Slow Down to Speed Up:

Longer Pause Time Before Solving Problems Relates to Higher Strategy Efficiency

Jenny Yun-Chen Chan*

Erin R. Ottmar

Ji-Eun Lee

Social Science and Policy Studies, Worcester Polytechnic Institute

100 Institute Road, Worcester, MA 01609 USA

jchan2@wpi.edu

erottmar@wpi.edu

jlee13@wpi.edu

* Corresponding author: jchan2@wpi.edu

Acknowledgments
The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through an Efficacy and Replication Grant (R305A180401) to Worcester Polytechnic Institute, and two Small Business Innovation Research Grants (#91990018C0032; 91990019C0034) to Graspable Inc. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. We thank Katharine Sawrey, Taylyn Hulse, Alisionna Iannacchione, Jennifer St. John, David Brokaw, and members of the Math Abstraction Play Learning and Embodiment Lab for their work on this project, and the participating teachers and students for their support.

Author Contribution Statement
Jenny Yun-Chen Chan: Conceptualization, Investigation, Formal analysis, Writing, Visualization, Project administration. **Erin Ottmar**: Conceptualization, Investigation, Writing, Supervision, Funding acquisition. **Ji-Eun Lee**: Data curation, Writing, Visualization.

Disclosure Statement
Erin Ottmar is a designer and a developer of From Here to There! and owns a 10% equity in Graspable Inc. This has been disclosed to WPI's Conflict Management Committee, and a conflict management plan has been implemented.

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

Abstract

We examined the influences of pre-solving pause time, algebraic knowledge, mathematics self-efficacy, and mathematics anxiety on middle-schoolers' strategy efficiency in an algebra learning game. We measured strategy efficiency using (a) the number of steps taken to complete a problem, (b) the proportion of problems completed on the initial attempt, and (c) the number of resets prior to completing the problems. Using the log data from the game, we found that longer pre-solving pause time was associated with more efficient strategies, as indicated by fewer solution steps, higher initial completion rate, and fewer resets. Higher algebraic knowledge was associated with higher initial completion rate and fewer resets. Mathematics self-efficacy and mathematics anxiety was not associated with any measures of strategy efficiency. The results suggest that pause time may be an indicator of student thinking before problem-solving, and provide insights into using data from online learning platforms to examine students' problem-solving processes.

Keyword: pause time, strategy efficiency, algebra problem-solving, online learning environment, metacognitive skills

Slow Down to Speed Up: Longer Pause Time before Solving Problems Relates to Higher Strategy Efficiency

Efficient and flexible problem-solving is a primary goal in mathematics education (Common Core State Stands, 2010), and noticing mathematical structures is an important foundation for learning algebra (Kaput, 1998; Venkat et al., 2019). However, students struggling with algebra often do not notice the structures that afford efficient strategies (Carpenter et al., 1980; Star & Rittle-Johnson, 2008). For example, $3(2+x) = 18$ can be solved either by using a three-step standard strategy: 1) distribute 3 into the parentheses, 2) subtract 6 from both sides, and 3) divide both sides by 3; or a two-step more efficient strategy: 1) divide both sides by 3, and 2) subtract 2 from both sides. Although the latter versus former strategy involves fewer steps and more simplified computations, it requires students to notice the multiplicative relation between 3 and 18. Students rushing into the problem may apply the distribution procedure by rote (i.e., PEMDAS: Parenthesis, Exponents, Multiplication, Division, Addition, Subtraction) without noticing or leveraging the relations between numbers. To better support the development of strategic problem-solving, researchers have examined how factors such as mathematical knowledge and affective dispositions influence students' strategy selection and efficiency (Newton et al., 2020; Pajares, 1996; Ramirez et al., 2016; Star & Rittle-Johnson, 2008). We extend prior research by examining whether *pause time before solving* predicts strategy efficiency.

