the fact that the IPT videos are identical across conditions, one of these videos was blocked while the others were not. No discernible reason could be discovered. Second, the Change Story Walk-Through Q1 was given twice in a row by error; since this Question is only shown to subjects in the Experimental condition who get Q1 incorrect originally and after the Hint videos, this only affected a few students. For data purposes, only answers to the first presentation of the Question are examined, since the second is likely unreliable. These errors were fixed for future testing, and thankfully, none of them were significant enough to cause issues affecting interpretation. Because all of these issues added slight delays to the Experimental condition, if there were to be any effect (which I do not expect), it would only negatively affect Experimental subjects.

5.1.2 Test Site 2

Test Site 2 (TS2) was an alternative high school for at-risk youth in Santa Ana, California. It is part of the Orange County Department of Education's Alternative, Community, and Correctional Schools and Services (ACCESS) program. The school's population consists entirely of students who are referred to attend the school from either their previous school district or county services; many ACCESS students live in group homes or are homeless, are incarcerated or on probation, work full-time, or have children. The school does not have a traditional class structure; instead, students earn credits from attending classes and submitting assignments on an independent basis. The school works with students to get them to graduate high school so they are career- or college-ready. The study was approved by the principal and teachers prior to student contact (the school is not part of a district).

According to the school's teachers and administrators, while some students there have completed Algebra-level coursework, the majority of students are lower-performing in math. While this school is clearly different from the others, it was chosen in order to assess how well the materials work for nontraditional students for whom typical educational materials may not be effective. However, due to the fact that the students have had more exposure to algebra materials, it is expected that they will do better on average than other sites.

All students at the high school were invited to participate. Out of 50 enrolled students total, with approximately 35 on average attending each day, 19 submitted the appropriate documents to participate. Both sessions were administered during normal school hours in an extra room. The online session (Session 1) was administered by myself. The Delayed Post-Test (Session 2) was administered by teachers and was held one week after the online portion (Session 1). Two students were absent during both follow-up days but later completed Session 2 after a three-week delay (one student) or a four-week delay (the second); both were in the Control condition. Additionally, three of the 19 students dropped out of the study between the two sessions: one stated that they did not want to participate in Session 2, and two stopped attending the school. Of these three, two had been randomly assigned to Experimental and one to Control. All three have been dropped from all analyses.

5.1.3 Test Site 3

Test Site 3 (TS3) consisted of 6th-grade students from a public elementary school in Newport Beach, California. According to the school's most recent School Accountability Report Card (SARC) report, there are 620 total students enrolled from Kindergarten through 6th grade, with 110 students in 6th grade (although only one of three classes was targeted). School-wide, 70% of students are white, 12.4% Hispanic or Latino, 9.2% Asian, 6.3% two or more races, 1.1% African-American, 0.5% Filipino, 0.3% Native Hawaiian or Pacific Islander, and 0.2% American Indian. Additionally, 11.9% of students are classified as socioeconomically disadvantaged, 6.5% as English language learners, 8.5% as having disabilities, and 0.5% as being foster youths. The study was approved by the principal and the students' teacher prior to student contact.

All of the teacher's students were invited to participate. Out of 35 students in the class, 15 submitted the appropriate documents to participate. Both sessions were administered by the teacher and two assistants. All 15 students participated in both sessions. The study was conducted after school and off-campus at the teacher's request for practical reasons.

Unfortunately, there were intermittent internet connection issues for many students during the online portion (Session 1). Four of the five students in the Experimental condition and one of the five students in the Baseline condition were unable to finish the videos because of bad connectivity; all five Control subjects were able to finish, however. The disproportionate number of Experimental subjects with internet issues may have been caused by the fact that the Experimental condition has the most videos (because all versions of every video in each branch must be in the same "Lesson" page on Zaption). Regardless of the cause, only one student in the Experimental condition was able to finish, which severely limits the ability to analyze data for Test Site 3.

In order to use at least some data from the testing, we randomly selected one student from each of the Control and Baseline conditions to use in the analysis. This random selection is justifiable by virtue of the fact that, in essence, the one Experimental student who was able to finish was also "randomly selected" by internet connectivity from amongst the group of five Experimental subjects. Thus, only one student from each condition is analyzed from Test Site 3. Because this is clearly not sufficient to examine on its own, these data are only analyzed in the aggregate with all Test Sites combined. Such aggregation would not be statistically warranted if the full Control and Baseline data were also included, as it would introduce a significant confound between Condition and Site.

5.1.4 Test Site 4

Test Site 4 (TS4) consisted of 6th-grade students from a public middle school in La Habra, California. The study was approved by the district, principal, and two math teachers prior to student contact. All students in the two math teachers' 6th grade math classes were invited to participate. A total of 40 students (14 in one class, 26 in the other) submitted the appropriate documents to participate.

Unfortunately, severe internet connection issues were experienced for all students. Despite three separate attempts across a period of a few weeks, only 5 of the 40 students from Site 4 were able to complete the entire learning session. Since none of the 5 students who were able to complete the experiment had been randomly assigned to the Experimental condition, none of TS4's data are analyzed below (adding subjects from TS4 to Control and Baseline only would bias the results across conditions).

5.1.5 Test Site 5

Test Site 5 (TS5) will consist of 6th-grade students from a public middle school in San Diego, California. The study was approved by the district, principal, and math teacher prior to student contact. All students in the math teacher's three 6th grade math classes have been invited to participate. At least 40 students are expected to participate.

5.2 Methods

Within each test site, students were randomly assigned to one of the three conditions. Data from Session 1 (Learning Phase + Immediate Post-Test) were collected through Zaption, and data from Session 2 (Delayed Post-Test) were collected on paper. Open-response answers from Session 1 were hand-coded for correctness. All answers for Session 2 were also hand-coded, as well as the apparent solution strategies used by each student. This coding was performed by two condition-blind research assistants working independently; inter-rater disagreements were flagged for review, and all such disagreements were successfully resolved between the two research assistants. Coded solution methods included both 'formal' and 'informal' applications of the taught procedure (elimination for Change analogues and substitution for Relational), where 'formal' indicates the use of formal notation, and 'informal' involves carrying out the math entailed by the formal procedure without use of the notation. The other method codes for all problem types were 'Guess & Check' (if students guessed values for the unknown and checked mathematically to see if they worked in the problem), 'Other' (if the work shown did not relate to their given answer or the work was otherwise indiscernible), 'No Work' (if they simply answered without showing their work), and 'No Answer' (if they skipped the problem entirely).

Due to current limitations in Zaption's conditional branching capabilities, it was necessary to use multiple-choice options to perform the branching (this limitation is set to be changed soon by Zaption's development team). However, given the relative ease with which students can correctly guess multiple-choice by chance or work backward from the choices to check each one mathematically, open-response questions are much more preferable, especially for encouraging constructive struggling. Thus, a compromise was struck so that students first encountered a question in its open-response format (in which they can enter any text as their answer); then, after submitting their response, students were immediately given a multiple-choice question with the correct response and three incorrect lures, and they were asked to indicate which of the choices *matched* their answer in the open response (and if their open response answer was not listed, to choose a fifth option that stated that it didn't match). Unfortunately, students did not always report this match consistently with their open response, which resulted in some incongruous branching (disproportionately adversely affecting the Experimental condition due to its highly branching structure). It is possible that despite the instructions, students treated the multiple-choice question as a 'second chance' for answering a question correctly if they were not confident in their open-response answer, but of course, this is speculative.

For the analyses below, students' open responses were coded to assess whether they answered a question correctly; however, it is important to remember that the branching structure encountered by the student did not always match this (if they were not consistent between open response and multiple choice). In order to quantify to what extent each student chose the matching multiple-choice option, each open response was hand-coded for whether it matched the corresponding multiple choice selection. Then for each student, a total proportion of responses was calculated for how often they were consistent between open response and multiple choice, referred to as their 'match score' (e.g., a 'match score' of .85 indicates that the student's open response and multiple choice matched 85% of the time). This allows for analyses to be separated by students' matching proportions to see if there are resulting differences in outcomes. Since the total number of chances for agreement differ by condition (Baseline only has four such problems, the 4 Immediate Post-Test items; Control always has 14; and Experimental varies because of the branching), a median-split was calculated within each group, so that ‘High Match’ Experimental students could be compared to ‘High Match’ agreement Control, etc.

Students in Experimental and Control conditions are also analyzed according to whether their open-response and multiple-choice answers matched for a given problem type. For example, in looking at Experimental and Control students’ performance on a post-test analogue of the Change problem, students in each condition can be categorized based on whether their answers matched or not during the Change Story in the learning phase. Because the branching structures in the Change and Relational Stories are crucial to the Experimental condition, this allows for students who followed the intended pathway (i.e., ‘Matchers’) to be compared across the Experimental and Control conditions, and within each condition, relative to those who did not follow the intended pathway (i.e., those who ‘Switched’ answers from open response to multiple choice).

5.3 Experiment Results

In total, 43 subjects from Test Sites 1-3 are included in the analyses. A total of 55 additional subjects’ had to be excluded from analyses due to dropout (3/19 from TS2) or internet connection issues in Session 1 (12/15 from TS3; 40/40 from TS4). Of the 43 valid subjects, 24 were from TS1, 16 from TS2, and 3 from TS3. And of these analyzed subjects, 14 of them had been randomly assigned to the Experimental condition, 16 to Control, and 13 to Baseline.

5.3.1 Learning Phase

The mean time-on-task for Session 1 (Learning Phase + Immediate Post-Test) was 59.36 minutes for Experimental (*Median* = 59.5, *SD* = 9.84), 56.0 minutes for Control (*Median* = 56.0, *SD* = 13.86), and 58.08 minutes for Baseline (*Median* = 57.0, *SD* = 8.41). An ANOVA revealed no significant effect of condition on time-on-task, $F(2,40) = 0.35$, $p = .71$. Thus, any differences between conditions are indicative of differences in task content or structure, not the amount of time spent learning.

5.3.1.1 Conditional Branching Analysis (Experimental condition only)

The conditional branching structure of the Experimental condition allows for an in-depth look at students' paths through the material. Considering the complexity of the conditional branching in the Change (2.2) and Relational (3) Stories, it is worth examining the different branching pathways taken by students, as well as the performance of students taking different paths. However, given the inconsistencies in some students' open response and multiple choice answers (see 5.2), a branching analysis must consider both whether a student answered correctly on the open response (a more reliable measure of their ability) and also whether they provided the same response on the subsequent multiple choice question (which determined the conditional branching pathway they actually took).

In the flowchart pathway visualizations (Figures 6 and 7), we can see how students moved through the conditional branching paths; the charts only represent the Questions and whether students answered them correctly or incorrect (the Story, Context, Hint, and Walk-Through videos are not shown in order to reduce clutter, although see the full branching structure in the Appendix C for reference). These flowcharts also help to visualize where students answered both the open response and multiple choice correctly ('True Correct'), both incorrectly ('True Incorrect'), or switched answers between open response and multiple choice ('Switchers'). In the discussion below, True Correct and True Incorrect students may collectively be referred to as Matchers because their open-response and multiple-choice answers matched (thus, their branching was appropriate).

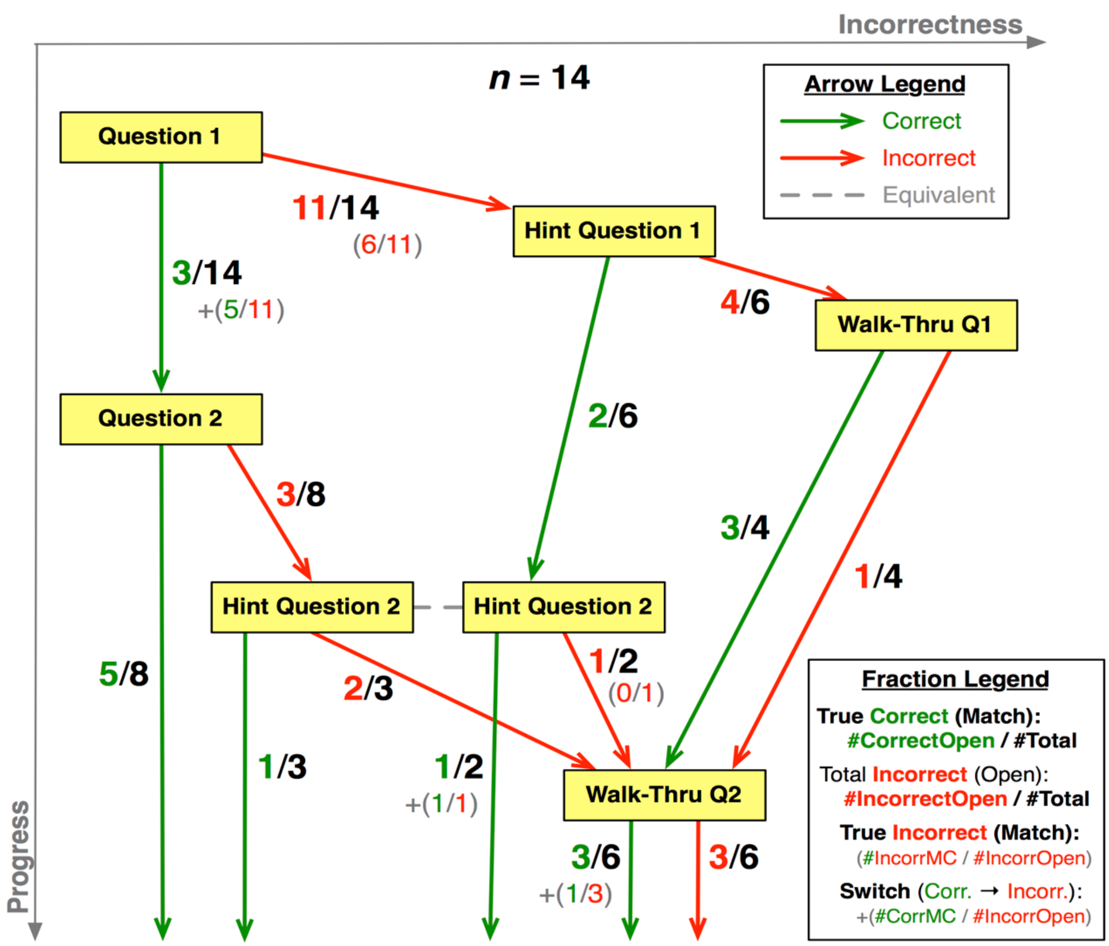


Figure 6. Flowchart visualizations of students' pathways through the conditional branching in the Change Story. Fractions in the large, bold font indicate the proportion of students who answered an open response correctly (green; 'True Correct') or incorrectly (red). Fractions in smaller font and parentheses preceded by a '+' indicate students added to a pathway by virtue of switching their multiple-choice response ('Switchers'); smaller-font fractions without the '+' indicate those who stayed in the appropriate pathway (i.e., did not switch; 'True Incorrect').

Looking at Figure 6, we can examine how students moved through the Change Story. For example, 3 out of 14 Experimental subjects correctly answered Question 1 (the price of 1 slice of cheese pizza) on the initial open response (i.e., they were True Corrects), although an additional 5 (among the 11 incorrect open responses) answered correctly on the multiple choice (i.e., there were 5 Switchers for Q1). All 3 of the incorrect responses to Question 2 from among the 8 True Corrects and Switchers (the leftmost pathway) came from the Switchers (i.e., 3/3 True Corrects and 2/5 Switchers answered Q2 correctly). Since Switching is determined by the student (not randomly assigned by the researcher), we cannot assess causation, but it may be true that Switchers would have benefitted from instead being branched appropriately (receiving Hints).

Overall, the most common pathway for the Change Story was answering both Q1 and Q2 correctly on the first attempt, if we include the Switchers on Q1; otherwise, the next most common pathway was to get Q1 incorrect twice before answering it correctly after the Walk-Throughs. It is possible that these two common pathways indicate two distinct subsets of students: those who intuitively are able to solve the problems on their own (either due to high pre-experiment ability or successful unaided constructive struggling) and those who benefit from the prompts. The size of this latter group, as visualized in the flowchart, gives a general sense of how effective the Hint and Walk-Throughs are for those students who could not solve the problem unaided. The Hints for Q1 helped a third (2/6) of the students who received it, while the Q1 Walk-Throughs helped 3/4 of those who didn't answer Hint Q1 correctly. Only 1 of the 14 Experimental students did not arrive at the correct answer for Q1 during the branching, which suggests a highly effective Hint and Walk-Through system. Looking at Q2, the Hints helped 2/5 total (collapsing across the two Q2 Hint pathways), while the Walk-Throughs helped an additional 3/6 (or 4/6 counting the one Switcher).

Looking at Figure 7, we see a broadly similar pattern of paths for the Relational Story, although with considerably fewer Switchers (only 2 total). Out of the 7 students who received the Hint (Matchers; does not include the one Switcher), the Hint was effective for 5 of them. Of the 2 remaining students, the Walk-Through (word equation practice and advanced hints) was not effective for either of them, although 1 Switched.

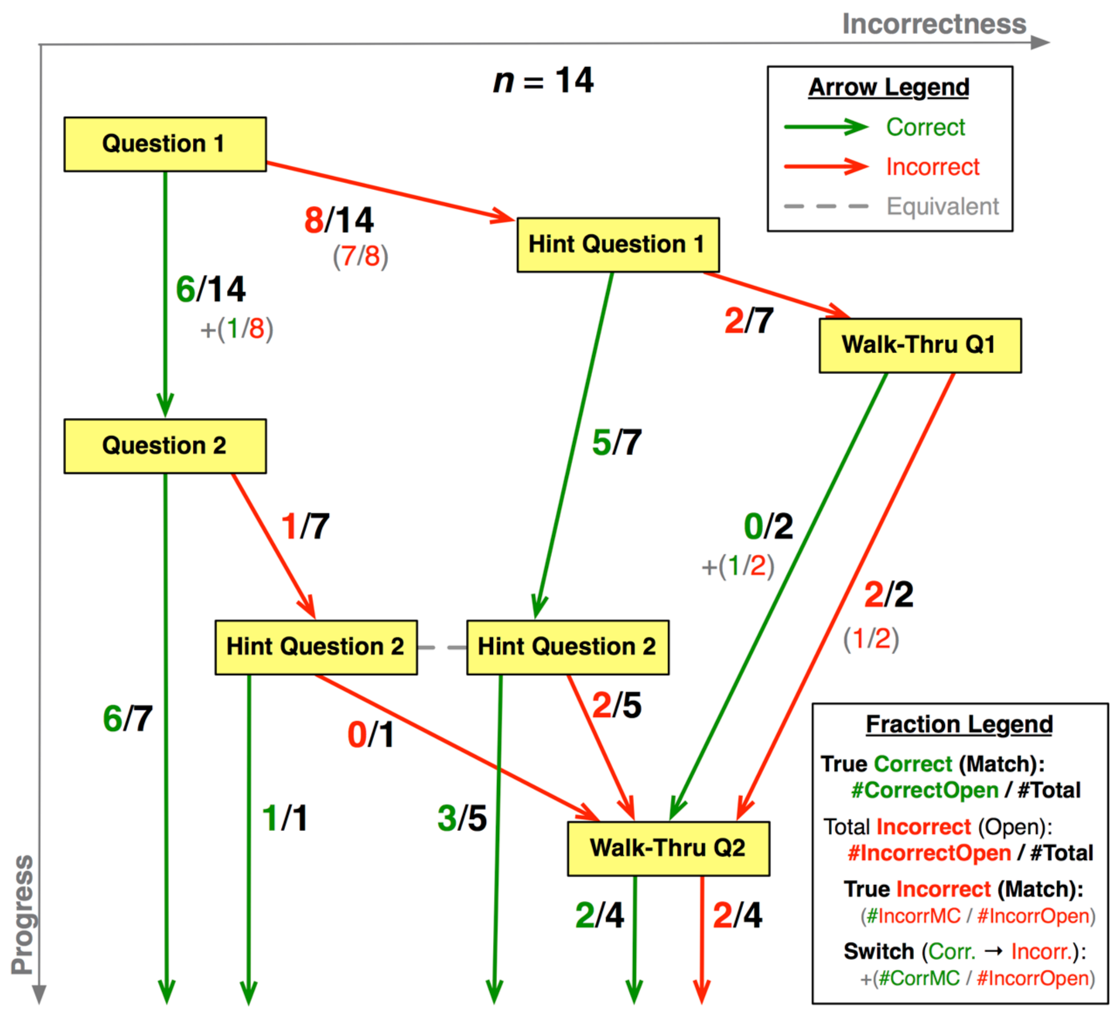


Figure 7. Flowchart visualizations of students' pathways through the conditional branching in the Relational Story. Fractions in the large, bold font indicate the proportion of students who answered an open response correctly (green; 'True Correct') or incorrectly (red). Fractions in smaller font and parentheses preceded by a '+' indicate students added to a pathway by virtue of switching their multiple-choice response ('Switchers'); smaller-font fractions without the '+' indicate those who stayed in the appropriate pathway (i.e., did not switch; 'True Incorrect').

Overall for the Relational Story, the modal pathway was Q1 Correct => Q2 Correct, while the next most common was Q1 Incorrect => Q1 Hint Correct => Q2 Hint Correct. As with the Change Story, this suggests the possibility of two distinct subsets of students: one who are able to solve the problems without assistance and one who are able to solve them with Hints. As with the Change Story, most students are able to arrive at the correct answer either on their own or with Hints (12/14 not counting the Switcher, or 13/14 counting them). This suggests that students in the Experimental condition do learn through this constructive struggling process, and the conditional branching structure is beneficial to those who need some assistance.

Taken together, the flowcharts for the Change and Relational Stories suggest that the Hints and Walk-Throughs are effective during the learning phase. However, it is clear that several students did not receive the appropriate conditional branching for their open response answer, which deprives these Switchers of the opportunity to learn in a more effective manner. As such, in analyzing the post-test data below, it is worthwhile to not only examine differences across conditions in the aggregate, but also to assess whether subjects at varying levels of 'matching' or branching correctness significantly differ on their post-tests.

5.3.2 Immediate Post-Test Analysis

The Immediate Post-Test (IPT) was given to all three conditions. Statistical comparisons of correct response proportions across the conditions were performed using two-tailed Fisher's exact test because some cell counts were too low to use a chi-square test. Due to the minor conditional branching in the Experimental condition's Change-analogue IPT for those who answer the initial Question 1 incorrectly, it is inappropriate to compare Experimental to the other conditions on the Change analogue Question 2.

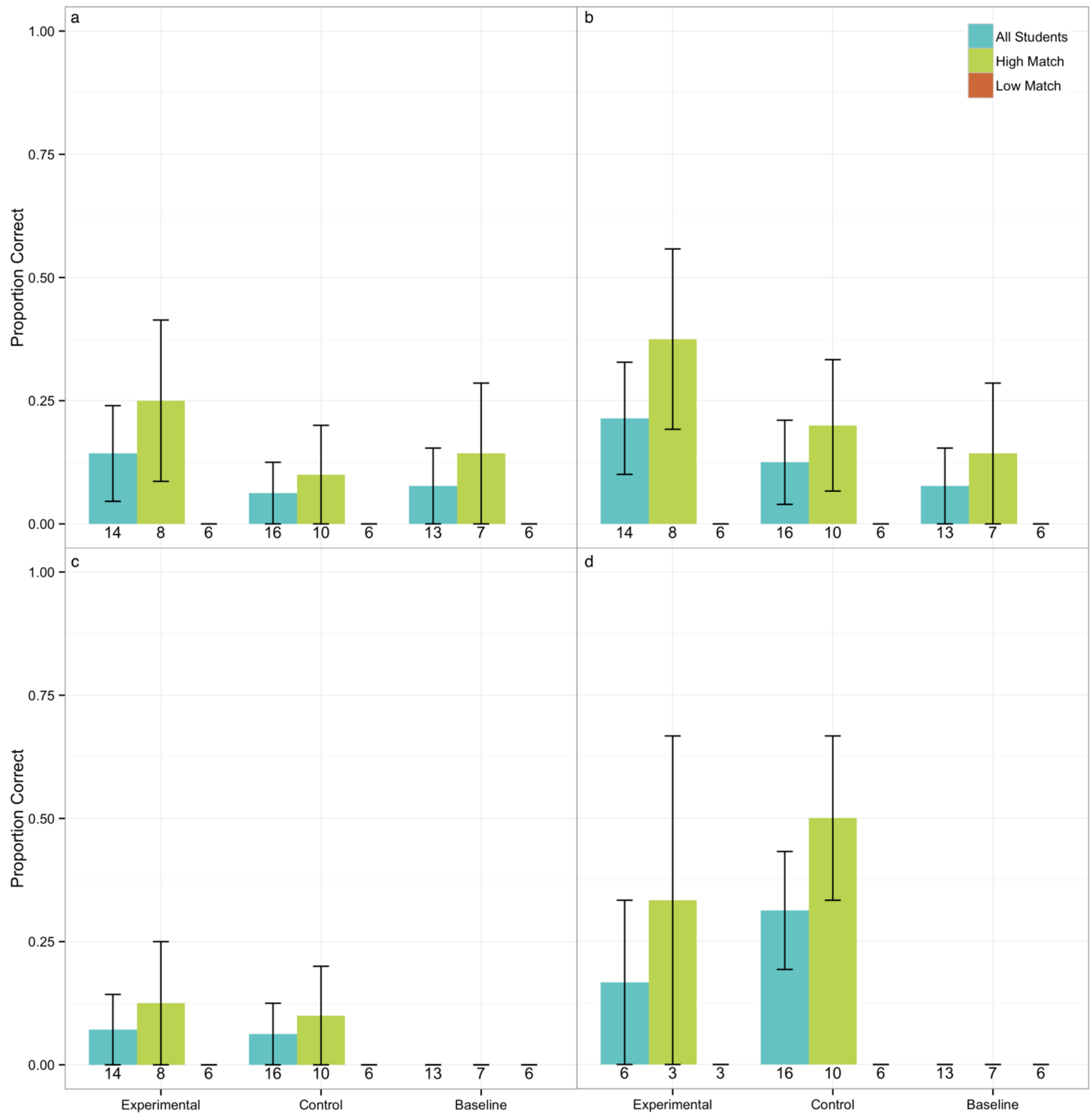


Figure 8. Immediate Post-Test results for each item: proportion correct for each condition, broken down by students' overall agreement between their open-response and multiple-choice answers throughout Learning Phase. 'High Match' students (green bars) were at or above their condition's median match agreement and 'Low Match' (red) were below; aggregate totals (blue) are also included for each condition. Error bars represent the standard error of the mean (SEM). Numbers below each bar represent the *n* for that group.

[a] Change analogue Question 1 (unknown changing quantities between equations) [b] Change analogue Question 2

[c] Relational analogue Question 1 (unknown that is multiple of other unknown) [d] Relational analogue Question 2

The proportion of correct responses on the IPT are shown in Figure 8, separated by Condition as well as by the median-split for students' 'match score'. There is an apparent floor effect throughout all Questions on the IPT, although less so for 'High Match' students. Most likely, this is due to the complexity of the numbers used for the problem (the total prices were \$20.45, \$39.45, and \$5.70, whereas multiples of 25¢ were used in the learning phase). These more difficult numbers were chosen in order to encourage students away from 'Guess and Check' procedures and toward attempting to implement the elimination/substitution procedures they had learned with word or letter equations (in Experimental and Control conditions, respectively). While student performance during piloting suggested these problems were still solvable by students in this age range, performance was dampened for students in the experiment. It is quite possible that this floor effect was also due at least in part to student fatigue, since the entire Session 1 (Learning + IPT) lasted an hour on average, and for some students, up to 75 minutes.

The analyses revealed no significant effect of condition on any the IPT questions (all comparisons made collapsed across match score). However, there was a marginally significant effect of condition for the Relational analogue's Question 2, $p = .08$, with the Control condition performing best (31.3% correct), followed by Experimental (16.7%), and then Baseline (0.0%). Most importantly, the Baseline condition performs poorly across all Questions, with only one subject in Baseline getting each of the Change analogue Questions correct but no Baseline subject getting either Relational Question correct. The lack of significant differences between the Experimental and Control conditions is not surprising; while it would actually be reasonable for Control conditions to outperform Experimental on the Immediate Post-Test (since Control receives more practice), such a pattern is not apparent here: on 2 of the 3 comparable items (Change analogue Q2 cannot be compared due to Experimental's branching), Experimental outperforms Control, although none of these differences are significant (but Control is marginally significantly better than Experimental on the Relational analogue Q2).

We can also examine students' conditional branching pathways within the Change

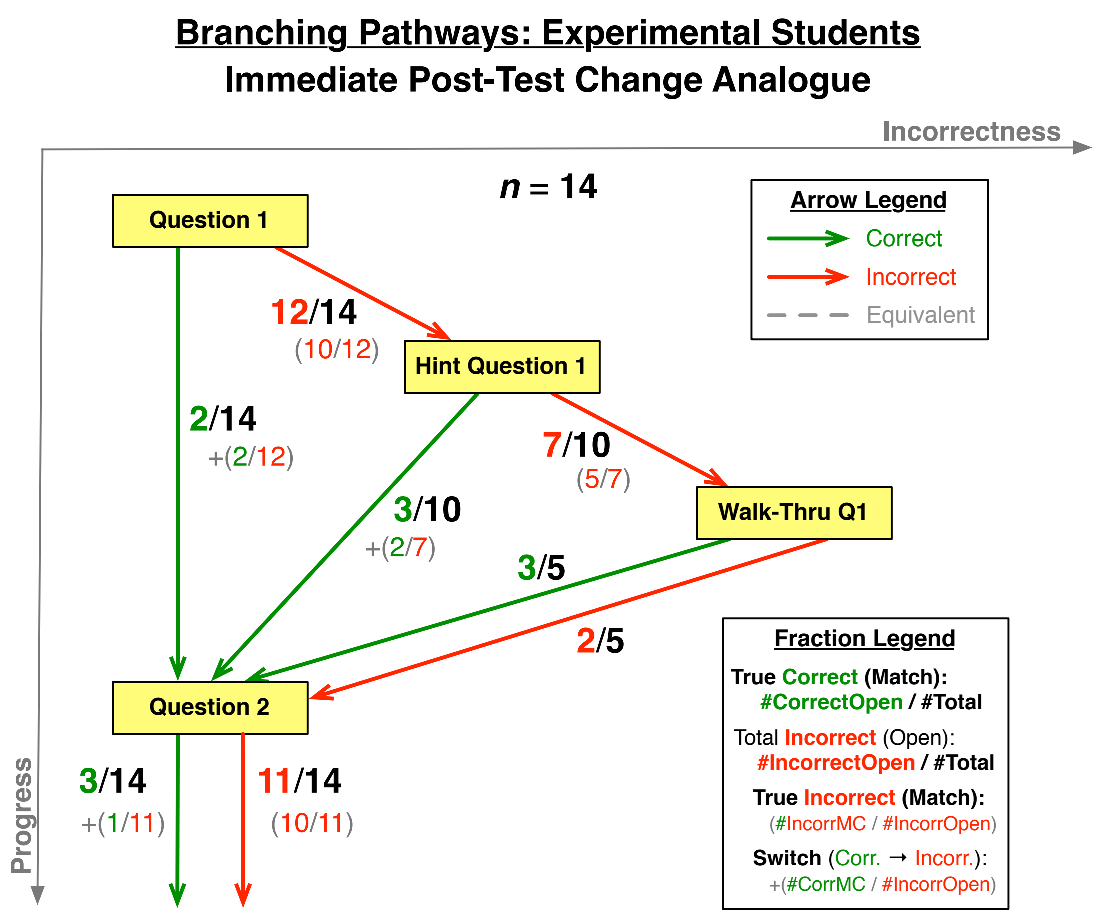


Figure 9. Flowchart visualizations of students' pathways through the conditional branching in the Immediate Post-Test Change analogue. Fractions in the large, bold font indicate the proportion of students who answered an open response correctly (green; 'True Correct') or incorrectly (red). Fractions in smaller font and parentheses preceded by a '+' indicate students added to a pathway by virtue of switching their multiple-choice response ('Switchers'); smaller-font fractions without the '+' indicate those who stayed in the appropriate pathway (i.e., did not switch; 'True Incorrect').

analogue IPT problem in Figure 9. Specifically, while only 2/14 answer the first Question correctly (and another 2 students Switch into that path), only 2 students fail to arrive at the correct answer by the end of the Walk-Through. As with the Learning Phase branching analysis, this suggests that the Hints and Walk-Throughs in this Change analogue are effective in leading students toward finding the answer.

It is also worth looking at how students in the Experimental condition who needed the full extent of the branching during the Learning Phase (i.e., used both the Hint and Walk-Through) went on to perform on the IPT for the analogous problem. However, given the rather low success rate on the IPT overall, it is difficult to assess any patterns. For example, of the 4 students who could not solve the Change problem during the Learning Phase even after the Hint, 3 of them answered the Change problem correctly after the Walk-Through, but none of those students went on to solve the Change analogue IPT problem correctly (although 2 of the 4 Switched before the Hint and another 1 after the Hint). Thus, given the current low number of subjects and the issues with Switching, it is too difficult to assess.

Considering the issues with Switching and the incongruous branching that it causes, it is worth examining how 'High Match' students (whose overall conditional branching during learning was more appropriate to their understanding) compare to their 'Low Match' counterparts (whose overall conditional branching, especially in the Experimental condition, did not give them the pedagogical support they should have received). For example, in Figure 8, we can see that among Experimental subjects, 25% of the 8 students at or above the median match score during the Learning Phase went on to answer correctly on the Change analogue IPT Question 1, while none of the 6 below the median did so. In fact, across all conditions, no student in the lower half of match scores got a single IPT item correct. Of course, because

Switching has only been observed among students moving from an incorrect open-response answer to a correct multiple-choice answer (and not the other way around), it is empirically the case for at least the present data that Switching would be somewhat negatively correlated with correctness, since all observed switches began as incorrect responses. However, this does not necessitate that performance should also be poor on a subsequent transfer task, since it is conceivable that Switchers had nonetheless learned from their initial incorrect response. Thus, it is still informative to see that students with poor matching during the learning phase do considerably more poorly on the transfer. It is possible that these students worked backwards from the multiple choice options to find the correct response to switch to, and in doing so, they were then deprived of the conditional branching that would have encouraged them to think more intuitively in an algebraic manner; thus, these students did not receive the intended benefits of the Experimental materials.

Another way to examine the relationship between branching and post-test performance is by comparing how students were branched on the first Question for the Change and Relational items to their eventual performance on the first Question of the IPT analogue of the respective problem type. We can categorize students' Q1 branching into the three groups described above (in 5.2): True Correct, True Incorrect, and Switchers. For the Change Story (see Figure 6), there were 3 True Correct students, 6 True Incorrect, and 5 Switchers. Of these, 2 of the 3 True Correct students answered the analogous IPT problem correctly as well, while none of the True Incorrect or Switchers did so. And in looking at the Relational Story (see Figure 7), 1 of the 6 True Correct students answered the IPT analogue correctly, while none of the 7 True Incorrect or 1 Switcher did. Taken together, these patterns suggest that there is likely a real distinction between the True Correct students and the Switchers, either in the strategies they employ in

Table 3

Matching Analysis: Immediate Post-Test Performance Based on Analogous Problem Matching During Learning Phase

	Experimental				Control			
	Matchers[a]		Switchers		Matchers[a]		Switchers	
Problem	Correct	*n*	Correct	*n*	Correct	*n*	Correct	*n*
Change Analogue Q1	22.2%	9	0%	5	7.7%	13	0%	3
Change Analogue Q2	22.2%	9	0%	5	15.4%	13	0%	3
Relational Analogue Q1	7.7%	13	0%	1	6.7%	15	0%	1
Relational Analogue Q2	20.0%[b]	5[b]	0%	1	33.3%	15	0%	1

Note. For each Immediate Post-Test problem (Change and Relational), students are classified as Matchers or Switchers based on whether their open-response and multiple-choice answers matched or were switched on the analogous problem during learning.

[a] 'Matchers' include both 'True Correct' and 'True Incorrect'. [b] No data could be collected for Relational Analogue Q2 from Experimental subjects at Test Site 1 due to an internet firewall issue for only that video.

answering questions, the learning that they received from the conditionally branched materials, or some mixture thereof.

We can also combine True Correct and True Incorrect into a Matcher group, all of whom received the intended branching for the analogous problem during learning. Matchers and Switchers from the Experimental and Control conditions are compared in Table 3. This allows us to more directly compare the effects of each condition's learning in the analogous problem, because if there are any non-experimental differences between Matchers and Switchers (e.g., Switchers may be paying less attention), these non-experimental qualities are expected to be equivalent across conditions due to random assignment. As such, comparing Experimental Matchers to Control Matchers provides a clearer picture of the difference between conditions under appropriate branching. None of the Switchers for either problem type in either condition answered any of the four IPT Questions correctly. Importantly, even with an apparent floor effect, a higher proportion of Experimental Matchers than Control Matchers provided correct answers for 3 of the 4 IPT Questions; the only Question for which this pattern did not hold was the Relational Story analogue's Q2, in which the 8 Experimental subjects from TS1 were not able to view the Question due to a computer error. While this pattern is not significant due to the low sample size, it is suggestive that with more students and less complex arithmetic, a clearer advantage of Experimental over Control may be apparent even on the IPT.

5.3.3 Delayed Post-Test

The mean number of days in the delay between Session 1 and Session 2 was 16.6 for Experimental (*Median* = 20, *SD* = 7.72), 18.69 for Control (*Median* = 20, *SD* = 6.36), and 14.31 for Baseline (*Median* = 12, *SD* = 6.98). Since the distribution of delays is bimodal (because the sites differed in the amount of time between sessions), a two-tailed nonparametric Kruskal-Wallis test was used, which showed that the distributions of delays did not significantly differ between condition, $\chi 2(2) = 3.05$, $p = .22$. Thus, any differences between conditions are indicative of differences in task content or structure, not differential amounts of time that elapsed between sessions.

The proportion of correct responses on the Delayed Post-Test (DPT) are shown in Figure 10 (for near-transfer items) and Figures 11–12 (for far-transfer), separated by Condition as well as by the median-split for students' match score. Statistical comparisons across the conditions (collapsed across match scores) were performed using Fisher's exact test.

5.3.3.1 Near-Transfer Items on the Delayed Post-Test

The results from each Near-Transfer DPT item are shown in Figure 10. There was a significant effect of condition for the first question on the DPT Change analogue, $p = .008$. Post-hoc tests with Fisher's exact test revealed that Experimental students (64.3% correct) were significantly better than Baseline (7.7%), $p = .004$. While Experimental students were not significantly better than Control, $p = .14$, they answered correctly more than *twice* as often (64.3% vs. 31.3%). Though there were no significant effects of condition for any of the other three near-transfer DPT questions (all ps > .2), the Experimental condition's students answered correctly at the highest rate for every single question, followed by Control, with Baseline at the lowest. Even more so than the IPT items, there is a clear trend in the current data with Experimental outperforming Control, which outperforms Baseline, but more subjects are needed in order to see if the trend holds up.

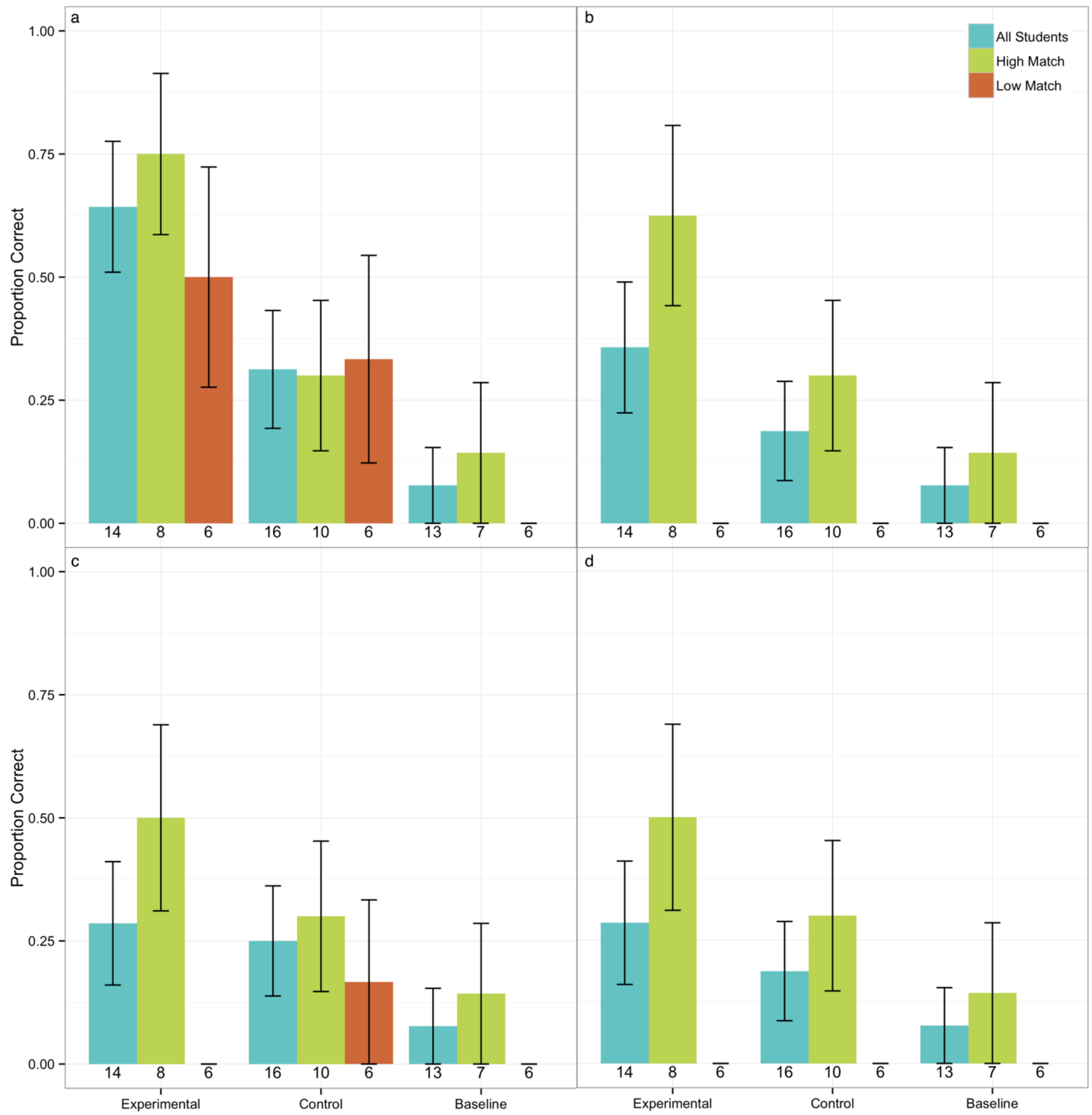


Figure 10. Delayed Post-Test results for each Near-Transfer item: proportion correct for each condition, broken down by students' overall agreement between their open-response and multiple-choice answers throughout Learning Phase. 'High Match' students (green bars) were at or above their condition's median match agreement and 'Low Match' (red) were below; aggregate totals (blue) are also included for each condition. Error bars represent the standard error of the mean (SEM). Numbers below each bar represent the *n* for that group.