1.1 Strategy Efficiency

Strategy efficiency has been operationalized as selecting a strategy that uses the fewest steps and/or the computation that involves small, whole numbers rather than large numbers or fractions (Xu et al., 2017). Strategy efficiency is an important skill in mathematics because it reflects students' understanding of mathematical structures (Robinson et al., 2006; Venkat et al., 2019), and solving problems efficiently allows students to reserve cognitive resources for challenging contents. It is hypothesized that the ability to use efficient strategies requires specialized mathematical knowledge (e.g., understanding of equivalence), and such knowledge may be more important than general cognitive factors (e.g., intelligence; Hoffman & Schraw, 2010). Other researchers argue that affective factors, such as mathematics self-efficacy and anxiety, also contribute to students' uses of efficient strategies, and thus should not be ignored (Mayer, 1998; Ramirez et al., 2016).

Related to but distinct from efficient strategies, many students believe that mathematics problems should be solved quickly within five minutes (Schoenfeld, 1992). These kinds of beliefs about mathematics may contribute to students' tendency to rush through problems rather than practice *pausing,* which may indicate thinking about the problem at hand. Some studies have shown a positive relation between pausing and mathematical performance (e.g., Paquette et al., 2014). Specifically, algebra students with long pause time between steps (≥ six seconds) tend to show improved learning in an online environment compared to those with short pause time (≤ five seconds). However, little is known about the impact of pause time on students' strategy use in equation-solving.

Here, we use pre-solving pause time (hereafter pause time)—the time between the start of a problem and students' first action—as a proxy indicator of thinking before problem-solving. We examine the unique influences of pause time on students' strategy efficiency within an online algebra game. In this game, students can reset and reattempt problems multiple times, and all their actions are time-stamped and recorded. Therefore, in addition to the traditional measure of strategy efficiency (i.e., the number of steps taken on a problem), we include two relevant measures of strategy efficiency—initial success of problem-solving and frequency of resetting problems—to further investigate its relation with pause time. In the following sections, we review the literature on the influences of pause time, mathematical knowledge, and affective dispositions on problem-solving strategies.

1.2 Pause Time as a Proxy Indicator of Thinking Before Problem-Solving

To apply efficient strategies, students need to notice the mathematical structure of problems, and inhibit applying standard procedures by rote. However, students often try to solve problems quickly without noticing important structures or implementing efficient strategies (Schoenfeld, 1992). A recent study indicated that fifth- and sixth-graders varied in whether they took the time to process and understand mathematical problems before solving, and this behavior was associated with more accurate judgement of problem-solving performance (García, Rodríguez, González-Castro, et al., 2016). Pausing, therefore, may support understanding of the problem, and help students suppress more automatic yet incorrect or inefficient strategies in favor of more efficient but not so obvious strategies.

Metacognition refers to the knowledge of one's own cognitive processes, and the ability to regulate and monitor the processes (Flavell, 1976). Some strategic regulation of behaviors, such as orienting to the task and planning the actions, reflect metacognitive skills and contribute to students' success in mathematical tasks (Garofalo & Lester, 1985). For example, fifth- and sixth-graders who spent more time on a problem were more likely to solve the problem correctly (García et al., 2019), and they reported planning strategies to a greater extent than their peers who solved the problem incorrectly. Further, prompting students to think about what, when, and why certain strategies should be applied, improved ninth-graders' mathematical learning in an online environment (Kramarski & Gutman, 2006). Similarly, fifth-graders who received an intervention that involved understanding the problem and devising strategies showed improvement on problem-solving (Vula et al., 2017). Together, these studies suggest that spending time to understand the problem and devise a plan *before* solving may be a pathway to efficient problem-solving.

We posit that pause time may be a proxy indicator of thinking about and planning for the problem at hand. Adults who paused longer prior to their first move in the Tower of Hanoi task (rearranging disks according to specified rules) completed the task with fewer moves, and they reported using that time for planning (Welsh et al., 1995). Similarly, students who spent more versus less time on developing plans prior to responding perform better on critical thinking tasks that involved hypothesis testing and argument analysis (Ku & Ho, 2010). Further, researchers have used the number and duration of pauses as indicators of students' cognitive engagement in an online science intelligent tutoring system (Gobert et al., 2015). Similarly, in Cognitive Tutor Algebra, long pause time before solving or requesting support (≥ six seconds) was used as an indicator of thinking, whereas short pause time (≤ five seconds) was used as an indicator of guessing (Paquette et al., 2014). Even though longer pause time may indicate thinking and planning, and associate with more efficient strategies, it may also contribute to and extend the total problem-solving time, leading to longer time on a problem. Therefore, Li and colleagues (2015) accounted for students' total problem-solving time by computing the percent pause time (i.e., pre-solving pause time / total problem-solving time), and found that longer percent pause time still significantly predicted more efficient strategies, as measured by the number of steps taken to solve a puzzle.