[a] Change analogue Question 1 (unknown changing quantities between equations) [b] Change analogue Question 2

[c] Relational analogue Question 1 (unknown that is multiple of other unknown) [d] Relational analogue Question 2

We may also consider the median-split match scores from the learning phase with these DPT items to see whether the same pattern observed with the IPT items holds up after a delay. Looking at the near-transfer DPT items (Figure 10), in all but one case, the High Match subset of a condition had better performance than the Low Match (the exception being the Change analogue Question 1 in the Control condition, in which the Low Match just barely exceeded the High Match). So again, there seems to be an overall pattern in which students with higher overall matching during learning answered more questions correctly, even after a delay. Moreover, within the High Match subset of each condition, there is a very stark pattern of Experimental > Control > Baseline. For example, looking at the Change analogue Question 1 (which had a significant effect of condition across all subjects), the differences are even greater among the High Match students, with Experimental students answering correctly 2.5 times as often as Control (75% vs. 30%), and over 5 times as often as Baseline (75% vs. 14.3%), although of course the sample sizes are lower for this subset than the aggregate.

The branching categorization we used for comparing Experimental students' learning phase branching to IPT scores can also be applied to DPT. In comparing across Change-type problems, 3 out of the 3 True Correct students from the Change Story in the learning phase went on to answer the analogous Change problem correctly after a delay, while 3 out of 6 True Incorrect and 3 out of 5 Switchers did the same. For the Relational-type problems, 2 out of 5 of the True Correct students from the Relational Story in the Learning Phase went on to correctly answer the analogous Relational problem on the DPT, while 2 of the 7 True Incorrect was correct and the 1 Switcher was incorrect.

As with the IPT, we can also compare across the Experimental and Control conditions in terms of Matchers vs. Switchers on the learning problem analogous to each DPT problem (see

Table 4

Matching Analysis: Delayed Post-Test Near-Transfer Performance Based on Analogous Problem Matching During Learning Phase

	Experimental				Control			
	Matchers[a]		Switchers		Matchers[a]		Switchers	
Problem	Correct	*n*	Correct	*n*	Correct	*n*	Correct	*n*
Change Analogue Q1	66.7%	9	60.0%	5	30.7%	13	33.3%	3
Change Analogue Q2	44.4%	9	20.0%	5	23.1%	13	0%	3
Relational Analogue Q1	30.7%	13	0%	1	26.7%	15	0%	1
Relational Analogue Q2	30.7%	13	0%	1	20.0%	15	0%	1

Note. For each Delayed Post-Test problem (Change and Relational), students are classified as Matchers or Switchers based on whether their open-response and multiple-choice answers matched or were switched on the analogous problem during learning.

[a] 'Matchers' include both 'True Correct' and 'True Incorrect'.

Table 4), which again lets us more directly compare the effects of each condition. First, across each problem type and each condition, in only 1 of the 8 comparisons did Switchers outperform Matchers (Control condition's Change analogue Q1), and the difference is small and the sample size for Control Switchers is quite low (3). Most importantly, across all 4 near-transfer DPT Questions, the Experimental Matchers outperform the Control Matchers, again suggesting an overall benefit to the Experimental learning for those who are correctly branched.

It is also possible to examine, to some extent, the efficacy of the Hints and Walk-Throughs from the Experimental condition's Learning Phase in terms of transfer on the DPT. Of particular interest is whether students who correctly solved the original Change problem only after a Walk-Through (not originally) were better on the DPT analogue of the Change problem than their counterparts who saw the Walk-Through but did not answer it correctly then. Since all students in the Experimental condition were eventually shown the explanation of how to intuitively solve a Change-type problem, this comparison will give a general sense of the 'added benefit' from the Walk-Through cuing the student's thought in the right direction. Of the 3 students who answered the Change problem correctly from the Walk-Through's advanced hints, 2 of them went on to correctly answer the analogous problem on the delayed test. By contrast, the 1 student who still did not get the Change problem after the Walk-Through also did not get its analogue correct after a delay. However, more data is needed to see if this pattern holds up.

Lastly, we can perform an analysis of the methods used by students in solving these near-transfer DPT items, particularly among High Match students and for those who were Matchers for the analogous problem during the Learning Phase. For the Change analogue on the DPT (Table 5), 6 out of the 8 High Match students used Informal Elimination (carried out the math entailed by elimination without using formal notation) and answered correctly, whereas only 2 of

Table 5

Method Analysis: Delayed Post-Test Change Analogue (High-Matchers Only)

Method Used	Experimental (n = 8)		Control (n = 10)		Baseline (n = 7)	
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Formal Elimination[a]	0%	0%	0%	0%	0%	0%
Informal Elimination[b]	75.0%	12.5%	20.0%	10.0%	0%	0%
Guess & Check[c]	0%	0%	0%	0%	0%	0%
Other[d]	0%	12.5%	0%	0%	0%	42.9%
Answered, No Work[e]	0%	0%	10.0%	50.0%	14.3%	28.6%
No Attempt[f]	—	0%	—	10.0%	—	14.3%

Note. Solution methods were coded by two independent, condition-blind research assistants.

[a] Used the elimination method on formal symbols (e.g., letter or word variables). [b] Used the same mathematical operations but without formal symbols. [c] Guessed numbers and checked mathematically to see if they worked. [d] Unrelated or indiscernible work. [e] Provided an answer without showing any work. [f] Made no attempt at answering (no answer or work).

10 Control High Match students did the same, and none of the 7 Baseline did. This suggests that not only were Experimental students more likely to answer correctly, but they were also more likely to do so using the taught procedure, despite the fact that the Control students actually had more practice using the procedure. Additionally, 1 High Match student in each of the Experimental and Control conditions attempted to use Informal Elimination but did so incorrectly. Collapsing across correct and incorrect uses, there was a significant effect of condition on use of Elimination (Formal and Informal), p = .001. Pairwise post-hoc tests

revealed that Experimental students were significantly more likely to use Elimination (7/8) than both Control (3/10, $p = .02$) and Baseline (0/7, $p = .001$).

Looking only within the Experimental condition's students based on their branching on the Change problem during learning, all 3 True Correct students went on to correctly answer the Change analogue on the DPT using Informal Elimination, while 3 of 6 True Incorrect students did the same, and only 1 of 5 Switchers did. This pattern suggests that the True Incorrect students benefitted from the branching to better learn how to solve the problem type, relative to the Switchers who did not receive the appropriate Hints and Walk-Throughs.

Examining students' methods used in solving the Relational analogue on the DPT (Table 6), 1 out of 8 High Match students in the Experimental condition used Informal Substitution to correctly find the answer, while 1 of 10 for Control did the same, and none of the 7 Baseline students did. It is possible that Experimental students did not use the Substitution procedure on the Relational analogue as often as they used Elimination on the Change analogue because they spent less time overall on navigating the Relational problem structure during the Learning Phase (since there are multiple contrast comparisons for the Change problem, but not the Relational). Moreover, Substitution is likely to be a less intuitive procedure (which is why it is logically sequenced after Elimination, for which students' implicit causal learning process can be evoked when considering change). Additionally, one High Match student in each condition divided by 3 instead of 4, suggesting that they may have attempted Informal Substitution but used the coefficient of the unknown being substituted in instead of the subsequent total; adding more feedback and practice during learning may help students avoid this mistake in the future. Three out of 7 Baseline High Match students simply divided the total cost by 2, ignoring the multiplicative relation between item costs altogether; none of the High Match students in

Table 6

Method Analysis: Delayed Post-Test Relational Analogue (High-Matchers Only)

Method Used	Experimental (*n* = 8)		Control (*n* = 10)		Baseline (*n* = 7)	
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Formal Substitution[a]	0%	0%	0%	0%	0%	0%
Informal Substitution[b]	12.5%	0%	10.0%	0%	0%	0%
Divided Total by 2[c]	—	0%	—	0%	—	42.9%
Divided Total by 3[d]	—	12.5%	—	10.0%	—	14.3%
Guess & Check[e]	25.0%	12.5%	10.0%	0%	14.3%	0%
Other[f]	0%	12.5%	0%	10.0%	0%	0%
Answered, No Work[g]	12.5%	0%	10.0%	30.0%	0%	28.6%
No Attempt[h]	—	12.5%	—	20.0%	—	0%

Note. Solution methods were coded by two independent, condition-blind research assistants.

[a] Used the substitution method on formal symbols (e.g., letter or word variables). [b] Used the same mathematical operations but without formal symbols. [c] Divided total cost by 2 (correct substitution would have divided by 4). [d] Divided total cost by 3. [e] Guessed numbers and checked mathematically to see if they worked. [f] Unrelated or indiscernible work. [g] Provided an answer without showing any work. [h] Made no attempt at answering (no answer or work).

Experimental or Control condition made this mistake, indicating they may be more likely to pay attention to the multiplicative relation between the item costs, even if they do not always know how to use that relation to solve the problem (e.g., when they divide by 3 or use 'Guess & Check' instead). There was no significant effect of condition on Substitution use.

Importantly, Experimental High Match students were still more likely to answer the Relational analogue correctly, but they more often did so via 'Guess & Check': (2/8

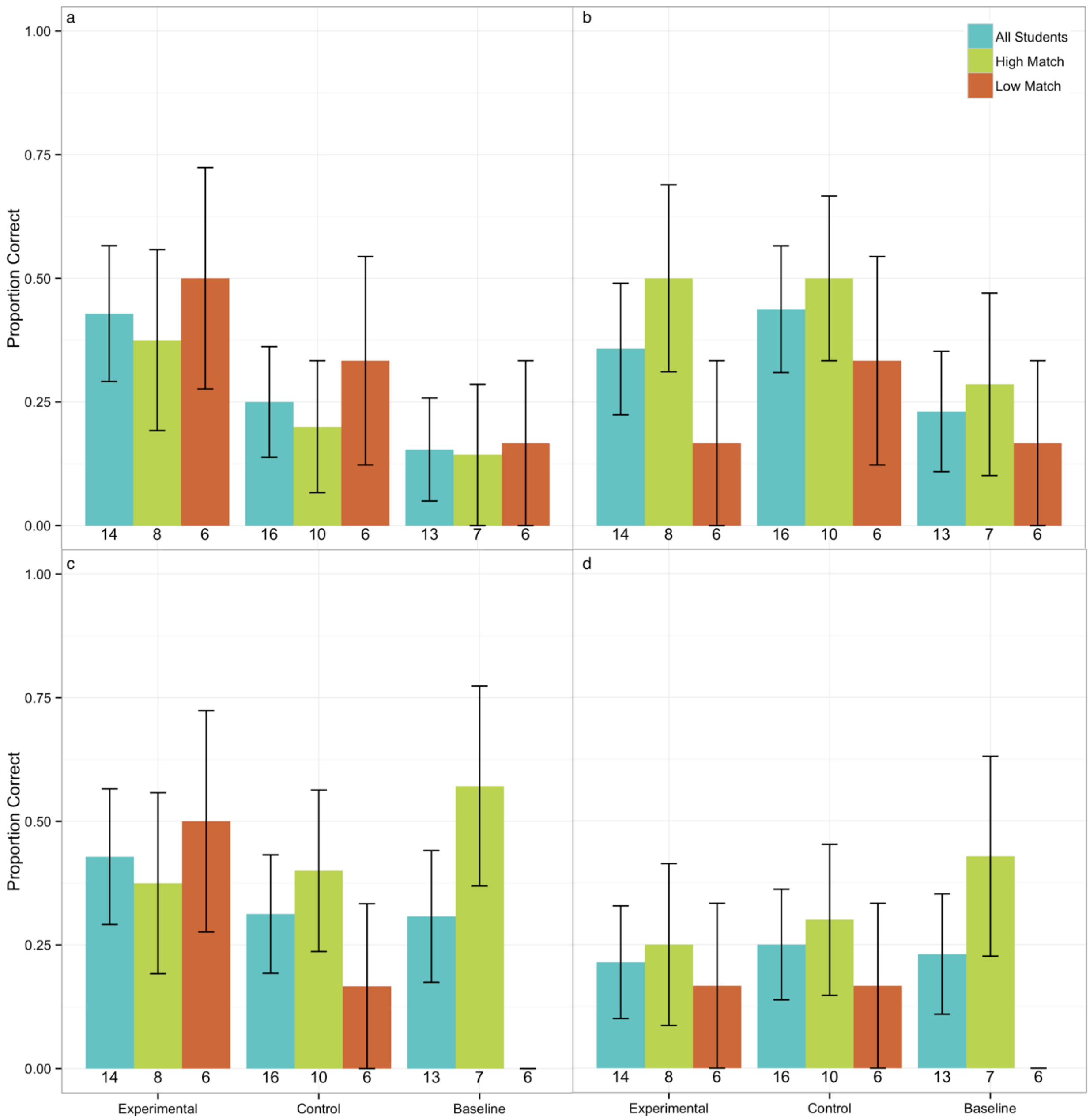


Figure 11. Delayed Post-Test results for the first two Far-Transfer problems' items: proportion correct for each condition, broken down by students' overall agreement between their open-response and multiple-choice answers throughout Learning Phase. 'High Match' students (green bars) were at or above their condition's median match agreement and 'Low Match' (red) were below; aggregate totals (blue) are also included for each condition. Error bars represent the standard error of the mean (SEM). Numbers below each bar represent the *n* for that group.

[a] Change Role Reversal Question 1 (item quantities unknown, prices known) [b] Change Role Reversal Question 2

[c] Relational Symbolic Question 1 (formal symbolic algebra, no context) [d] Relational Symbolic Question 2

Experimental vs. 1/10 Control and 1/7 Baseline). This suggests that Experimental students were more comfortable with the mathematical structure, even if they did not cue the Substitution procedure. While of course more data is necessary to see if these patterns in solution procedures hold up, early indications are that Experimental students are more likely to use the procedure they learned gradually and intuitively than Control students are to use the same procedure taught formally upfront, despite the latter group's additional practice.

5.3.3.2 Far-Transfer Items on the Delayed Post-Test

Among the 6 total far-transfer items on the DPT (3 problems with 2 questions each; Figures 11–12), there were no significant effects of condition (all *p*s > .2). However, there was still somewhat of a trend (although less so than with IPT and the near-transfer DPT items) in favor of Experimental > Control > Baseline. Of these 6 far-transfer items, 3 had this exact pattern, while 3 others did not.

For the Role-Reversal Change analogue (wherein students were given the items' prices but not the quantities purchased), the Experimental students outperformed Control and Baseline for Question 1 (Figure 11a), but not for Question 2 (in which Control did the best; Figure 11b). Only one student overall recognized that Informal Elimination could be used to solve the problem, which they did correctly; that student came from the Experimental condition and was both a High Match and had answered the Learning Phase's Change problem correctly (True Correct). Other Experimental High Match students answering correctly used Guess and Check (Table 7), again suggesting more comfort with the problem type even if the Elimination procedure was not cued. Looking at Experimental students' performance on the Change problem during the Learning Phase, Switchers again performed worse than Matchers on this DPT

Table 7

Method Analysis: Delayed Post-Test Change Role-Reversal (High-Matchers Only)

	Experimental (n = 8)		Control (n = 10)		Baseline (n = 7)	
Method Used	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Formal Elimination[a]	0%	0%	0%	0%	0%	0%
Informal Elimination[b]	12.5%	0%	0%	0%	0%	0%
Guess & Check[c]	25.0%	25.0%	10.0%	20.0%	14.3%	14.3%
Other[d]	0%	12.5%	0%	20.0%	0%	28.6%
Answered, No Work[e]	0%	0%	10.0%	20.0%	0%	28.6%
No Attempt[f]	—	25.0%	—	20.0%	—	14.3%

Note. Solution method were coded by two independent, condition-blind research assistants.

[a] Used the elimination method on formal symbols (e.g., letter or word variables). [b] Used the same mathematical operations but without formal symbols. [c] Guessed numbers and checked mathematically to see if they worked. [d] Unrelated or indiscernible work. [e] Provided an answer without showing any work. [f] Made no attempt at answering (no answer or work).

analogue, with 4 out of 6 True Incorrect students responding correctly, 1 out of 3 True Correct, and 1 out of 5 Switchers. This pattern again indicates that appropriate branching may support students' learning better.

For the Symbolic Relational analogue (using formal algebra notation and no context), there was a slight trend among all students of Experimental > Control > Baseline (Figure 11cd), although Baseline High Match students performed the best out of all subsets (suggesting those randomly assigned to Baseline may have had the most prior experience with formal algebra

Table 8

Method Analysis: Delayed Post-Test Relational Symbolic (High-Matchers Only)

	Experimental (n = 8)		Control (n = 10)		Baseline (n = 7)	
Method Used	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Formal Substitution[a]	12.5%	0%	0%	0%	14.3%	0%
Informal Substitution[b]	0%	0%	10.0%	0%	0%	0%
Guess & Check[c]	12.5%	12.5%	0%	0%	14.3%	0%
Other[d]	0%	0%	0%	0%	0%	0%
Answered, No Work[e]	12.5%	37.5%	30.0%	40.0%	28.6%	42.9%
No Attempt[f]	—	12.5%	—	20.0%	—	0%

Note. Solution methods were coded by two independent, condition-blind research assistants.

[a] Used the substitution method on formal symbols (e.g., letter or word variables). [b] Used the same mathematical operations but without formal symbols. [c] Guessed numbers and checked mathematically to see if they worked. [d] Unrelated or indiscernible work. [e] Provided an answer without showing any work. [f] Made no attempt at answering (no answer or work).

notation, since they received no exposure to algebra during learning). Among High Match students (Table 8), 1 in each of the Experimental and Baseline conditions correctly answered Question 1 using Formal Substitution, while 1 Control student answered correctly with Informal Substitution. Many High Match students did not show any work at all (sometimes still answering correctly), indicating that they may not be comfortable yet with manipulating formal algebraic symbols, so they could have guessed and checked using mental math instead.

Students overall performed poorly on the final DPT item, the ‘Puzzle Problem’ Change analogue (Figure 12), with barely any discernible pattern between conditions across the two

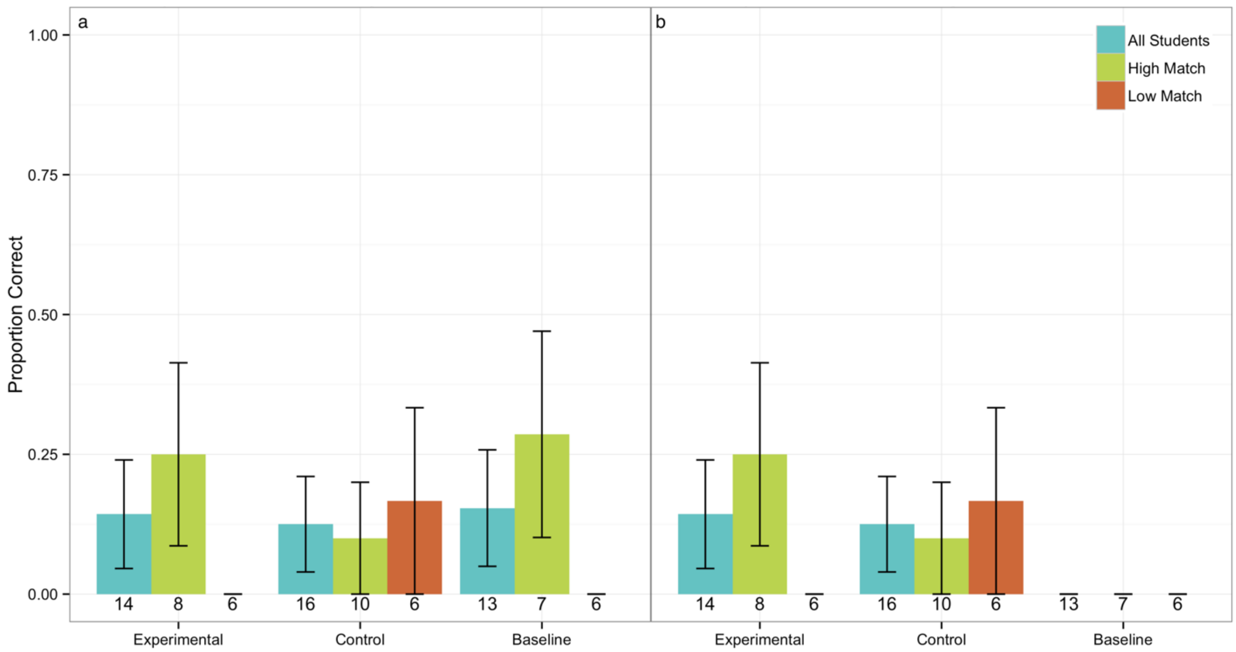


Figure 12. Delayed Post-Test results for the third Far-Transfer problem's items: proportion correct for each condition, broken down by students' overall agreement between their open-response and multiple-choice answers throughout Learning Phase. 'High Match' students (green bars) were at or above their condition's median match agreement and 'Low Match' (red) were below; aggregate totals (blue) are also included for each condition. Error bars represent the standard error of the mean (SEM). Numbers below each bar represent the *n* for that group.

[a] Change 'Puzzle Problem' Question 1 (non-PRM context) [b] Change 'Puzzle Problem' Question 2

Questions. Moreover, no student in any of the conditions used an Elimination procedure (Formal or Informal; Table 9), which likely means that the 'Puzzle' context was too disparate to identify the common internal mathematical structure.

Additionally, in all but 1 of the 6 items, Matchers in the Experimental condition went on to answer correctly more often than Matchers in the Control condition in the analogous far-transfer problems on the DPT (Table 10). The exception was the Symbolic Relational analogue Q2, in which Control slightly outperformed Experimental (26.7% to 23.1%). Again, this recurring pattern suggests that among students who received the appropriate conditional branching, the

Table 9

Method Analysis: Delayed Post-Test Change 'Puzzle Problem' (High-Matchers Only)

	Experimental (n = 8)		Control (n = 10)		Baseline (n = 7)	
Method Used	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Formal Elimination[a]	0%	0%	0%	0%	0%	0%
Informal Elimination[b]	0%	0%	0%	0%	0%	0%
Guess & Check[c]	12.5%	0%	10.0%	0%	0%	28.6%
Other[d]	0%	0%	0%	0%	0%	28.6%
Answered, No Work[e]	12.5%	62.5%	0%	50.0%	0%	28.6%
No Attempt[f]	—	12.5%	—	40.0%	—	14.3%

Note. Solution method were coded by two independent, condition-blind research assistants.

[a] Used the elimination method on formal symbols (e.g., letter or word variables). [b] Used the same mathematical operations but without formal symbols. [c] Guessed numbers and checked mathematically to see if they worked. [d] Unrelated or indiscernible work. [e] Provided an answer without showing any work. [f] Made no attempt at answering (no answer or work).

Experimental condition's manipulations lead to more reliable learning, even on far-transfer problems after a delay.

5.3.3.3 Conceptual Items on the Delayed Post-Test

Composite scores were computed for the 'conceptual' items on the Delayed Post-Test (Figure 13). Since there were nine items, one of which had two questions, students could receive a score anywhere from 0 to 10 for this part of the test. Experimental students had a mean score of 6.0

Table 10

Matching Analysis: Delayed Post-Test Far-Transfer Performance Based on Analogous Problem Matching During Learning Phase

	Experimental				Control			
	Matchers[a]		Switchers		Matchers[a]		Switchers	
Problem	Correct	*n*	Correct	*n*	Correct	*n*	Correct	*n*
Change Role-Reversal Q1	55.6%	9	20.0%	5	15.4%	13	66.7%	3
Change Role-Reversal Q2	55.6%	9	0%	5	38.5%	13	66.7%	3
Relational Symbolic Q1	46.1%	13	0%	1	33.3%	15	0%	1
Relational Symbolic Q2	23.1%	13	0%	1	26.7%	15	0%	1
Change ‘Puzzle Problem’ Q1	22.2%	9	0%	5	7.7%	13	33.3%	3
Change ‘Puzzle Problem’ Q2	22.2%	9	0%	5	7.7%	13	33.3%	3

Note. For each Delayed Post-Test far-transfer problem, students are classified as Matchers or Switchers based on whether their open-response and multiple-choice answers matched or were switched on the analogous problem (Change or Relational) during learning.

[a] ‘Matchers’ include both ‘True Correct’ and ‘True Incorrect’.

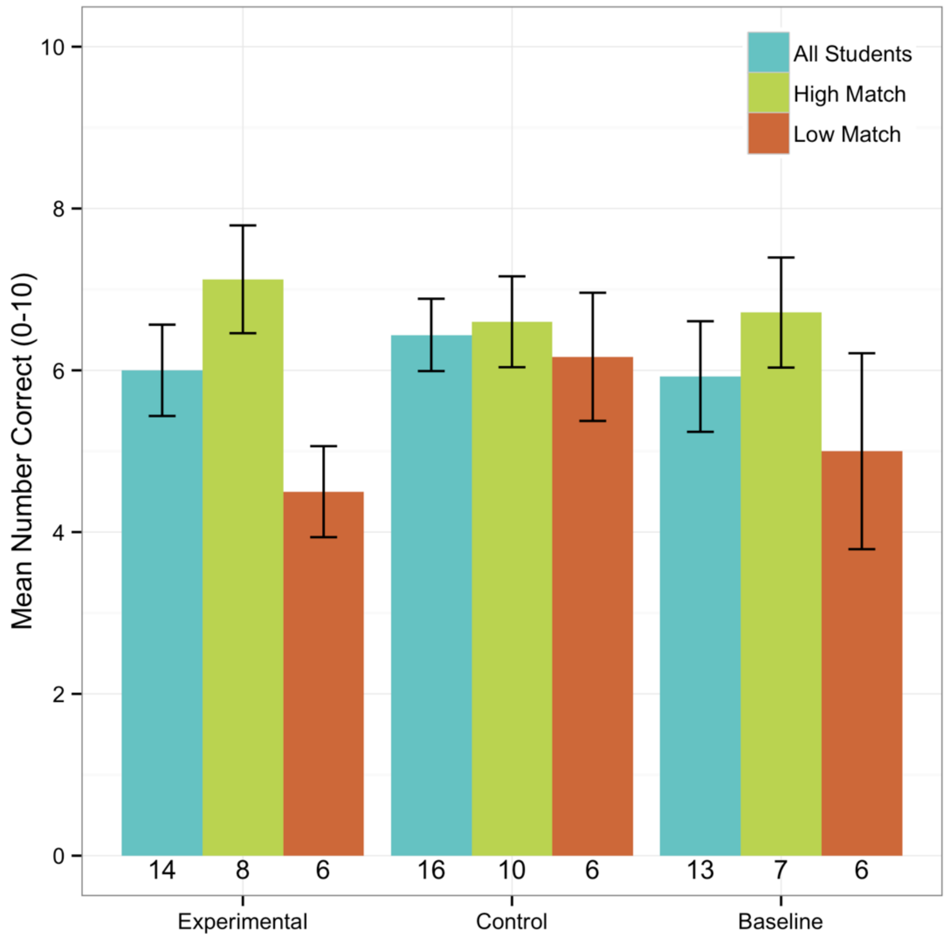


Figure 13. Composite scores (0–10) for the nine Conceptual problems (10 total items) on the Delayed Post-Test: total number correct for each condition, broken down by students' overall agreement between their open-response and multiple-choice answers throughout the Learning Phase. 'High Match' students (green bars) were at or above their condition's median match agreement and 'Low Match' (red) were below; aggregate totals (blue) are also included for each condition. Error bars represent the standard error of the mean (SEM). Numbers below each bar represent the *n* for that group.

(*SD* = 2.11), Control students a mean of 6.4 (*SD* = 1.79), and Baseline a mean of 5.9 (*SD* = 2.47). The scores did not significantly differ across conditions according to a Kruskall-Wallis rank sum test, $\chi 2(2) = 0.29$, $p = .87$. The high performance on these items in the Baseline condition suggests that the populations sampled from had a reasonably high baseline understanding of algebra, which was also indicated in their performance on the Symbolic Relational analogue. However, despite this relatively high baseline, Baseline students were greatly outperformed by Experimental students on the vast majority of non-symbolic post-test items.

5.4 Conclusions

While additional data is needed to draw more substantive conclusions, the results collected thus far are promising. When the teaching materials allowed and encouraged students to infer the purpose of a variable before its formal representation was introduced, they performed best on Delayed Post-Test transfer problems. The most essential indicators of the materials' efficacy are students' performance on the near- and far-transfer Delayed Post-Test items: the analogues to the Change and Relational problems from the Learning Phase, and their isomorphic variants that reversed the contextual roles, were purely symbolic, or had an unrelatable "Puzzle Problem" context. The Change problem type is particularly important due to its role as a focal point of the Learning Phase with many contrast comparisons.

As predicted, there was a significant effect of condition for the Change analogue on the Delayed Post-Test, driven by the Experimental condition's strong performance. This effect was even starker when comparing only the upper half of each condition on a 'match' score that quantifies how often students followed the instructions that allowed for the appropriate conditional branching. Thus, among students who received the closest to the intended experience within each condition, Experimental students outperformed the others to an even greater extent.

While the effect of condition is not significant for the Relational analogue on the Delayed Post-Test, it fits the same qualitative pattern along with the Change analogue, the immediate post-test items, and half of the far-transfer items, with the Experimental condition performing best, followed by Control, and then Baseline, as predicted by the hypotheses.

Despite the fact that students in the Baseline condition performed decently well on the symbolic analogue in the Delayed Post-Test as well as the largely symbolic 'conceptual' items, they performed significantly worse than Experimental students on the contextual items. By contrast, Experimental students, and Control students to a lesser degree, were much more likely to answer contextual items correctly while performing approximately equivalently on the symbolic problems.

Even though students in the Experimental condition only learned from one example problem of each type while the Control condition saw a worked example and completed two practice problems of each type, the Experiment condition consistently scored higher on these transfer problems than the Control. While previous research (e.g., Koedinger & Anderson, 1998) has emphasized the role of schematic comparisons across repeated examples to generate transfer, the current results suggest that encouraging the learner to develop their own intuitive sense of the purpose of a variable may lead to even better transfer performance than the more schematic Control approach.

Additionally, the conditional branching data are encouraging signs for the Experimental condition's efficacy of Hints and Walk-Throughs. While the mismatches between some students' open responses and their subsequent multiple choice answers (Switchers) likely add noise to the current data and make it harder to assess, we can still use the open response data to see how effective the Hints and Walk-Throughs are. Indeed, the modal response pathways for both the

Change and Relational problems suggest that students tend to be able to either solve it on their own or solve it after some assistance, with few students failing to solve a problem after receiving intuitive hints toward its solution.

Because of the unanticipated Switching that led to some students not receiving the intended instructional path, it became necessary to examine differences between students who did and did not match their open response to the multiple choice so that their branching was appropriate. Students in the Experimental condition who consistently matched their multiple choice selection to their open response (High Match) routinely outperformed those who did not on subsequent transfer tasks. Moreover, those students who answered correctly on both open response and the subsequent multiple choice (True Correct) also routinely outperformed those who Switched on the same subsequent transfer tasks, and even those who answered incorrectly on both response formats (True Incorrect) typically outperformed Switchers, suggesting that students benefit from the Experimental condition's branching structure.

5.5 Future Directions

While data collection at Test Site 5 is underway, it is still likely that more data will be necessary, especially for analyzing these finer-grained hypotheses. Moreover, upcoming changes to Zaption's capabilities for text parsing will completely fix the issues encountered so far with 'Switchers', which will make the data a lot less noisy. This feature will be implemented into the current materials as soon as possible.

There are also many possible follow-up experiments to carry out. First, it will be informative to test the three Experimental manipulations one at a time in order to assess the effect sizes of each manipulation and to what extent they may statistically interact with each other. Second, currently popular educational videos, such as those from Khan Academy and ALEKS, could serve as effective 'traditional' controls to compare against these materials, including as a comparison for the current Control condition. Third, these purpose-driven materials could be explicitly compared against more typical schematic comparison materials (e.g., Koedinger & Anderson, 1998).

There are many extensions, improvements, and additions that are planned for the materials as well, such as using the Double Change problem from the written materials, which gets at an even more abstract notion of substitution because it involves substitution of an algebraic expression (e,g., "$20 – 4*h*") that is made as intuitive as possible (as, for example, the change left over after using a $20 bill to pay for 4 hot dogs). More Active Learning and Contrast Comparison prompts can be added as well, such as questions that more specifically target representation.

Additionally, the current materials focus on evoking a sense of purpose for variables while using relatively consistent relatable contexts, but subsequent extensions of the materials could include schematic abstraction across varied examples with the same internal structure as a way to encourage generalization of the intuitive knowledge to more abstract or unfamiliar contexts. Once students have an intuitive sense of the purpose of a variable, having them schematize over concrete examples (as Koedinger & Anderson, 1998 do, for example) may be much more effective than if that purpose had not been focused on originally. Similarly, a more 'concreteness fading' approach (Goldstone & Son, 2005) could have the connection between the initial food item context and the introduced representations slowly fade into more idealized forms, for example, by gradually removing perceptual features from the images until they become abstract ‘blobs’ that nonetheless are meant to symbolically represent some unknown quantity.

Further improvements can be made to the assessment materials as well, such as adding new conceptual items that measure proceptual thinking, in order to directly assess the extent to which students may learn the ‘object’/’product’ conception of a variable from these materials. Metacognitive judgments could also be added, especially for ‘thinking’ active learning prompts such as the questions for the Unsolvable Two-Unknown problem.

The results thus far from this project are encouraging for the future of this approach. There are many future directions to take to further test the hypotheses behind these materials. Moreover, the results from each future test can inform how to improve the materials through iterative cycles of improvement. Altogether, these materials provide a means to collect a rich source of student, to the extent that despite the students’ data lost to internet issues, a large number of inferences can still be made about how students progressed through the materials and how that relates to what they learned. The conditional branching pathways offer an interesting way to observe students’ learning process. The materials are thus a *product* that can be used as both an educational resource for the student population at large, as well as a research tool for examining student learning and thinking. Applying the Purpose-Driven Progressive Formalization framework to further topic areas will create even more effective and rich tools for experimental research and student learning.

APPENDIX A:

Experimental Condition Video Scripts

Narration Legend

- Underlined text: narration that also appears as text on screen
- **Bold text**: narrated word timed to the onset of text on screen
- *Italicized text*: vocal emphasis in narration
- Green text: narration only in the variant of the video on the Correct branch
- Red text: narration only in the variant of the video on the Incorrect branch

Instructions

0 Instructions

- "The following is a series of videos that tell stories and ask you some questions. It is ***not a test*** and you will *not be graded*."
- "You're going to see some short stories told by a young student who buys lunch at or near their school. **Each** story has one or two questions to answer."
- "**Please** use your packet of paper to *show your work* for any math you may do."
- "All the information you need will be shown onscreen as you need it, but **you** may take notes on your paper if you want to."
- "This is a learning experience, so some of these questions may be new to you and seem difficult to answer. **You** are not expected to know how to answer every question. But please try your best."

Story 1: One-Unknown

1-S Story (Gloria's Restaurant)

- "Story 1"
- "I went to Gloria's Restaurant…"
- "And **I** bought two tacos for four dollars."

1-C Context

- **Narrator**: "My friend sat next to me at lunch."
- ***Friend #1***: "**Hey**, those tacos look good! I want to get one tomorrow. How much money did *one* taco cost?"
- ***Narrator***: "**All** I can remember is that I paid *four dollars* total."

1-Q Question

- "**How** much money did *one taco* cost?"

Story 2.1: Unsolvable Two-Unknowns

2.1-S Story (Joe's Pizzeria)

- "Story 2."
- "I went to Joe's Pizzeria…

- "And I **bought** *two* slices of cheese pizza [pause] and ***one*** bottle of water [pause] for ***five*** *dollars*."

2.1-C Context

- **Narrator**: "Two friends sat next to me at lunch."
- ***Friend #1***: "**Hey**, that pizza looks good!"
- ***Friend #2***: "**I** agree. How much money did one slice of cheese pizza cost?"
- ***Narrator***: "**I** already threw away the receipt, but I remember that I paid *five* dollars total."

2.1-Q Question

- "**How** much money did one slice of cheese pizza cost?"

  [See questions on MC Feedback page below]

Story 2.2: Change (Two Unknowns, Two Equations, One Unknown Changes Quantities)

2.2-C0 Pre-Story Context

- "Story 2, continued."
- "I decided to surprise my friends the next day by getting them *each* one slice of cheese pizza from Joe's."

2.2-S Story (Joe's Pizzeria)

- "On **Day** 2, I went *back* to Joe's Pizzeria and **bought** *four* slices of cheese pizza and one bottle of water for nine dollars and fifty cents."

2.2-C1 Context 1

- **Narrator**: "**I** gave each of my friends one slice of cheese pizza and kept the rest for my own lunch."
- ***Friend #1***: "**Hey**, this pizza is great! Thank you!"
- ***Friend #2***: "**I** agree. I want to pay you back. How much money did one slice of cheese pizza cost?"
- ***Narrator***: "**Oh** no! I threw away the receipt again."
- ***Narrator***: "**But** I remember that my order yesterday cost *five* dollars and today's cost *nine dollars and fifty cents*."

2.2-Q1 Question 1

- "Remember what I ordered on each day."
- "**How** much money did *one* slice of cheese pizza cost?"

2.2-[All]-C2 Context 2

2.2-C2 Context 2

2.2-[Q1IncCor]-H-C2 Context 2

2.2-[Q1IncIncX]-W-C2 Context 2

- ***Friend #1***: "**Hey**, don't they sell that same bottle of water here at school?"
- ***Friend #2***: "**Yeah**, it costs seventy-five cents here."
- ***Narrator***: "**Hmm**… I wonder if it would save me money to buy water *at school* next time."

- ***Narrator***: "<u>**But** I don't remember how much money one bottle of water cost at Joe's.</u>"

2.2-Q2 Question 2

- "<u>**How** much money did one *bottle of water* cost at Joe's?</u>"

2.2-[Q1Inc]-H1 Hint 1

- "Since you didn't get it the first time, let's try again with a hint."
- "<u>**Think** about what I ordered each day</u> [pause] <u>**and** how much it cost</u>."
- "<u>**How** did my order *change* between the two days?</u>"

2.2-[Q1Inc]-H2 Hint 2

- "<u>**How** did the *total cost change* between the two days?</u>"

2.2-[Q1Inc]-H-Q1 Hint Question 1

- "So <u>**how** much money did one slice of cheese pizza cost?</u>"

2.2-[Q1IncCor]-H-Q2 Hint Question 2

- "<u>**Think** about my order and total cost for *one* of the days, now that you know the price of one slice of cheese pizza.</u>"
- "<u>**How** much money did *one bottle of water* cost at Joe's?</u>"

2.2-[Q1CorQ2Inc]-H-Q2 Hint Question 2 [Incorr. Q2]

- "Since you didn't get the price of one bottle of water, let's try again with a hint."

- "**Think** about my order and total cost for *one* of the days, now that you know the price of one slice of cheese pizza."
- "**How** much money did *one bottle of water* cost at Joe's?"

2.2-[MultiCor]-W1 Walk-Through 1 [Corr. Tracks]

2.2-[Q1CorQ2Cor]-W1 Walk-Through 1 [Correct]

2.2-[Q1IncCorQ2Cor]-W1 Walk-Through 1 [Correct Hint]

- "Great job! There are a few different ways to find the items' prices. You may have done something similar, but let's walk through the story again *together* so that we can *visualize* *one* way to find the prices."
- "First, let's write what happened in the story in a *new* way to find out if it helps you to better see the math."
- "**We** can represent my order [pause] **and** the total cost [pause] **from** each day using *word equations*:"
- "On **Day** 1, [pause] **the** price of two slices of cheese pizza [pause] + [pause] **the** price of 1 bottle of water [pause] **equaled** [pause] **five** dollars."
- "Then on **Day** 2, [pause] **the** price of *four* slices of cheese pizza [pause] + [pause] **the** price of 1 bottle of water [pause] **equaled** [pause] **nine** dollars and fifty cents."
- "Compare the sentences at the top of the screen **[box]** to the word equations on the bottom **[box]**."
- "**Which** way is *more useful* for finding the price of 1 slice of cheese pizza?"

2.2-[MultiQ2Inc]-W1 Walk-Through 1 [Incorr. Q2 Tracks]

2.2-[Q1CorQ2IncInc]-W1 Walk-Through 1 [Incorr. Q2]

2.2-[Q1IncCorQ2Inc]-W1 Walk-Through 1 [Incorr. Q2 Hint]

- "In order to learn an easier way to answer questions like these, let's walk through the story together one more time."
- "First, let's write what happened in the story in a *new* way to find out if it helps you to better see the math."
- "**We** can represent my order [pause] **and** the total cost [pause] **from** each day using *word equations*:"
- "On **Day** 1, [pause] **the** price of two slices of cheese pizza [pause] + [pause] **the** price of 1 bottle of water [pause] = [pause] **five** dollars."
- "Then on **Day** 2, [pause] **the** price of *four* slices of cheese pizza [pause] + [pause] **the** price of 1 bottle of water [pause] = [pause] **nine** dollars and fifty cents."
- "Compare the sentences at the top of the screen **[box]** to the word equations on the bottom **[box]**."
- "**Which** way is *more useful* for finding the price of 1 slice of cheese pizza?"