Together, these prior studies demonstrate the importance of pause time on students' problem-solving performance; however, they do not account for the potential influences of knowledge and affective factors, which are known to predict strategy efficiency. Therefore, we simultaneously estimate the effects of pause time, mathematical knowledge, self-efficacy, and anxiety to examine their relative influences on students' strategy efficiency. Further, we compute and use students' percent pause time in the analyses to account for the potential individual variation in students' total problem-solving time.

1.3 Mathematical Knowledge

To tailor strategies to a problem, students need to have the content knowledge—knowledge of underlying concepts and procedures to carry out solutions (Star & Rittle-Johnson, 2008). This association between knowledge and strategy use has been demonstrated across age groups and mathematics topics. For example, second-graders with low mathematics achievement tend to use one familiar strategy when solving arithmetic equations, whereas students with higher mathematics achievement show more variation in their strategy use, and are more likely to select efficient strategies to achieve the answer more quickly and accurately (Torbeyns et al., 2006). Similarly, middle-schoolers with higher versus lower mathematics achievement are more likely to use a strategy that involves fewer steps when solving algebraic equations (Newton et al., 2020; Wang et al., 2019).

Beyond the correlational findings, the influence of knowledge on strategy use has been examined in experiments that aim to increase students' algebraic knowledge. For example, after an instructional intervention that demonstrates multiple strategies of solving algebraic equations, sixth-graders are more likely to use more efficient strategies that involve fewer steps (Star & Rittle-Johnson, 2008). Although mathematical knowledge and strategy use seem to be tightly related to each other, they are only moderately correlated ($r = .27$; Xu et al., 2017). Students with the knowledge of multiple strategies still do not use the most efficient strategy frequently or consistently, likely because the knowledge tends to precede the ability of implementing efficient strategies. This finding also suggests that other factors, such as students' affective dispositions, may contribute to their strategy selection.

1.4 Mathematics Self-Efficacy and Anxiety

Students' affective dispositions towards mathematics, including self-efficacy and anxiety, are associated with their general mathematics performance, and strategy efficiency specifically. Self-efficacy refers to an individual's belief in their own ability to perform the behaviors that achieve a specific outcome (Bandura, 1977). *Mathematics self-efficacy* refers to the belief in one's own ability to do and learn mathematics. Students with high mathematics self-efficacy tend to perform well in mathematics (Fast et al., 2010; Pajares, 1996). *Mathematics anxiety*, first introduced by Dreger and Aiken as number anxiety (1957), refers to the feelings of fear, tension, and apprehension when engaging with mathematics (Ashcraft, 2002). Students with higher versus lower mathematics anxiety tend to perform worse, especially on complex equations and problems (Ramirez et al., 2018; Wu et al., 2012). Although mathematics self-efficacy and mathematics anxiety seem to be related to each other, they capture distinct aspects of affective dispositions and are not mutually exclusive. For example, a student can have high mathematics self-efficacy but still suffer from high mathematics anxiety.

Prior work suggests that students' affective dispositions may influence their mathematics performance through learning approaches and strategy choices. One on hand, fifth- and sixth-graders with higher versus lower mathematics self-efficacy tend to approach mathematical learning with the goal of a deep understanding; on the other hand, students with higher versus lower mathematics anxiety tend to approach mathematical learning with the goal of passing the course. This surface level approach to learning negatively predicts mathematics achievement

(García, Rodríguez, Betts, et al., 2016). Further, first- and second-graders' strategy choice mediates the association between mathematics anxiety and achievement (Ramirez et al., 2016). Students with higher mathematics anxiety tend to use advanced retrieval-based strategies less frequently, leading to lower mathematics achievement.

Focusing on strategy efficiency, students' mathematics self-efficacy tends to be positively associated with strategy efficiency. For example, undergraduate and graduate students' mathematics self-efficacy positively predicts equation-solving efficiency, as measured by accuracy and reaction time, above and beyond working memory capacity (Hoffman, 2010). Furthermore, increasing ninth-graders' self-efficacy is related to more accurate and efficient mathematics problem-solving that involves fewer help-seeking behaviors (Bernacki, et al., 2015). Conversely, students' mathematics anxiety is negatively related to their strategy efficiency. For example, undergraduate students with higher versus lower mathematics anxiety tend to select less efficient strategies for addition problems (Ashkenazi & Najjar, 2018). The negative impact of mathematics anxiety is also observed in primary (Ramirez et al., 2016) and secondary (Passolunghi, et al., 2016) students' approaches to calculations. Students with higher versus lower mathematics anxiety tend to be slower at written calculations and use fact retrieval less often.

Together, these findings suggest that strategy efficiency is positively associated with mathematics self-efficacy (Bernacki et al., 2015; Hoffman, 2010) and negatively associated with mathematics anxiety (Ashkenazi & Najjar, 2018; Passolunghi et al., 2016; Ramirez et al., 2016). Given that mathematical knowledge, self-efficacy, and anxiety are known correlates of strategy efficiency, it is important to consider these factors simultaneously with pause time to estimate their unique relations with strategy efficiency. Doing so will allow researchers and educators to identify predictors of strategy efficiency and inform instructional designs that support students' problem-solving.

1.5 The Current Study

We examine the influences of pause time, mathematical knowledge, mathematics self-efficacy, and mathematics anxiety on students' strategy efficiency, as measured by solution steps (Figure 1a), initial success in problem-solving (Figure 1b), and resets (Figure 1c) in *From Here to There! (FH2T)*. FH2T is an interactive game where mathematical terms are turned into digital objects that can be manipulated, and students transform starting expressions into specified goals (Figure 2). The data in the current study came from a larger Randomized Control Trial (RCT) testing the efficacy of FH2T (Chan, et al., 2021). Previously, we found that students who received FH2T showed improved mathematical understanding compared to students who received online problem sets with hints and immediate feedback. Going beyond the intervention effects, we focused on students in FH2T to examine their problem-solving processes and potential mechanisms of learning. Here, we used students' mouse- and keyboard-action data within FH2T to investigate the influence of pause time on strategy efficiency. Strategy efficiency, rather than algebraic knowledge, was our focal outcome because we aimed to examine whether and how students' pause time relates to their problem-solving. We asked whether students' pause time, mathematical knowledge, mathematics self-efficacy, and mathematics anxiety predict (1) the number of steps they take to solve a problem, (2) their completion rate of problem-solving on the initial attempt, and (3) the number of times they reset and reattempt problems prior to completing the problem?