2.2-[Q1IncInc]-W1 Walk-Through 1 [Incorr. Q1]

- "In order to learn an easier way to answer questions like these, let's walk through the story together one more time."
- "To figure out how much money 1 slice of cheese pizza costs…"
- "**We** can represent my order [pause] **and** the total cost [pause] **from** each day using *word equations*:"

- “On **Day** 1, [pause] **the** price of two slices of cheese pizza [pause] + [pause] **the** price of 1 bottle of water [pause] = [pause] **five** dollars.”
- “Then on **Day** 2, [pause] **the** price of *four* slices of cheese pizza [pause] + [pause] **the** price of 1 bottle of water [pause] = [pause] **nine** dollars and fifty cents.”
- “Compare the sentences at the top of the screen **[box]** to the word equations on the bottom **[box]**.”
- “**Which** way is *more useful* for finding the price of 1 slice of cheese pizza?”

2.2-[All]-W2 Walk-Through 2 [All Tracks]

2.2-[Q1CorQ2Cor]-W2 Walk-Through 2 [Correct]

2.2-[Q1IncCorQ2Cor]-W2 Walk-Through 2 [Correct Hint]

2.2-[Q1CorQ2IncInc]-W2 Walk-Through 2 [Incorr. Q2]

2.2-[Q1IncCorQ2Inc]-W2 Walk-Through 2 [Incorr. Q2 Hint]

2.2-[Q1IncInc]-W2 Walk-Through 2 [Incorr. Q1]

- “**Now** practice *writing down* these *word equations* [box] on your scratch paper.”

2.2-[All]-W3 Walk-Through 3 [All Tracks]

2.2-[Q1CorQ2Cor]-W3 Walk-Through 3 [Correct]

2.2-[Q1CorQ2IncInc]-W3 Walk-Through 3 [Incorr. Q2]

2.2-[Q1IncInc]-W3 Walk-Through 3 [Incorr. Q1]

- “So, looking at our word equations…”
- “**How** did Day 2 *differ* from Day 1 in the *number* of items ordered?”

2.2-[All]-W4 Walk-Through 4 [All Tracks]

2.2-[Q1CorQ2Cor]-W4 Walk-Through 4 [Correct]

2.2-[Q1CorQ2IncInc]-W4 Walk-Through 4 [Incorr. Q2]

2.2-[Q1IncInc]-W4 Walk-Through 4 [Incorr. Q1]

- "Day 2's order differed from Day 1 because **[box]** 2 *extra* slices of cheese pizza were ordered."
- "Now… **How** did Day 2 differ from Day 1 in the total cost?"

2.2-[Q1IncInc]-W-Q1 Walk-Through Q1 [Incorr. Q1]

- "Day 2's order *also* differed from Day 1's in that the **[box]** *total cost* was four dollars and fifty cents *more*."
- "Now… thinking about the difference in the items ordered and the difference in the total cost, **how** much money did 1 slice of cheese pizza cost?"

2.2-[Multi]-W-(Q1) Walk-Through of Q1 [Cor, Q2Inc, & Q1IncIncCor Tracks]

2.2-[Q1CorQ2Cor]-W-(Q1) Walk-Through of Q1 [Correct]

2.2-[Q1CorQ2Inc]-W-(Q1) Walk-Through of Q1 [Incorr. Q2]

2.2-[Q1IncIncCor]-W-(Q1) Walk-Through of Q1 [IncCorr. Q1]

- "So, the ***difference*** between Day 1 and Day 2's order **[box]** was **two** *extra* slices of cheese pizza."
- "And the *difference* between Day 1 and Day 2's *total cost* **[box]** was an *extra* **four** dollars and fifty cents."
- "So, the price of two slices of cheese pizza [pause] ***equaled*** four dollars and fifty cents."

- "Therefore, as you correctly answered, **the** price of one slice of cheese pizza [pause] **equaled** [pause] **two** dollars and twenty-five cents."

2.2-[Q1IncIncInc]-W-(Q1) Walk-Through of Q1 [Incorr. Q1] [got 2.2-[Q1IncInc]-W-Q1 incorrect]

- "Since you didn't find the price of 1 slice of cheese pizza, let's walk through *one* way of finding it."
- "The ***difference*** between Day 1 and Day 2's *order* **[box]** was **two** *extra* slices of cheese pizza."
- "The *difference* between Day 1 and Day 2's *total cost* **[box]** was an *extra* **four** dollars and fifty cents."
- "The increase in the cost was due to the extra pizza ordered. So, the price of two slices of cheese pizza [pause] ***equaled*** four dollars and fifty cents.
- "Therefore, **the** price of 1 slice of cheese pizza [pause] **equaled** [pause] **two** dollars and twenty-five cents."

2.2-[QXInc]-W-Q2 Walk-Through Q2 [Incorr. QX] [joins Q1CorQ2Inc and Q1IncIncX tracks]

- "Let's focus on the word equation for Day 1. **[fade out Day 2]**"
- "Now that you know the price of 1 slice of cheese pizza [pause] can you figure out **how** much money 1 bottle of water cost at Joe's?"

2.2-[Multi]-W-(Q2) Walk-Through of Q2 [Corr. & QXIncQ2Cor Tracks]

2.2-[Q1CorQ2Cor]-W-(Q2) Walk-Through of Q2 [Correct]

2.2-[QXIncQ2Cor]-W-(Q2) Walk-Through of Q2 [Incorr. QX, Corr. Q2]

- "As we found out before, **the** price of the *extra two* slices of cheese pizza [pause] **equaled** [pause] **four** dollars and fifty cents."
- "Now let's rewrite Day 1's word equation… instead of using the *words* [pop], let's *replace* them with the *number* [pop] they represent."
- "So, ***four** dollars and fifty cents* [move, pause] plus the price of 1 bottle of water equaled five dollars."
- "Therefore, as you correctly answered, **the** price of 1 bottle of water [pause] **equaled** [pause] **fifty** cents."
- "Since the price of one bottle of water at school is seventy-five cents, it's cheaper at Joe's."

2.2-[QXIncQ2Inc]-W-(Q2) Walk-Through of Q2 [Inc. QX, Inc. Q2]

- "Since you didn't find the price of 1 bottle of water, let's walk through *one* way of finding it."
- "Let's focus again on the word equation for Day 1. **[fade out Day 2]**"
- "Now that we *know* the price of the cheese pizza, we can *take* this information **[]** and *use* it in our *word equation* to figure out how much of the *total cost* **[box around cost]** was due to the bottle of water **[box around water]**."
- "As we found out before, **the** price of the *extra two* slices of cheese pizza [pause] **equaled** [pause] **four** dollars and fifty cents."

- "Let's copy Day 1's word equation below."
- "Now we can replace the *words* [box] with the *number* [box] they represent. Notice they are *equal* [pop]."
- "So, ***four** dollars and fifty cents* [move, pause] plus the price of 1 bottle of water equaled five dollars."
- "Out of the total cost of five dollars **[box]**, four dollars and fifty cents **[box]** was for the pizza. This leaves fifty cents for the water."
- "So, **the** price of 1 bottle of water [pause] **equaled** [pause] **fifty** cents."
- "Since the price of one bottle of water at school is seventy-five cents, it's cheaper at Joe's."

2.2-B Big Picture

- "Let's think about the big picture from this story."
- "**Why** could we *not* find the price of one slice of cheese pizza on Day 1, [fade in Day 1] but we *could* find the price of one slice of cheese pizza after *both* Day 1 and Day 2? [fade in Day 2]"

Story 3: Relational

3-C0 Pre-Story Context

- **Narrator**: “Story 3.”
- ***Friend #1***: “I’m going to Patty’s Bakery on the way to school tomorrow to get a cupcake. Would you like me to bring you something?”
- ***Narrator***: “That would be great! Could you please get me a donut? I’ll pay you back!”

3-S Story (Patty’s Bakery)

- “My friend went to Patty’s Bakery…”
- “And she <u>**bought** *one* donut</u> [pause] <u>and ***one*** cupcake</u> [pause] <u>for ***four*** *dollars and fifty cents*</u>.”

3-C1 Context 1

- **Narrator**: “The next day at school, my friend gave [pause] **the** donut to me and kept [pause] **the** cupcake for herself.”
- ***Narrator***: “<u>**Thanks** for the donut! I need to pay you back. How much money did one donut cost?</u>”
- ***Friend #1***: “<u>**Oops**, I can’t remember! The receipt only says that the total cost was four dollars and fifty cents.</u>”
- ***Friend #1***: “<u>**But** I do remember noticing that the price of one cupcake was *twice* the price of one donut.</u>”

3-Q1 Question 1

- "So… **my** friend bought one donut and one cupcake for four dollars and fifty cents."
- "And… **the** price of one cupcake was twice the price of one donut."
- "**How** much money did one donut cost?"

3-[All]-C2 Context 2 (To Copy)

3-C2 Context 2

3-[Q1IncCor]-H-C2 Context 2

3-[Q1IncIncX]-W-C2 Context 2

- ***Narrator***: "**Your** cupcake looks good. I might want to get one next time, if it's not too expensive."
- ***Narrator***: "**How** much money did one cupcake cost?"

3-Q2 Question 2

- "**How** much money did 1 cupcake cost?"

3-[Q1Inc]-H Hint

- "Since you didn't get it the first time, let's try again with a hint."
- "**Let's** make use of the *relationship* between the price of one cupcake [pause] and the price of one donut."
- "**What** would the total cost have been if my friend got one donut for me and two donuts for herself *instead* of the cupcake?"

3-[Q1Inc]-H-Q1 Hint Question 1

- "So how much money did one donut cost?"

3-[Q1IncCor]-H-Q2 Hint Question 2

- "**Think** about my friend's *order* and the *total cost*, now that you know the price of one donut."
- "**How** much money did one cupcake cost?"

3-[Q1CorQ2Inc]-H-Q2 Hint Question 2 [Incorr. Q2]

- "Since you didn't get the price of one cupcake, let's try again with a hint."
- "**Think** about my friend's *order* and the *total cost*, now that you know the price of one donut."
- "**How** much money did one cupcake cost?"

3-[Q1CorQ2Cor]-W1 Walk-Through 1 [Correct] [got 3-Q1&2 or 3-[Q1Inc]-H-Q1&2 correct]

- "Great job! There are a few different ways to find the items' prices. You may have done something similar, but let's walk through the story again *together* so that we can *visualize* *one* way to find the prices."
- "Let's use **word** equations again [pause] **to** represent the total cost [pause] **and** the *relationship* between the items' prices."
- "For the total cost, we can write that [pause] the price of 1 donut [pause] plus [pause] the price of 1 cupcake [pause] equaled [pause] four dollars and fifty cents."
- "For the relationship, we can write that [pause] the price of 1 cupcake [pause] equaled

[pause] <u>two</u> times [pause] <u>the</u> price of 1 donut [pause] and because 'two times the price of 1 donut' is equal to [pause] <u>the</u> price of 2 donuts, we can instead write that the price of 1 cupcake is equal to [pause] <u>the</u> price of 2 donuts."

- "<u>Are these word equations more *useful* for finding the price of 1 donut? If so, how?</u>"

3-[Q1CorQ2Inc]-W1 Walk-Through 1 [Incorr. Q2] [got 3-[Q2Inc]-H-Q2 or 3-[Q1Inc]-H-Q2 incorrect]

- "In order to learn an easier way to answer questions like these, let's walk through the story together one more time."
- "Let's use <u>**word** equations</u> again [pause] <u>**to**</u> represent the total cost [pause] <u>**and**</u> the *relationship* between the items' prices."
- "For the <u>total cost</u>, we can write that [pause] <u>the</u> price of 1 donut [pause] <u>plus</u> [pause] <u>the</u> price of 1 cupcake [pause] <u>equaled</u> [pause] <u>four</u> dollars and fifty cents."
- "For the <u>relationship</u>, we can write that [pause] <u>the</u> price of 1 cupcake [pause] <u>equaled</u> [pause] <u>two</u> times [pause] <u>the</u> price of 1 donut… [pause] and because 'two times the price of 1 donut' is equal to [pause] <u>the</u> price of 2 donuts… [pause] the price of 1 cupcake is equal to [pause] <u>the</u> price of 2 donuts."
- "<u>Are these word equations more *useful* for finding the price of 1 donut? If so, how?</u>"

3-[Q1IncInc]-W1 Walk-Through 1 [Incorr. Q1]

- "In order to learn an easier way to answer questions like these, let's walk through the story together one more time."
- "To figure out how much money 1 donut cost…"

- “Let’s use **word** equations again [pause] **to** represent the total cost [pause] **and** the *relationship* between the items’ prices.”
- “For the total cost, we can write that [pause] the price of 1 donut [pause] plus [pause] the price of 1 cupcake [pause] equaled [pause] four dollars and fifty cents.”
- “For the relationship, we can write that [pause] the price of 1 cupcake [pause] equaled [pause] two times [pause] the price of 1 donut.”
- “Now, because ‘two times the price of 1 donut’ is equal to [pause] the price of 2 donuts, we can *instead* write that [pause] the price of 1 cupcake [pause] equals [pause] the price of 2 donuts.”
- “Are these word equations more *useful* for finding the price of 1 donut? If so, how?”

3-[All]-W2 Walk-Through 2 [All Tracks]

3-[Q1CorQ2Cor]-W2 Walk-Through 2 [Correct]

3-[Q1CorQ2Inc]-W2 Walk-Through 2 [Incorr. Q2]

3-[Q1IncInc]-W2 Walk-Through 2 [Incorr. Q1]

- “Now let’s focus on the relationship word equation.”
- “How many donuts could we buy for the same price as one cupcake?”

3-[All]-W3 Walk-Through 3 [All Tracks]

3-[Q1CorQ2Cor]-W3 Walk-Through 3 [Correct]

3-[Q1CorQ2Inc]-W3 Walk-Through 3 [Incorr. Q2]

3-[Q1IncInc]-W3 Walk-Through 3 [Incorr. Q1]

- “We could buy *two donuts* [box] for the same price as one cupcake [box], because their prices are equal [pop]."

- "Now looking across *both* word equations…"
- "How many donuts could we buy for four dollars and fifty cents?"

3-[Q1IncInc]-W-Q1 Walk-Through Q1 [Incorr. Q1]

- "We could buy *three donuts* for four dollars and fifty cents, because a donut and a cupcake cost four dollars and fifty cents [box Total Cost], and the cupcake costs the same as two more donuts [box Relationship]."
- "So… **How** much money did *one* donut cost?"

3-[Multi]-W-(Q1) Walk-Through of Q1 [Cor & Q2IncTracks]

3-[Q1CorQ2Cor]-W-(Q1) Walk-Through of Q1 [Correct]

3-[Q1CorQ2Inc]-W-(Q1) Walk-Through of Q1 [Incorr. Q2]

- "We could buy *three donuts* for four dollars and fifty cents, because a donut and a cupcake cost four dollars and fifty cents [box Total Cost], and the cupcake costs the same as two more donuts [box Relationship]."
- "Let's rewrite the total cost word equation… and *replace* 'the price of one cupcake' [box] with 'the price of two donuts' [box]."
- "The price of one donut plus the price of *two donuts* [move] equaled four dollars and fifty cents."
- "So, **the** price of *three* donuts [pause] **equaled** [pause] **four** dollars and fifty cents."
- "Therefore, as you correctly answered, [pause] **the** price of *one* donut [pause] **equaled** [pause] **one** dollar and fifty cents."

3-[Q1IncIncCor]-W-(Q1) Walk-Through of Q1 [IncCorr. Q1]

- "Let's rewrite the total cost word equation… and *replace* 'the price of one cupcake' [box] with 'the price of two donuts' [box]."
- "The price of one donut plus the price of *two donuts* [move] equaled four dollars and fifty cents."
- "So, **the** price of *three* donuts [pause] **equaled** [pause] **four** dollars and fifty cents."
- "Therefore, as you correctly answered, [pause] **the** price of *one* donut [pause] **equaled** [pause] **one** dollar and fifty cents."

3-[Q1IncIncInc]-W-(Q1) Walk-Through of Q1 [Incorr. Q1] [got 3-[Q1IncInc]-W-Q1 incorrect]

- "Since you didn't find the price of one donut, let's walk through *one* way of finding it."
- "We know that the price of one cupcake [box] *equaled* [pop] the price of two donuts [box]."
- "Let's copy the total cost word equation below."
- "Now, we can *replace* 'the price of one cupcake' [box] with 'the price of two donuts' [box], because they are *equal* [pop]."
- "So, the price of one donut plus the price of *two donuts* [move] equaled four dollars and fifty cents."
- "And since the price of one donut plus the price of two more donuts is equal to **the** price of *three* donuts, the price of three donuts [move] [pause] **equaled** [pause] **four** dollars and fifty cents."
- "In other words, if my friend had bought *three* donuts, the total cost would *still* be four

dollars and fifty cents."

- "Therefore, [pause] **the** price of *one* donut [pause] **equaled** [pause] **one** dollar and fifty cents."

3-[QXInc]-W-Q2 Walk-Through Q2 [Incorr. QX] [joins Q2Inc, Q1IncInc, and Q1IncIncInc tracks]

- "Focus on *either* one of the word equations."
- "Now that you know the price of one donut, [pause] can you figure out **how** much money one cupcake cost?"

3-[Multi]-W-(Q2) Walk-Through of Q2 [Corr. & QXIncQ2Cor Tracks]

3-[Q1CorQ2Cor]-W-(Q2) Walk-Through of Q2 [Correct]

3-[QXIncQ2Cor]-W-(Q2) Walk-Through of Q2 [Incorr. QX, Corr. Q2]

- "As we found out before, **the** price of one donut [pause] **equaled** [pause] **one** dollar and fifty cents."
- "Now let's rewrite the total cost word equation… instead of using the *words* [pop], let's *replace* them with the *number* [pop] they represent."
- "So, ***one** dollar and fifty cents* [pause] plus the price of one cupcake equaled four dollars and fifty cents."
- "Therefore, as you correctly answered, **the** price of one cupcake [pause] **equaled** [pause] **three** dollars."

3-[QXIncQ2Inc]-W-(Q2) Walk-Through of Q2 [Inc. QX, Inc. Q2]

- "Since you didn't find the price of one cupcake, let's walk through *one* way of finding it."
- "Let's focus again on the word equation for the total cost. **[fade out Relationship]**"
- "Now that we *know* the price of the donut, we can *take* this information and *use* it in our *word equation*."
- "As we found out before, **the** price of one donut [pause] **equaled** [pause] **one** dollar and fifty cents."
- "Let's copy the total cost word equation below."
- "Now, we can replace the *words* [box] with the *number* [box] they represent."
- "So, ***one** dollar and fifty cents* [pause] plus the price of one cupcake equaled four dollars and fifty cents."
- "Out of the total cost of four dollars and fifty cents **[box]**, one dollar and fifty cents **[box]** was for the donut. This leaves *three dollars* for the cupcake."
- "So, **the** price of one cupcake [pause] **equaled** [pause] **three** dollars."

Lessons

L1-Q Lesson 1 Question

- "Now that you've worked through the stories, let's go through a *lesson* together."
- "In the walk-throughs for Stories 2 and 3, **we** wrote *word equations* using *terms* such as: [pause] '**the** price of 1 slice of cheese pizza'."

- "When we write 'the price of 1 slice of cheese pizza' we are using these *words* [] to *represent* [] a *number* []: **the** amount of money that 1 slice of cheese pizza cost."
- "Using *words* [pop] to represent a *number* [pop] is useful when we don't know what that number *is* yet [?]."
- "But you may have noticed that we had to write "the price of 1 slice of cheese pizza" [pop] many times. **Writing** *all* of these words *every* time makes the equations *very long*, and we might get tired of doing so much writing."
- "**How** might we make these word equations *shorter*?"

L1-L Lesson 1 Lesson

- "Instead of writing the whole phrase '**the** price of 1 slice of cheese pizza', we can use the *letter* ***p*** (for "pizza") to *abbreviate* [], or use a shorter way of writing [*aside*], all of those *words*."
- "Just as 'the price of 1 slice of cheese pizza' [pop] *represented* [] a *number* [] that we didn't know yet [?], we will use the letter *p* [pop] to *represent* [] that *same number* [pop]."
- "For this story, whenever you see the letter *p* [pop], you can know that it represents *the number* [pop] that is *also* represented by the words 'the price of 1 slice of cheese pizza' [pop]."

L2-Q1 Lesson 2 Question 1

- "Now, let's think about the word equation from Day 1 in Story 2."
- "**If** *p* represents the price of 1 slice of cheese pizza… [] **how** could you represent the price

of 2 slices of cheese pizza in the word equation above?"

L2-Q2 Lesson 2 Question 2

- "**2*p*** or **2** times *p* can be used to represent the price of 2 slices of cheese pizza because they stand for *two times* the price of 1 slice of cheese pizza, which is *p*."
- "**How** could you *represent* the price of 1 bottle of water in the word equation above?"

L2-Q3 Lesson 2 Question 3

- "*One* way to represent the price of one bottle of water is with the letter ***w***."
- "We could also use the letter *b* or any other. The letter that we choose is not important as long as we remember what it represents."
- "Now… **try** rewriting the word equation for Day 1, replacing the word phrases with *letters.*"

L2-L Lesson 2 Lesson

- "Instead of writing the word equation for Day 1 []…"
- "We can replace 'the price of 2 slices of cheese pizza' [pop] with [] 2*p*, and 'the price of 1 bottle of water' [pop] with [] *w*."
- "So, **2*p*** [] **plus** [] ***w*** [] **equaled** [] **five** dollars."
- "We can do the same thing for the word equation for Day 2 []…"
- "So, **4*p*** [] **plus** [] ***w*** [] **equaled** [] **nine** dollars and fifty cents."

L3-Q Lesson 3 Question

- "Thinking about **Story** 2, there are *three* different ways we have *represented* the information in the story."
- "First, we used ***sentences***, which describe what *happened* in the story."
- "Second, we used ***word** equations*, which show the *math* in the story."
- "Third, we just learned ***letter** equations*, which *shorten* the word equations."
- "In your opinion… Which of these three *representations* is the most *useful* for finding the items' prices?"

L3-L Lesson 3 Lesson

- "Many people find ***letter** equations* to be the most useful way to represent information because they clearly show the *math* [pop symbols; words fade in] and are *shorter* than word equations, [words move in] while still *representing* the same numbers."
- "Let's use the Day 1 letter equation as an example [fade out Day 2]."
- "As we learned before, '2*p*' means the *same* thing [copy up and out] as '2 *times* [x] *p*'."
- "And the letter *p* [copy up] comes from '***pizza***', but it abbreviates the *whole phrase*, '**the** price of one slice of cheese pizza'."
- "Both the letter *p* [pop] and the *phrase* [pop] represent the same number, even if we don't know what that number is yet [?]."
- "So the phrase 'the price of one slice of cheese pizza' [pop] is *contained* in the letter 'p' [move in], and the *multiplication* [pop x] of 2 and *p* is *contained* in '2*p*' [move in]."
- "So for this story, whenever you see '2*p*' [pop], you can read it as 'two times the price of one slice of cheese pizza' [move out], which is *equal to* 'the price of two slices of cheese

pizza' [move down]."

- "And *all* of that information is *contained* in '2*p*' [move in]."
- "Similarly, the letter *w* abbreviates the *whole phrase*, '**the** price of one bottle of water' [move out] and represents that number [move in]."
- "[fade in Day 2] And as with the *words*, the *letters p* and *w* in Day 2's equation [pop Day 2's letters] represent the same numbers as in Day 1's [pop Day 1 letters]."

L4-Q Lesson 4 Question

- "Remember that in **Story** 2… [pause] **The** *change* in *total cost* from Day 1 to Day 2… [pause] **was** *caused* by… [pause] **the** *change* in the *order* from Day 1 to Day 2."
- "Let's look again at our [pause] *letter equations* for Day 1 and Day 2."
- "We want to find the [pause] ***change*** between Day 1 and Day 2."
- "What *mathematical operation* would allow you to find the *change* between the numbers?"

L4-[Sub]-L Lesson 4 Lesson [Subtraction]

- "Since we want to know what *changes* from Day 1 to Day 2… we can take *Day 2* [copy move] and *subtract* [-] Day 1 [copy move]."
- "Now, let's look at each ***change*** individually [fade out symbols]."
- "The change in *p* [box] between Day 1 and Day 2 is found by 4*p* minus [-] 2*p*, which equals ***2p***."
- "The change in *w* [box] [-] equals ***zero***."
- "And the change in the *total cost* [box] [-] equals ***four** dollars and fifty cents*."
- "So, 2*p plus* [+] zero equals four dollars and fifty cents."

- "We have taken the Day 2 equation and *subtracted* the Day 1 equation, to find a *new* equation that does *not* have *w* in it. **By** *eliminating w*, we can find *p* directly."
- "So, ***2p*** [] ***equals*** [] ***four*** *dollars and fifty cents*."
- "And [] ***p*** [] ***equals*** [] ***two*** *dollars and twenty-five cents*."
- "Which is a shorter way of writing 'the price of one slice of cheese pizza' [move out] ***equals*** [] ***two*** *dollars and twenty-five cents*."

L4-[Add]-L Lesson 4 Lesson [Addition]

- "Addition *would* allow you to find the change between Day 1 and Day 2, but let's visualize how *subtraction could also* allow you to find the change."
- "Since we want to know what *changes* from Day 1 to Day 2… we can take *Day 2* [copy move] and *subtract* [-] Day 1 [copy move]."
- "Now, let's look at each ***change*** individually [fade out symbols]."
- "The change in *p* [box] between Day 1 and Day 2 is found by 4*p* minus [-] 2*p*, which equals ***2p***."
- "The change in *w* [box] [-] equals ***zero***."
- "And the change in the *total cost* [box] [-] equals ***four*** *dollars and fifty cents*."
- "Notice that you could *also* find these numbers by thinking about *addition* [move numbers] since Day 1's numbers *plus* [+] the *changes* equal Day 2's numbers."
- "So, 2*p* *plus* [+] zero equals four dollars and fifty cents."
- "We have taken the Day 2 equation and *subtracted* the Day 1 equation, to find a *new* equation that does *not* have *w* in it. **By** *eliminating w*, we can find *p* directly."
- "So, ***2p*** [] ***equals*** [] ***four*** *dollars and fifty cents*."

- “And [] ***p*** [] ***equals*** [] ***two** dollars and twenty-five cents*.”
- “Which is a shorter way of writing ‘the price of one slice of cheese pizza’ [move out] ***equals*** [] ***two** dollars and twenty-five cents*.”

L4-[MD]-L Lesson 4 Lesson [Mult-Div]

- “*Subtraction* would allow us to find the *change* between the numbers.”
- “Since we want to know what *changes* from Day 1 to Day 2… we can take *Day 2* [copy move] and *subtract* [-] Day 1 [copy move].”
- “Now, let’s look at each ***change*** individually [fade out symbols].”
- “The change in *p* [box] between Day 1 and Day 2 is found by 4*p* minus [-] 2*p*, which equals ***2p***. In other words, there is two *more p* on Day 2 than Day 1.”
- “The change in *w* [box] [-] equals ***zero***, because any number subtracted from itself is zero, even if we don’t know what it is. In other words, there is *no change* in *w*.”
- “And the change in the *total cost* [box] is [-] equals ***four** dollars and fifty cents*. In other words, Day 2 cost four dollars and fifty cents *more* than Day 1.”
- “So, 2*p plus* [+] zero **equals** four dollars and fifty cents.”
- “We have taken the Day 2 equation [pop] and *subtracted* the Day 1 equation [pop], to find a *new* equation [pop] that does *not* have *w* in it. **By** *eliminating w*, we can find *p* directly.”
- “Since 2*p* plus zero is just 2*p*, [] ***2p*** [] ***equals*** [] ***four** dollars and fifty cents*.”
- “So [] ***p*** [] ***equals*** [] ***two** dollars and twenty-five cents*.”
- “Which is a shorter way of writing ‘the price of one slice of cheese pizza’ [move out] ***equals*** [] ***two** dollars and twenty-five cents*.”

Final Stories (Immediate Post-Tests: FS1=Change Analogue, FS2=Relational Analogue)

FS1-Q1 Final Story 1 Question 1

- "For our final story… use what you have learned about letter equations."
- "I went to Bill's Burgers and [] **bought** three hamburgers and two milkshakes for twenty dollars and forty-five cents."
- "Later I went back to Bill's Burgers and [] **bought** seven hamburgers and two milkshakes for thirty-nine dollars and forty-five cents."
- "**How** much money did one hamburger cost?"

FS1-[Q1Inc]-H-Q1 Final Story 1 Hint Question 1

- "**Think** about how we used letter equations to find the price of 1 slice of cheese pizza for Story 2."
- "**How** much money did one hamburger cost?"

FS1-[Q1IncInc]-W-Q1 Final Story 1 Walk-Through Question 1 [Q1 Incorr.]

- "Try using this **Day** 2 letter equation and *subtracting* this **Day** 1 letter equation."
- "Now… **how** much money did one hamburger cost?"

FS1-Q2 Final Story 1 Question 2

- "**How** much money did one milkshake cost?"

FS2-Q1 Final Story 2 Question 1

- “I went to Annie’s Cafe and I [] **bought** one coffee and one slice of pie for five dollars and seventy cents.”
- “And **the** price of one slice of pie was *five times* the price of one coffee.”
- “**How** much money did one coffee cost?”

FS2-Q2 Final Story 2 Question 2

- “**How** much money did one slice of pie cost?”

Multiple Choice Feedback

Story *incorrect*: “Sorry, that is incorrect. Let’s try again with a hint…”

Hint *incorrect*: “Sorry, that is incorrect. Let’s walk through the story together…”

Walk-Thru *inc.*: “Sorry, that is incorrect. Let’s walk through one way of finding it…”

FS Hint *incorrect*: “Sorry, that is incorrect. Let’s try again with another hint…”

FS Walk-Thru *inc.*: “Sorry, that is incorrect.”

Story 2.1 (Unsolvable Two-Unknowns) Questions (Contrast Comparisons)

- “How much money did 1 slice of cheese pizza cost?”

  [Open-ended response, no feedback, continue to next Q.]
- “Do we have enough information to know for sure how much 1 slice of cheese pizza cost?”

  Yes: Incorrect. [No text, continue to next Q.]

No: Correct. "There are multiple possible prices for 1 slice of cheese pizza, given the information we know." [Skip remaining Qs.]

- "Is it possible for 1 slice of cheese pizza to cost $2.75?"

  Yes: [No feedback, continue to next Q.]

  No: [No feedback, continue to next Q.]

- "Is it possible for 1 slice of cheese pizza to cost $1.75?"

  Yes: [No feedback, continue to next Q.]

  No: [No feedback, continue to next Q.]

- "Is it possible for 1 slice of cheese pizza to cost $2.10?"

  Yes: Correct. "$2.10 is a possible price. Although $2.75 is *impossible*, there are *multiple* possible prices (including $1.75) given the information we know. So we do *not* have enough information to know for sure how much 1 slice of cheese pizza cost." [Continue to next video.]

  No: Incorrect. "Actually, $2.10 is a possible price (2 slices would cost $4.20 total). Although $2.75 is *impossible*, there are *multiple* possible prices (including $1.75) given the information we know. So we do *not* have enough information to know for sure how much 1 slice of cheese pizza cost." [Continue to next video.]

Story 2.2 (Change) "Big Picture" Questions (Contrast Comparisons)

- "Why could we not find the price of 1 slice of cheese pizza on Day 1, but we could find the price of 1 slice of cheese pizza after both Day 1 and Day 2?"

  [Open-ended response, no feedback, continue to next Q.]

- "Now that we know both days' orders, is it possible for 1 slice of cheese pizza to cost $1.75?"

Yes: Incorrect. “It is not possible. Once we know both days' orders, we have enough information to know for sure that the price of 1 slice of pizza is \$2.25, and it cannot be \$1.75.” [Continue to next video.]

No: Correct. “Once we know both days' orders, we have enough information to know for sure that the price of 1 slice of pizza is \$2.25.” [Continue to next video.]

APPENDIX B:

Control Condition Video Scripts

Narration Legend

- Underlined text: narration that also appears as text on screen
- **Bold text**: narrated word timed to the onset of text on screen
- *Italicized text*: vocal emphasis in narration

Instructions

0 Instructions

- "The following is a series of videos that tell stories and ask you some questions. It is ***not*** *a test* and you will *not be graded*."
- "You're going to see some short stories told by a young student who buys lunch at or near their school. **Each** story has one or two questions to answer."
- "**Please** use your packet of paper to *show your work* for any math you may do."
- "All the information you need will be shown onscreen as you need it, but **you** may take notes on your paper if you want to."
- "This is a learning experience, so some of these questions may be new to you and seem difficult to answer. **You** are not expected to know how to answer every question. But please try your best."

Control Story 1: One-Unknown

C1-L0 Control 1 Lesson 0

4. “In this lesson, you will learn about ***variables***.”
5. “A ***variable*** is **a** *symbol* that represents a *number*.
6. “The symbol is **usually** a *letter*, like *x* or *y*.”
7. “Variables are *helpful* because they **can** represent a number we don’t *know* yet.”
8. “Let’s practice using a variable to solve a math problem.”

C1-S0 Control 1 Story 0

- “I went to Maria’s Restaurant…”
- “And **I** bought *three* tacos for *nine* dollars.”

C1-L1.1 Control 1 Lesson 1.1

- “Our **goal** is to **figure** out how much money one taco cost.”
- “For this problem, we can use the letter ***t*** as a ***variable*** that ***represents*** [] **the** price of one taco.”
- “In this case, the letter ’t’ comes from *taco*, but we could use any letter we want. The letter that we choose is not important as long as we remember what it represents.”
- “We know from the story that **the** price of *three* tacos is nine dollars.”
- “Since the price of three tacos is the *same* as **three** times the price of *one* taco [copy move down], we know that three times the price of one taco **is** also nine dollars [copy move down].”

- “Now we can take this information and write a math equation using a variable.”
- “We can represent ‘three times the price of one taco’ [pop] as **three** *times t*, because *t* represents ‘the price of one taco’ [copy move in].”
- “And since ‘three times the price of one taco’ is *nine dollars* [pop], three times *t* ***equals*** [] **nine** dollars.”

C1-L1.2 Control 1 Lesson 1.2

- “Three times *t* can also be written as just **three** *t*. Writing multiplication with a variable this way is shorter, and it also avoids using the multiplication sign [pop], which looks a lot like the letter *x*, a commonly used variable.”
- “So *t* represents a number, and *three times that number* **equals** **nine** dollars… therefore, ***t*** [] must **equal** [] ***three*** dollars.”
- “Notice that we just wrote *t* by itself, instead of writing **one** *t*. Just as we can avoid using the multiplication sign, we can avoid writing the numeral ‘one’ next to a variable because it’s shorter, and because any number multiplied by *one* is that same number.”
- “Now let’s double-check our work. First, let’s copy our original equation [copy move right].”
- “Since *t* and three dollars are *equal* [pop], we can *substitute* **three** dollars for *t* in the equation.”
- “So… three times *three dollars* [move down] equals nine dollars. This is true, so we have correctly found the number that *t* represents.”

C1-S1 Control 1 Story 1

- "I went to Gloria's Restaurant, where the prices are *different*…"
- "And **I** bought *two* tacos for *four* dollars."

C1-Q1 Control 1 Question 1

- "**How** much money did *one taco* cost at Gloria's?"

C1-[Q1Inc]-E1 Control 1 Explanation 1 [Q1 Incorrect]

- "Let's use the letter *t* as a variable again."
- "**The** price of two tacos is four dollars, so **two** times the price of *one* taco [move down] is also four dollars [move down]."
- "We can write **two** *t* equals four dollars to represent this information."
- "So, ***t*** must equal two dollars."

C1-S2 Control 1 Story 2

- "I went to Al's Restaurant…"
- "And **I** bought *four* tacos for *eleven* dollars."

C1-Q2 Control 1 Question 2

- "**How** much money did *one taco* cost at Al's?"

C1-[Q2Inc]-E2 Control 1 Explanation 2 [Q2 Incorrect]

- "Let's use the letter *t* again as a variable."

- “**The** price of four tacos is eleven dollars, so **four** times the price of *one* taco [move down] is also eleven dollars [move down].”
- “We can write **four** *t* equals eleven dollars to represent this information.”
- “So, ***t*** must equal *two* dollars and *seventy-five* cents.”

Control Story 2: Change

C2-L0 Control 2 Lesson 0

- “In this lesson, you will learn how to use ***multiple*** *variables* and ***multiple*** *equations*.”
- “In order to tell the variables apart, we **must** use a *different* letter for each variable.”
- “This allows us to **represent** *multiple* numbers we don’t *know* yet.”
- “You will also learn how to **use** the *elimination method* to find a variable’s value.”
- “Let’s practice using two variables to solve a math problem.”

C2-S0 Control 2 Story 0

- “On Day 1, I went to Lori’s Pizzeria…”
- “And I bought *one* slice of cheese pizza [pause] and *one* bottle of water [pause] for *four* dollars.”
- “On Day 2, I went *back* to Lori’s Pizzeria…”
- “And I bought *three* slices of cheese pizza [pause] and *one* bottle of water [pause] for *ten* dollars.”

C2-L1.1 Control 2 Lesson 1.1

- "Our ***goals*** are to ***figure** out* how much money one slice of cheese pizza cost, and **how** much money one bottle of water cost."
- "For this problem, we can use the letter ***p*** as a *variable* that ***represents*** [] **the** price of one slice of cheese pizza, and the letter ***w*** as a *variable* that ***represents*** [] **the** price of one bottle of water."
- "We know from Day 1 [fade in Day 1 equation] that the total price of one slice of cheese pizza and one bottle of water is four dollars."
- "Now we can take this information and write a math equation using our two variables [pop p and w]."
- "The price of one slice of cheese pizza is ***p*** and the price of one bottle of water is ***w***. Since the two prices are *added together* to get the total cost, *p* ***plus*** *w* ***equals*** [] **four** dollars."
- "So, *p* and *w* represent two different numbers, and the *sum* of those two numbers is four dollars."
- "We know from Day 2 [fade in Day 2 equation] that the total price of *three* slices of cheese pizza and one bottle of water is *ten* dollars."
- "The price of three slices of cheese pizza is ***three** p*, and the price of one bottle of water is ***w***. Since the prices are *added together* to get the total cost, three *p* ***plus*** *w* ***equals*** [] **ten** dollars."

C2-L1.2 Control 2 Lesson 1.2

- “Since we have *two* variables, we can *eliminate one* variable to find the *other* variable’s value.”
- “Comparing the two equations, notice that p [pop] *increases* from Day 1 to Day 2, w [pop] stays the same, and the total cost [pop] *also increases*.”
- “Let’s find the *difference* between the two equations by *subtracting* the first equation from the second. This will allow us to *eliminate* w.”
- “Let’s copy our Day 2 equation [copy move] and the Day 1 equation [copy move] below it, so we can ***subtract***.”
- “The *difference* in p [box] between Day 1 and Day 2 is found by $3p$ minus [-] p, which equals ***2p***.”
- “The difference in w [box] [-] equals ***zero***, because any number subtracted from itself is *zero*, even if we don’t know what it is.”
- “And the difference in the *total cost* [box] [-] equals ***six*** *dollars*.”
- “So, $2p$ *plus* [+] zero **equals** six dollars.”
- “We have taken the Day 2 equation [pop] and *subtracted* the Day 1 equation [pop], to find a *new* equation [pop] that does *not* have w in it. **By** *eliminating w*, we can find p directly.”
- “Since $2p$ plus zero is just $2p$, [] ***2p*** [] ***equals*** [] ***six*** *dollars*.”
- “So, [] ***p*** [] ***equals*** [] ***three*** *dollars*.”
- “So we found that ‘the price of one slice of cheese pizza’ [move out] ***is*** [] ***three*** *dollars*.”

C2-L1.3 Control 2 Lesson 1.3

- "Now that we have found the price of one slice of cheese pizza [cross out], we can *take* this information and *use* it in our *equation* to figure out the price of one bottle of water."
- "Remember that <u>***p*** equals three dollars</u>."
- "Now let's focus on just the Day 1 equation [fade out Day 2]."
- "Instead of using the letter *p* [pop], let's replace it with the number [pop] it represents. This is called *substitution*."
- "So, *three dollars* [copy down] plus *w* equals *four* dollars."
- "Therefore, <u>***w***</u> must <u>**equal**</u> [] <u>***one* dollar**</u>."
- "So we found that 'the price of one bottle of water' [move out] <u>***is***</u> [] <u>***one** dollar*</u>."

C2-S1 Control 2 Story 1

- "On Day 1, I went to Joe's Pizzeria, where the prices are *different*…"
- "And I bought *two* slices of cheese pizza [pause] and *one* bottle of water [pause] for *five* dollars."
- "On Day 2, I went *back* to Joe's Pizzeria…"
- "And I bought *four* slices of cheese pizza [pause] and *one* bottle of water [pause] for *nine dollars and fifty cents*."

C2-Q1.1 Control 2 Question 1.1

- "<u>**How** much money did *one slice of cheese pizza* cost at Joe's?</u>"

C2-[Q1.1Inc]-E1.1 Control 2 Explanation 1.1 [Q1.1 Incorrect]

- "Let's use the letters *p* and *w* again as our variables."
- "Looking at the information from **Day** 1, we can write that ***two*** *p* plus *w* equals five dollars."
- "And looking at the information from **Day** 2, we can write that ***four*** *p* plus *w* equals nine dollars and fifty cents."
- "Now we can eliminate *w* by taking our Day 2 equation [copy move] and *subtracting* [-] our Day 1 equation [copy move]."
- "The *difference* in *p* [box] between Day 1 and Day 2 is found by 4*p* minus [-] 2*p*, which equals ***2p***."
- "The difference in *w* [box] [-] equals ***zero***."
- "And the difference in the *total cost* [box] [-] equals ***four*** *dollars and fifty cents*."
- "So, 2*p* *plus* [+] zero **equals** four dollars and fifty cents."
- "So, ***two*** *p* equals four dollars and fifty cents."
- "Therefore, ***p*** equals two dollars and twenty-five cents."

C2-Q1.2 Control 2 Question 1.2

- "**How** much money did *one bottle of water* cost at Joe's?"

C2-[Q1.1Inc]-E1.2 Control 2 Explanation 1.2 [Q1.2 Incorrect]

- "Let's use the letters *p* and *w* again as our variables."
- "Looking at the information from **Day** 1, we can write that ***two*** *p* plus *w* equals five dollars."

- “Remember that ***p*** equals two dollars and twenty-five cents.”
- “We can substitute this number [pop] for the letter *p* [pop] in the equation.”
- “So, ***two times*** *two dollars and twenty-five cents* [copy down] **plus *w* equals five dollars**.”
- “Since two times two dollars and twenty-five cents is ***four** dollars and fifty cents*, four dollars and fifty cents **plus** *w* equals five dollars.”
- “Therefore, *w* must equal [] **fifty** cents.”

C2-S2 Control 2 Story 2

- “On Day 1, I went to Martin’s Pizzeria…”
- “And I bought *two* slices of cheese pizza [pause] and *one* bottle of water [pause] for *nine* dollars.”
- “On Day 2, I went *back* to Martin’s Pizzeria…”
- “And I bought *five* slices of cheese pizza [pause] and *one* bottle of water [pause] for *eighteen dollars and seventy-five cents*.”

C2-Q2.1 Control 2 Question 2.1

- “**How** much money did *one slice of cheese pizza* cost at Martin’s?”

C2-[Q2.1Inc]-E2.1 Control 2 Explanation 2.1 [Q2.1 Incorrect]

- “Let’s use the letters *p* and *w* again as our variables.”
- “Looking at the information from **Day** 1, we can write that ***two*** *p* plus *w* equals nine dollars.”

- "And looking at the information from **Day** 2, we can write that ***five*** *p* plus *w* equals eighteen dollars and seventy-five cents."
- "Now we can eliminate *w* by taking our Day 2 equation [copy move] and *subtracting* [-] our Day 1 equation [copy move]."
- "The *difference* in *p* [box] between Day 1 and Day 2 is found by 5*p* minus [-] 2*p*, which equals ***3p***."
- "The difference in *w* [box] [-] equals ***zero***."
- "And the difference in the *total cost* [box] [-] equals ***nine*** *dollars and seventy-five cents*."
- "So, 3*p* *plus* [+] zero **equals** nine dollars and seventy-five cents."
- "So, ***three*** *p* equals nine dollars and seventy-five cents."
- "Therefore, ***p*** equals three dollars and twenty-five cents."

C2-Q2.2 Control 2 Question 2.2

- "**How** much money did *one bottle of water* cost at Martin's?"

C2-[Q2.1Inc]-E2.2 Control 2 Explanation 2.2 [Q2.2 Incorrect]

- "Let's use the letters *p* and *w* again as our variables."
- "Looking at the information from **Day** 1, we can write that ***two*** *p* plus *w* equals nine dollars."
- "Remember that ***p*** equals three dollars and twenty-five cents."
- "We can substitute this number [pop] for the letter *p* [pop] in the equation."
- "So, ***two*** [] ***times*** [] *three dollars and twenty-five cents* [copy down] **plus** ***w*** **equals** **nine** **dollars**."

- "Since two times three dollars and twenty-five cents is ***six*** *dollars and fifty cents*, six dollars and fifty cents **plus** *w* equals nine dollars."
- "Therefore, *w* must equal [] **two** dollars and fifty cents."

Control Story 3: Relational

C3-L0 Control 3 Lesson 0

- "In this lesson, you will learn about ***substitution***."
- "Substitution is **another** method for solving problems with *two* variables."
- "In order to find a variable's value, you can **substitute** one variable for the *other*."
- "This is **useful** for problems where the elimination method is harder to use."
- "Let's practice using substitution to solve a two-variable problem."

C3-S0 Control 3 Story 0

- "My friend went to Ali's Bakery…"
- "And she **bought** one donut [pause] and one cupcake [pause] for eight dollars."
- "And **the** price of one cupcake was *three times* the price of one donut."

C3-L1.1 Control 3 Lesson 1.1

- "Our ***goals*** are to ***figure*** *out* how much money one donut cost, and **how** much money one cupcake cost."

- “For this problem, we can use the letter ***d*** as a *variable* that ***represents*** [] **the** price of one donut, and the letter ***c*** as a *variable* that ***represents*** [] **the** price of one cupcake.”
- “We know from my friend’s order [fade in Order info] that the total price of one donut and one cupcake is eight dollars.”
- “Now we can take this information and write a math equation using our two variables [pop d and c].”
- “The price of one donut is ***d*** and the price of one cupcake is ***c***. Since the two prices are *added together* to get the total cost, *d* ***plus*** *c* ***equals*** [] **eight** dollars.”
- “We also know about the *relationship* [fade in Relationship info] between the prices: that the price of one cupcake is *three times* the price of one donut.”
- “The price of one cupcake is ***c***, and three times the price of one donut is ***three*** *d*. Since these are equal to each other, we can write that *c* ***equals*** three *d*.”

C3-L1.2 Control 3 Lesson 1.2

- “Since we have *two* variables, we can *replace one* variable with the *other* variable to find the second variable’s value.”
- “We know that *c* [pop] equals *three d* [pop].”
- “Let’s copy the *order* equation below [copy move].”
- “Now, we can replace *c* [pop] with *three d* [pop], because they are *equal* [pop]. This is called ***substitution***.”
- “So, *d* plus *three d* [move] equals eight dollars.”
- “And since *d* plus three *d* is equal to ***four*** *d*, four *d* [move] [pause] **equals** [] **eight** dollars.”

- “Therefore, ***d*** [] **equals** [] ***two*** dollars.”
- “So we found that ‘the price of one donut’ [move out] ***is*** [] ***two** dollars*.”

C3-L1.3 Control 3 Lesson 1.3

- “Now that we have found the price of one donut [cross out], we can *take* this information and *use* it in our *equation* to figure out the price of one cupcake.”
- “Remember that ***d*** equals *two* dollars.”
- “Now let’s focus on the Order equation [fade out Relationship].”
- “We can substitute two dollars [pop] for the letter *d* [pop] in the equation.”
- “So, *two dollars* [copy down] plus *c* equals *eight* dollars.”
- “Therefore, ***c*** must **equal** [] ***six* dollars**.”
- “So we found that ‘the price of one cupcake’ [move out] ***is*** [] ***six** dollars*.”

C3-S1 Control 3 Story 1

- “My friend went to Patty’s Bakery, where the prices are *different*…”
- “And she **bought** *one* donut [pause] and ***one*** cupcake [pause] for ***four** dollars and fifty cents*.”
- “And **the** price of one cupcake was *twice* the price of one donut.”

C3-Q1.1 Control 3 Question 1.1

- “**How** much money did one donut cost at Patty’s?”

C3-[Q1.1Inc]-E1.1 Control 3 Explanation 1.1 [Q1.1 Incorrect]

- "Let's use the letters *d* and *c* again as our variables."
- "Looking at the information from my friend's ***order***, we can write that ***d*** plus *c* equals four dollars and fifty cents."
- "And looking at the ***relationship*** between the items, we can write that ***c*** equals *two d*."
- "Let's copy the *order* equation below [copy move]."
- "Now, we can *replace c* [pop] with *two d* [pop], because they are *equal* [pop]."
- "So, *d* plus *two d* [move] equals four dollars and fifty cents."
- "So ***three*** *d* [] **equals** [] **four** dollars and fifty cents."
- "Therefore, ***d*** [] **equals** [] ***one*** dollar and fifty cents."

C3-Q1.2 Control 3 Question 1.2

- "**How** much money did one cupcake cost at Patty's?"

C3-[Q1.2Inc]-E1.2 Control 3 Explanation 1.2 [Q1.2 Incorrect]

- "Let's use the letters *d* and *c* again as our variables."
- "Looking at the information from my friend's ***order***, we can write that ***d*** plus *c* equals four dollars and fifty cents."
- "Remember that ***d*** equals one dollar and fifty cents."
- "We can substitute one dollar and fifty cents [pop] for the letter *d* [pop] in the equation."
- "So, *one dollar and fifty cents* [copy down] plus *c* equals *four* dollars and fifty cents."
- "Therefore, ***c*** must **equal** [] ***three*** **dollars**."

C3-S2 Control 3 Story 2

- "My friend went to Isabella's Bakery…"
- "And she **bought** *one* donut [pause] and ***one*** cupcake [pause] for ***six*** *dollars and twenty-five cents*."
- "And **the** price of one cupcake was *four times* the price of one donut."

C3-Q2.1 Control 3 Question 2.1

- "**How** much money did one donut cost at Isabella's?"

C3-[Q2.1Inc]-E2.1 Control 3 Explanation 2.1 [Q2.1 Incorrect]

- "Let's use the letters *d* and *c* again as our variables."
- "Looking at the information from my friend's ***order***, we can write that ***d*** plus *c* equals six dollars and twenty-five cents."
- "And looking at the ***relationship*** between the items, we can write that ***c*** equals *four d*."
- "Let's copy the *order* equation below [copy move]."
- "Now, we can *replace c* [pop] with *four d* [pop], because they are *equal* [pop]."
- "So, *d* plus *four d* [move] equals six dollars and twenty-five cents."
- "So ***five*** *d* [] **equals** [] **six** dollars and twenty-five cents."
- "Therefore, ***d*** [] **equals** [] ***one*** dollar and *twenty-five* cents."

C3-Q2.2 Control 3 Question 2.2

- "**How** much money did one cupcake cost at Isabella's?"

C3-[Q2.2Inc]-E2.2 Control 3 Explanation 2.2 [Q2.2 Incorrect]

- "Let's use the letters *d* and *c* again as our variables."
- "Looking at the information from my friend's ***order***, we can write that ***d*** plus *c* equals six dollars and twenty-five cents."
- "Remember that ***d*** equals one dollar and twenty-five cents."
- "We can substitute one dollar and twenty-five cents [pop] for the letter *d* [pop] in the equation."
- "So, *one dollar and twenty-five cents* [copy down] plus *c* equals *six* dollars and *twenty-five* cents."
- "Therefore, ***c*** must **equal** [] ***five*** **dollars**."

Final Stories (Immediate Post-Test: FS1=Change Analogue and FS2=Relational Analogue)

C-FS1-Q1 Control Final Story Question 1

- "I went to Bill's Burgers and [] **bought** three hamburgers and two milkshakes for twenty dollars and forty-five cents."
- "Later I went back to Bill's Burgers and [] **bought** seven hamburgers and two milkshakes for thirty-nine dollars and forty-five cents."
- "**How** much money did one hamburger cost?"

C-FS1-Q2 Control Final Story Question 2

- "**How** much money did one milkshake cost?"

C-FS2-Q1 Control Final Story 2 Question 1

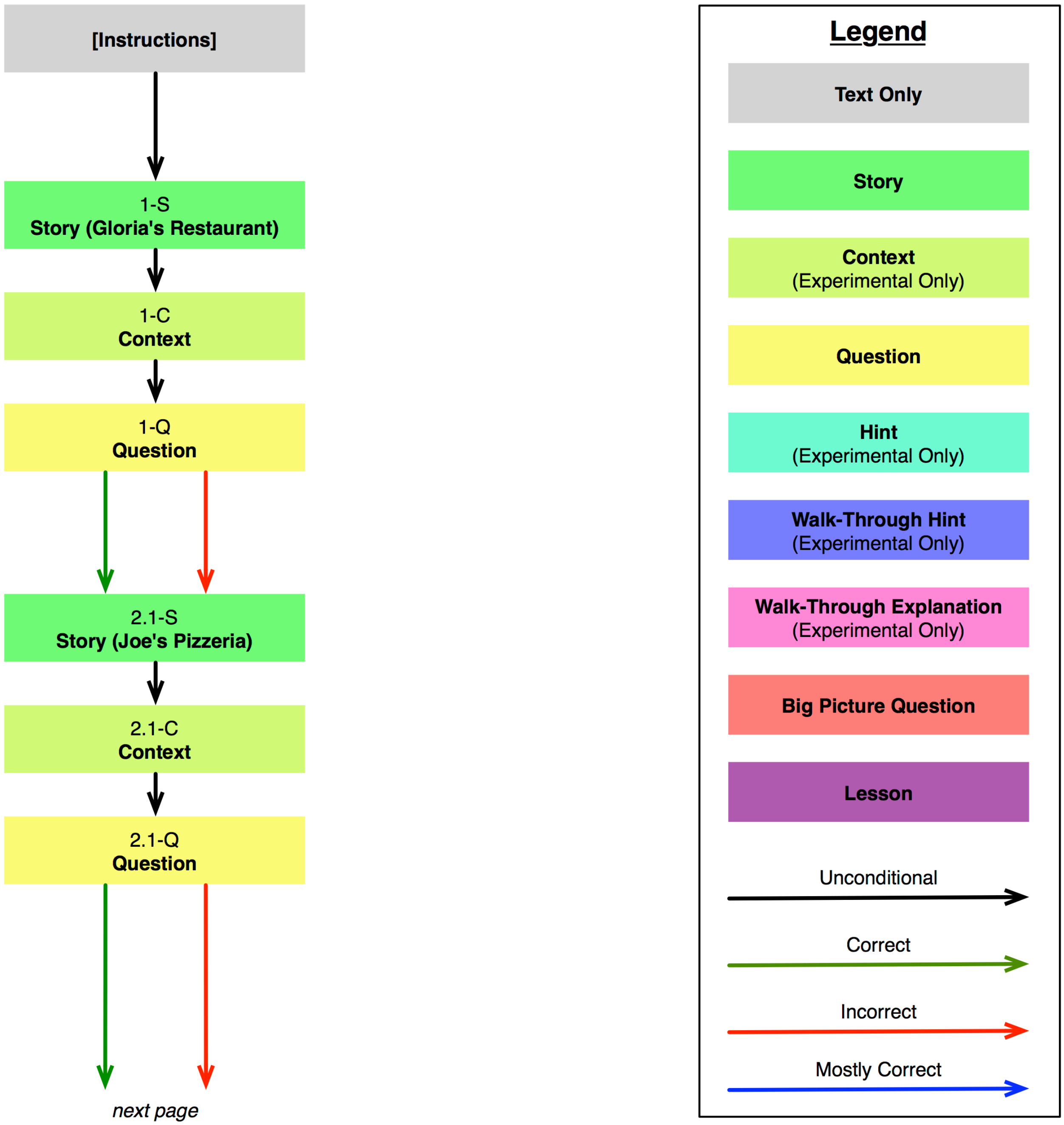
- "I went to Annie's Cafe and I [] **bought** one coffee and one slice of pie for five dollars and seventy cents."
- "And **the** price of one slice of pie was *five times* the price of one coffee."
- "**How** much money did one coffee cost?"

C-FS2-Q2 Control Final Story 2 Question 2

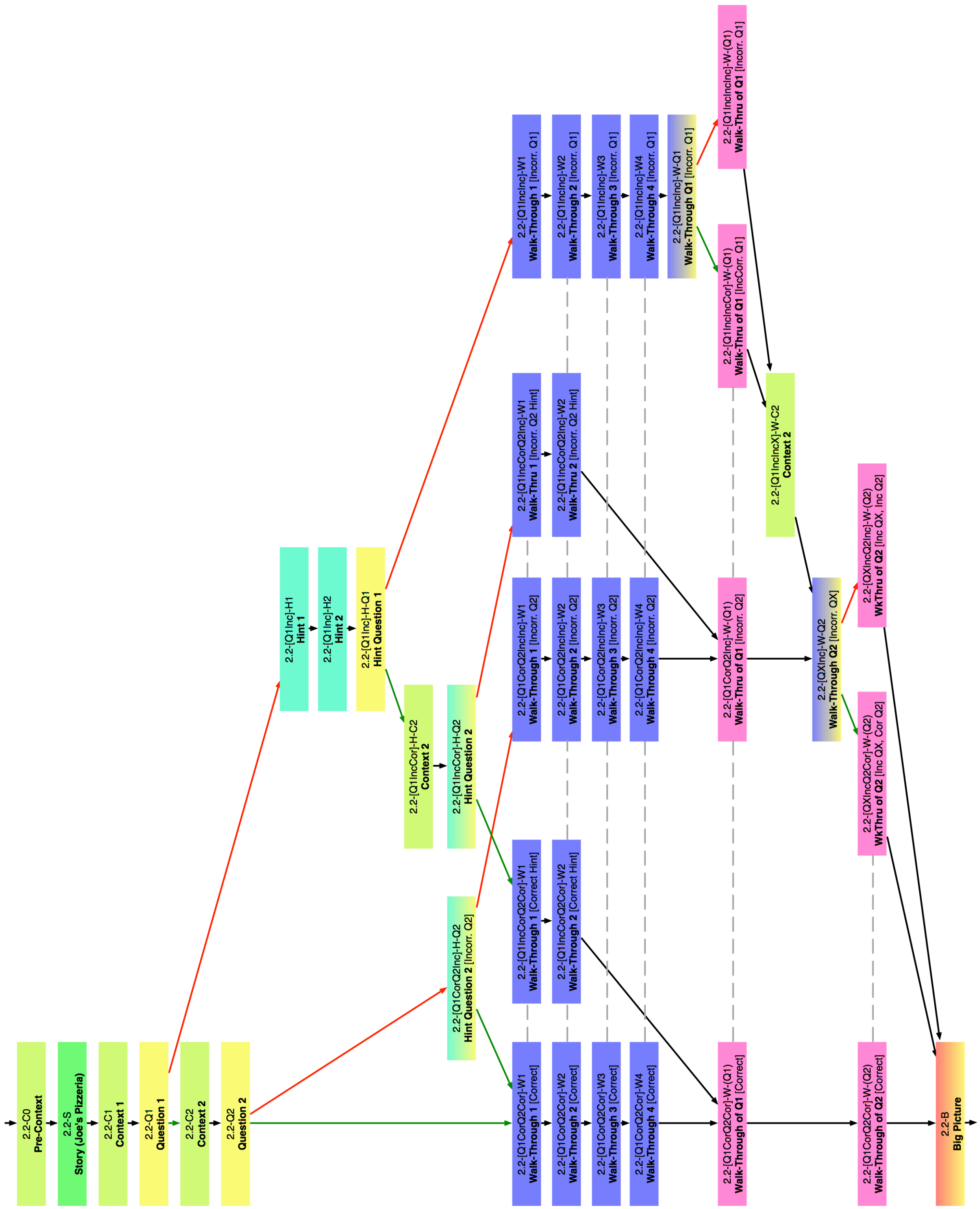
- "**How** much money did one slice of pie cost?"

APPENDIX C:

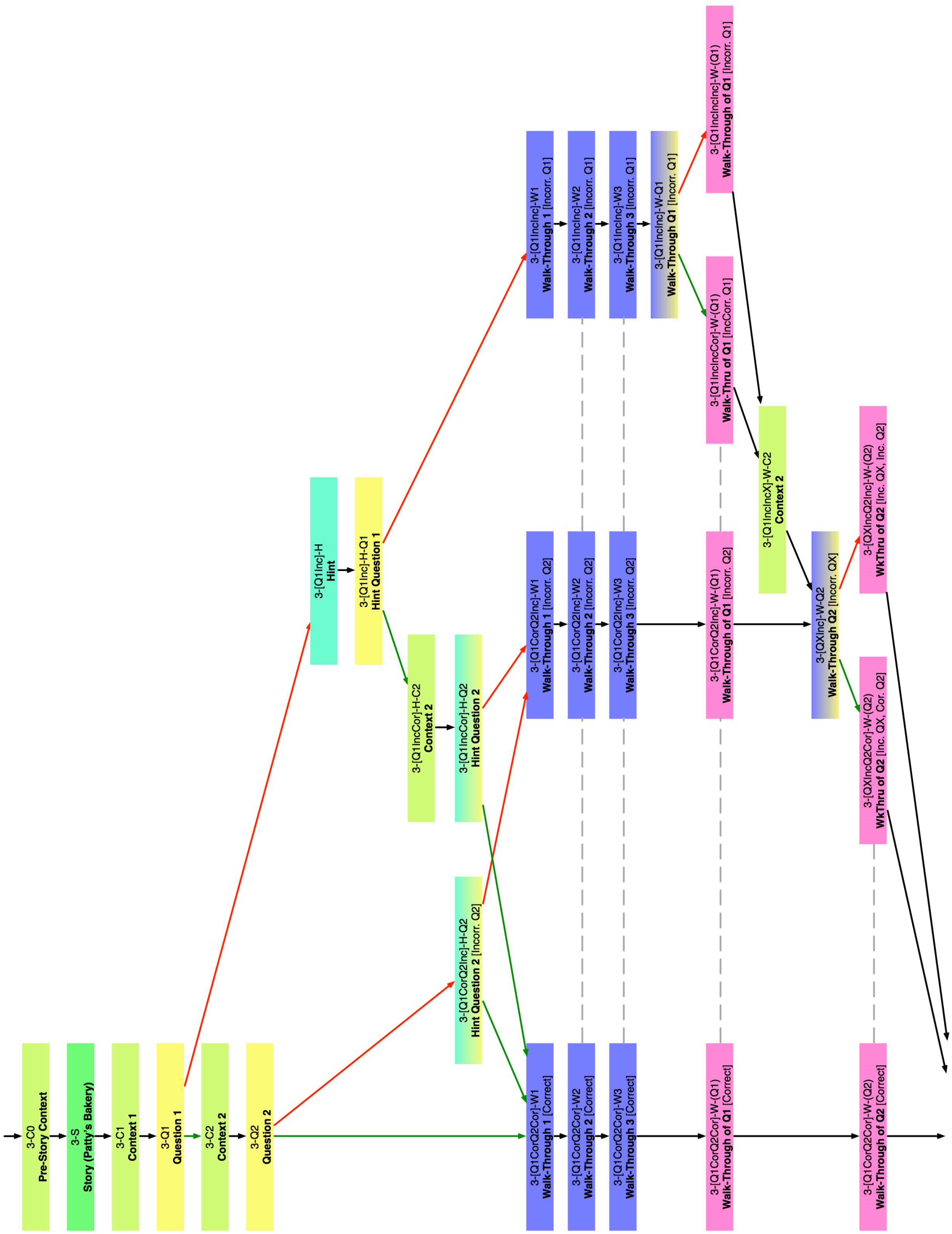
Experimental Condition Flowchart



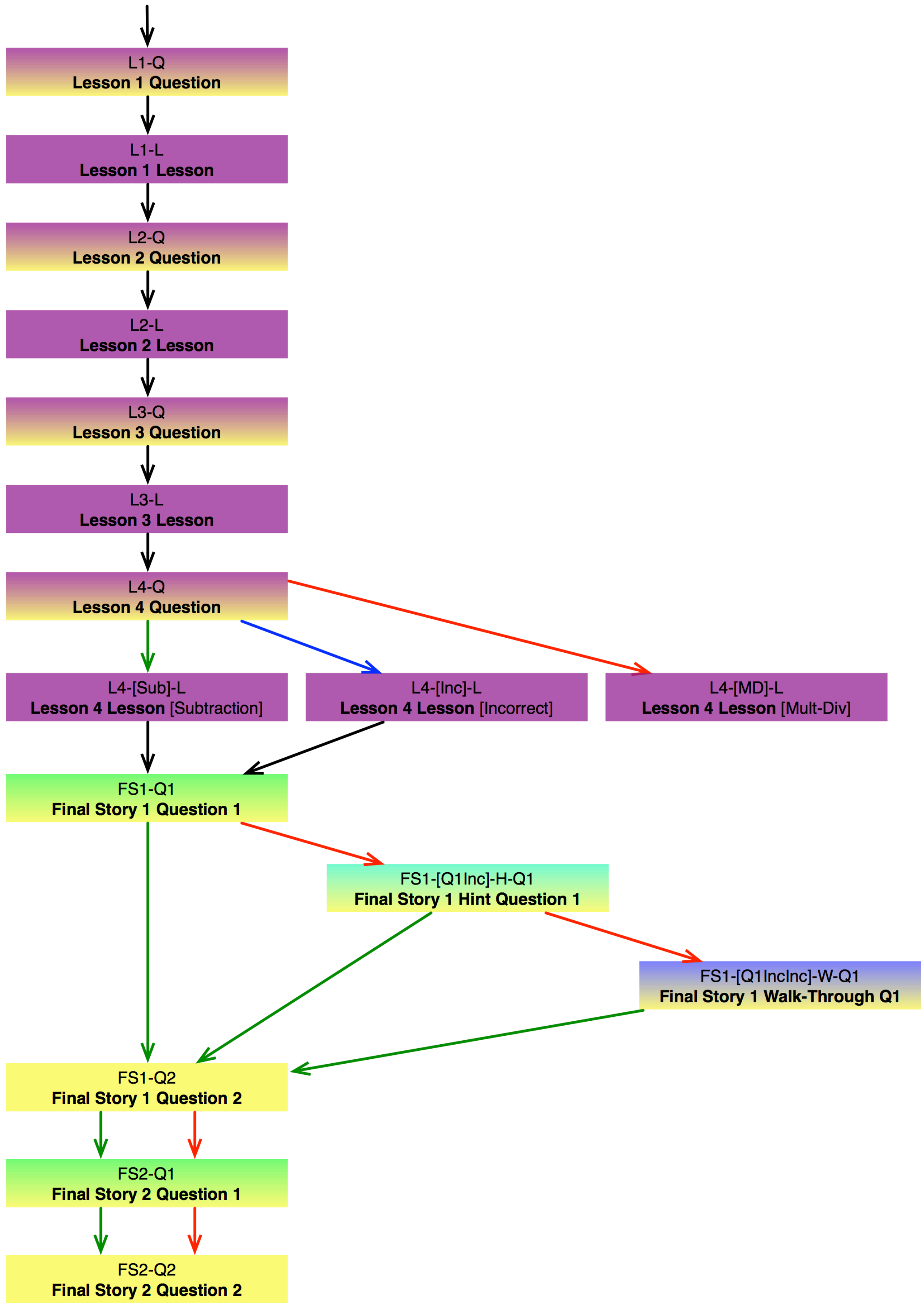
Instructions, Story 1 (One-Unknown), and Story 2.1 (Unsolvable Two-Unknowns)



Story 2.2 (Change)



Story 3 (Relational)



Lessons and Final Stories (Immediate Post-Tests: Change and Relational Analogues)

APPENDIX D:

Delayed Post-Test Items

1. I went to Terrell's Diner and I bought **4 salads** and **3 bowls of soup** for **$19.50**.

   The next week, I went back to Terrell's Diner and I bought **7 salads** and **3 bowls of soup** for **$28.50**.

   a. *How much money* did **1 salad** cost?
   b. *How much money* did **1 bowl of soup** cost?

2. My friend went to Mizuki's Café and she bought **1 sandwich** and **1 bag of chips** for **$13.00** total.

   The price of **1 sandwich** was *three times* the price of **1 bag of chips**.

   a. *How much money* did **1 sandwich** cost?
   b. *How much money* did **1 bag of chips** cost?

3. I went to Javier's Market and placed an order for some drinks. I ordered some **lemonades** for **$1.80** *each* and some **iced teas** for **$2.75** *each*. The total cost for all of the drinks together was **$15.45**.

   Later, I went to Maya's Groceries, and I placed the same order as I did at Javier's — I bought the same number of each type of drink. At Maya's, the **lemonades** cost **$2.10** *each*, and the **iced teas** cost **$2.75** *each*. The total cost for all of the drinks at Maya's was **$16.65**.

   a. *How many* **lemonades** did I buy from each place?

   b. *How many* **iced teas** did I buy from each place?

4. $\mathbf{3}\boldsymbol{x} + \boldsymbol{y} = \mathbf{14}$

   $\boldsymbol{y} = \mathbf{4}\boldsymbol{x}$

   a. *What number* does $\boldsymbol{x}$ represent?

   b. *What number* does $\boldsymbol{y}$ represent?

5. Sophia has a total of **24 camels** on her ranch: *some* with **1 hump** and *some* with **2 humps**. Altogether, her **camels** have a total of **31 humps**.

   a. *How many* **camels** with ***one*** **hump** does Sophia have?

   b. *How many* **camels** with ***two*** **humps** does Sophia have?

6. $7 = 3 + 4$ Is the equation to the left true? Circle one: YES NO

7. $n + n = 18$ What can n be to make this equation true?

8. $8 + 4 = x + 5$ What number does x need to be to make this equation true?

9. $2y – 6 = y + 2$ What number does y need to be to make this equation true?

10. a and b both stand for numbers, and $a = b + 2$. Which is the larger number, a or b?

    Circle one: a b

11. Is $3t + 5t = 8t$ true no matter what number t stands for? Circle one: YES NO

12. Cakes cost c dollars each and brownies cost b dollars each. Suppose I buy 1 cake and 2 brownies.

    What does $c + 2b$ stand for?

    Circle one: the number of desserts how much money the desserts cost

13. David is 10 inches taller than Carla. Carla is h inches tall. Which of the following expressions represents David's height in terms of h?

Circle one:

a) $h - 10$

b) $h + 10$

c) $10h$

Frank is twice as tall as Carla. What could you write to represent Frank's height in terms of h?

Frank's height =

14. A friend gives you some money. Can you tell which is larger:

- the amount of money your friend gives you plus six more dollars, OR
- three times the amount of money your friend gives you?

Circle one:

a) "The amount of money your friend gives you plus six more dollars" is larger

b) "Three times the amount of money your friend gives you" is larger

c) One or the other can be larger, depending on the amount of money your friend gives you.

REFERENCES

Ainley, J. (2012). Developing purposeful mathematical thinking: a curious tale of apple trees. *PNA, 6*(3), 85-103.

Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. *Educational researcher*, *25*(4), 5-11.

Baranes, R., Perry, M., & Stigler, J. W. (1989). Activation of real-world knowledge in the solution of word problems. *Cognition and Instruction, 6*(4), 287–318.

Bassok, M. (1996). Using content to interpret structure: Effects on analogical transfer. *Current Directions in Psychological Science*, *5*(2), 54-58.

Behr, M., Erlwanger, S., & Nichols, E. (1976). How children view equality sentences, Project for the Mathematical Development of Children. *ERIC Document Reproduction Services No. ED144802*.

Bell, A. (1995). Purpose in school algebra. *The Journal of Mathematical Behavior, 14*(1), 41–73.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. Shimamura (Eds.), *Metacognition: Knowing about Knowing* (pp. 185-205). Cambridge, MA: MIT Press.

Booth, L. R. (1988). Children's difficulties in beginning algebra. *The ideas of algebra, K-12*, 20-32.

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

Carpenter, T. P., & Levi, L. (2000). *Developing conceptions of algebraic reasoning in the primary grades* (Research Report #002). Madison, WI: National Center for Improving Student Learning and Achievement in Mathematics and Science.

Carraher, T. N., Carraher, D. W., & Schliemann, A. D. (1987). Written and oral mathematics. *Journal for Research in Mathematics Education*, 83-97.

Chazan, D., Sela, H., & Herbst, P. (2012). Is the role of equations in the doing of word problems in school algebra changing? Initial indications from teacher study groups. *Cognition and Instruction*, *30*(1), 1-38.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological review*, *104*(2), 367.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive psychology*, *17*(4), 391-416.

Chi, M., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, NJ: Erlbaum.

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive science*, *13*(2), 145-182.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational psychology review*, *3*(3), 149-210.

Clement, J., Lochhead, J., & Monk, G. S. (1981). Translation difficulties in learning mathematics. *The American Mathematical Monthly, 88*(4), 286–290.

Clement, J. (1982). Algebra word problem solutions: Thought processes underlying a common misconception. *Journal for research in mathematics education, 13*(1), 16–30.

Collis, K. F. (1975). *A Study of Concrete and Formal Operations in School Mathematics: A Piagetian Viewpoint*.

Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule induction on mathematical problem-solving transfer. *Journal of Educational Psychology, 79*, 347-362.

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology, 20*(4), 405–438.

Davis, R. B. (1975). Cognitive processes involved in solving simple algebraic equations. *Journal of Children's Mathematical Behavior 1*(3), 7–35.

DeLoache, J. S. (1995). Early symbol understanding and use. *Psychology of Learning and Motivation*, *33*, 65-116.

Evans, J. S. B. (2003). In two minds: dual-process accounts of reasoning. *Trends in cognitive sciences*, *7*(10), 454-459.

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

Freudenthal, H. (1987). Mathematics starting and staying in reality. In I. Wirszup & R. Street (Eds.), Proceedings of the USCMP conference on mathematics education on development in school mathematics education around the world. Reston, VA: National Council of Teachers of Mathematics.

Gellert, U., & Jablonka, E. (2009). "I am not talking about reality". In L. Verschaffel et al. (Eds.) *Words and worlds: Modelling verbal descriptions of situations* (pp. 39-53). Rotterdam: Sense Publishers.

Gerofsky, S. (1996). A linguistic and narrative view of word problems in mathematics education. *For the Learning of Mathematics, 16*(2), 36–45.

Gerofsky, S. (1999). *The word problem as genre in mathematics education.* (Unpublished doctoral dissertation). Simon Fraser University, Burnaby, BC.

Gerofsky, S. (2009). Genre, simulacra, impossible exchange, and the real. In L. Verschaffel et al. (Eds.) *Words and worlds: Modelling verbal descriptions of situations* (pp. 21-38). Rotterdam: Sense Publishers.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive psychology*, *12*(3), 306-355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive psychology*, *15*(1), 1-38.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences*, *14*(1), 69-110.

Gray, E. M., & Tall, D. O. (1994). Duality, ambiguity, and flexibility: A proceptual view of simple arithmetic. *Journal for research in mathematics education, 26*(2), 115–141.

Greeno, J. G. (1978). Natures of problem-solving abilities. *Handbook of learning and cognitive processes*, *5*, 239-270.

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42*(1), 335-346.

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

Hiebert, J. (1988). A theory of developing competence with written mathematical symbols. *Educational Studies in Mathematics*, *19*(3), 333–355.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. *Second handbook of research on mathematics teaching and learning*, *1*, 371-404.

Hinsley, D. A., Hayes, J. R., & Simon, H. A. (1977). From words to equations: Meaning and representation in algebra word problems. *Cognitive processes in comprehension*, *329*.

Hull, C. L. (1943). *Principles of behavior.* New York: Appleton-Century-Crofts.

Jacobs, V. R., Franke, M. L., Carpenter, T. P., Levi, L., & Battey, D. (2007). Professional Development Focused on Children's Algebraic Reasoning in Elementary School. *Journal for Research in Mathematics Education, 38*(3), 258–288.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science, 320*, 454-455.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. (2009). Transfer of mathematical knowledge: The portability of generic instantiations. *Child Development Perspectives*, *3*(3), 151-155.

Kaput, J. J. (1998). Transforming Algebra from an Engine of Inequity to an Engine of Mathematical Power by "Algebrafying" the K-12 Curriculum. In National Council of Teachers of Mathematics & Mathematical Sciences Education Board (Eds.), *The nature and role of algebra in the K-14 curriculum* (pp. 25–26).

Kaput, J. J. (1999). Teaching and learning a new algebra. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 133-155). Mahwah, NJ: Erlbaum.

Kieran, C. (1992). The learning and teaching of algebra. *Handbook of research on mathematics teaching and learning*, 390–419.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological review*, *92*(1), 109.

Klein, J. (1968). *Greek mathematical thought and the origin of algebra*. (E. Brann, Trans.). Cambridge, MA: MIT Press. (Original work published 1934-1936).

Kline, M. (1973). *Why Johnny can't add: the failure of the new math*. New York: St. Martin's Press.

Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. C. (2005). Middle school students' understanding of core algebraic concepts: Equivalence & Variable. *Zentralblatt für Didaktik der Mathematik, 37*(1), 68–76

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, 297–312.

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.

Koedinger, K. R. & Nathan, M. J. (2004). The real story behind story problems: Effects of representation on quantitative reasoning. The Journal of the Learning Sciences, 13(2), 129-164.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*(6), 2797-2822.

Kraft, N. (2013, May 5). Teaching Algebra without algebra [Blog post]. Retrieved from: http://nathankraft.blogspot.com/2013/05/teaching-algebra-without-algebra.html

Küchemann, D. (1978). Children's understanding of numerical variables. *Mathematics in School, 7*(4), 23–26.

Lee, L., & Wheeler, D. (1989). The arithmetic connection. *Educational Studies in Mathematics*, *20*(1), 41-54.

Linchevski, L. (1995). Algebra with numbers and arithmetic with letters: A definition of pre-algebra. *The Journal of Mathematical Behavior, 14*(1), 113-120.

Mayer, R. E. (1982). Different problem-solving strategies for algebra word and equation problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(5), 448–462.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, *38*(1), 43-52.

Nathan, M. J. (2008). An embodied cognition perspective on symbols, gesture and grounding instruction. *Symbols and embodiment: Debates on meaning and cognition*, 375-396.

Nathan, M. J. (2012). Rethinking formalisms in formal education. Educational Psychologist, 47(2), 125–148.

Nathan, M. J., & Koedinger, K. R. (2000a). An investigation of teachers' beliefs of students' algebra development. *Cognition and Instruction*, *18*, 207–235.

Nathan, M. J., & Koedinger, K. R. (2000b). Teachers' and researchers' beliefs about the development of algebraic reasoning. *Journal for Research in Mathematics Education*, *31*, 168–190.

Nathan, M. J., Koedinger, K. R., & Alibali, M. W. (2001). Expert blind spot: When content knowledge eclipses pedagogical content knowledge (pp. 644–648). Presented at the Proceedings of the Third International Conference on Cognitive Science, Beijing, China.

Nathan, M. J., Long, S. D., & Alibali, M. W. (2002). The symbol precedence view of mathematical development: A corpus analysis of the rhetorical structure of textbooks. *Discourse Processes, 33*(1), 1–21.

National Council for Teachers of Mathematics (2000). *Principles and standards for school mathematics.* Reston, VA: NCTM.

Nesher, P. (1980). The stereotyped nature of school word problems. *For the Learning of Mathematics, 1*(1), 41-48.

Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological review*, *65*(3), 151.

Piaget, J. (1970). *Science of Education and the Psychology of the Child,* Orion Press, New York.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of engineering education*, *93*(3), 223-231.

Pólya, G. *How to solve it.* Princeton, NJ: Princeton Univ. Press, 1957.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of comparative and physiological psychology*, *66*(1).

Romberg, T. A. (2001). Designing middle school mathematics materials using problems created to help students progress from informal to formal mathematical reasoning. In L. P. Leutzinger & S. P. Smith (Eds.), *Mathematics in the middle* (pp. 107–119). Reston, VA: National Council of Teachers of Mathematics.

Rosnick, P. (1981). Some misconceptions concerning the concept of variable. *The Mathematics Teacher, 74*(6), 418-450.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and brain sciences*, *3*(03), 417-424.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*(1), 1–36.

Sfard, A. (1995). The development of algebra: Confronting historical and psychological perspectives. *The Journal of Mathematical Behavior*, *14*(1), 15-39.

Sfard, A., & Linchevski, L. (1994). The gains and the pitfalls of reification—the case of algebra. In *Learning mathematics* (pp. 87-124). Springer Netherlands.

Sherman, M. F., Walkington, C., & Howell, E. (2016). A Comparison of Symbol-Precedence View in Investigative and Conventional Textbooks Used in Algebra Courses. *Journal for Research in Mathematics Education*, *47*(2), 134-146.

Simon, D., & Simon, H. (1978). Individual differences in solving physics problems. In R. Siegler, (Ed.), *Children's thinking: What develops?* NJ: Erlbaum.

Stanovich, K.E. and West, R.F. (2000) Individual differences in reasoning: Implications for the rationality debate. *Behav. Brain Sci. 23*, 645–726.

Stigler, J. W., Fuson, K. C., Ham, M., & Sook Kim, M. (1986). An analysis of addition and subtraction word problems in American and Soviet elementary mathematics textbooks. *Cognition and Instruction, 3*(3), 153-171.

Stigler, J. W., Geller, E. H., & Givvin, K. B. (2015). Zaption: A Platform to Support Teaching, and Learning about Teaching, with Video. *Journal of e-Learning and Knowledge Society*, *11*(2).

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. Simon and Schuster.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), 257-285.

Swetz, F. J. (2009). Word problems: Footprints from the history of mathematics. In L. Verschaffel et al. (Eds.) *Words and worlds: Modelling verbal descriptions of situations* (pp. 73-91). Rotterdam: Sense Publishers.

Uttal, D. H., Liu, L. L., & DeLoache, J. S. (1999). Taking a hard look at concreteness: do concrete objects help young children learn symbolic relations?.

Uttal, D. H., Liu, L. L., & DeLoache, J. S. (1999). Taking a hard look at concreteness: Do concrete objects help young children learn symbolic relations? In C. S. Tamis-LeMonda (Ed.), *Child psychology: A handbook of contemporary issues* (pp. 177–192). Philadelphia, PA: Psychology Press.

Uttal, D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: A new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology, 18*, 37–54.

Verschaffel, L., De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning and Instruction, 4*(4), 273–294.

Verschaffel, L., De Corte, E., & Borghart, I. (1997). Pre-service teachers' conceptions and beliefs about the role of real-world knowledge in mathematical modelling of school word problems. *Learning and Instruction, 7*(4), 339–359.

Walker, J. M., Cheng, P. W., & Stigler, J. W. (2014). Equations are effects: Using causal contrasts to support algebra learning. In P. Bello, M. Guarini, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1730–1735). Austin, TX: Cognitive Science Society.

Wheeler, D. (1989). Contexts for research on the teaching and learning of algebra. *Research issues in the learning and teaching of algebra, 4*, 278.