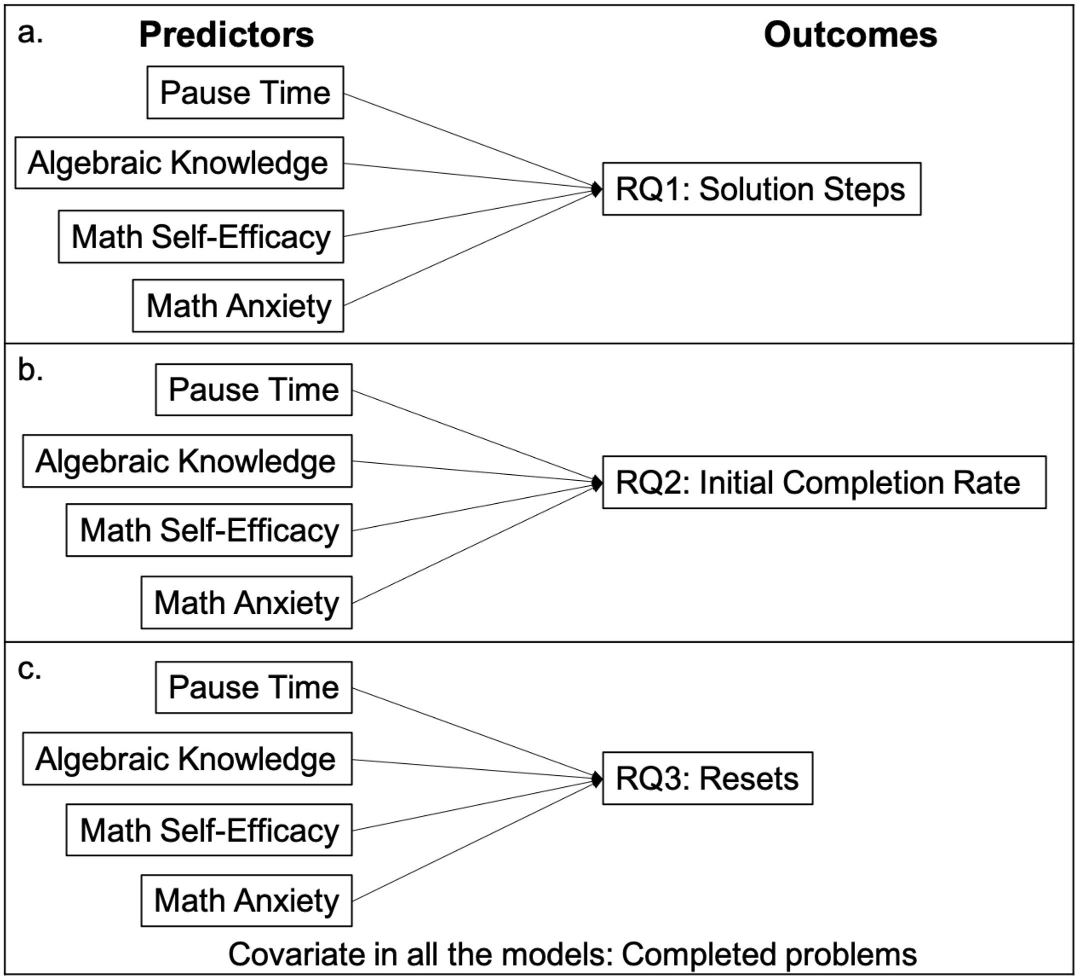

Figure 1. Conceptual models illustrating the research questions that test the effects of pause time, algebraic knowledge, mathematics self-efficacy, and mathematics anxiety on students' solution steps (a; RQ1), initial completion rate (b; RQ2), and reset frequency (c; RQ3)

By addressing these questions, we extend prior research in two ways. First, we simultaneously compare the relative influences of students' metacognitive (i.e., pause time), cognitive (i.e., mathematical knowledge), and affective (i.e., mathematics self-efficacy and mathematics anxiety) factors on their strategy efficiency. Second, we extend beyond solution steps and explore how student factors predict other aspects of strategy efficiency—specifically, students' initial success in problem-solving and the number of resets prior to completing a problem.

2 Methods

2.1 Participants

The sample was drawn from a RCT conducted in Fall 2019, which aimed to improve students' algebra learning through educational technologies (Chan et al., 2021). The students were recruited from 29 classrooms across six middle-schools in a large, suburban district in the Southeastern United States. In the RCT, 689 students were randomly assigned within classrooms at the student level to complete four 30-minute intervention sessions using FH2T (n = 348) or online problem sets (n = 341). Here, we focused on the students who used FH2T because FH2T recorded all student actions during problem-solving, providing detailed information on students' pause time and strategy efficiency. The online problem sets presented students with textbook problems (e.g., $3(2+x) = 18$) and students entered the answer in a textbox ($x = 4$); therefore, they did not provide the data for our research questions.

Among the 348 students, we removed 45 students who did not complete the pretest, and 18 students who made low progress in FH2T due to class scheduling constraints. Because the 45 students did not complete the pretest, we could not compare their cognitive and affective characteristics to the final analytic sample. The 18 students excluded due to low progress did complete the pretest; their algebraic knowledge (M = 0.37) and mathematics self-efficacy (M = 2.63) were descriptively lower than those of the final sample (algebra knowledge: M = 0.57; mathematics self-efficacy: M = 3.71), whereas their mathematics anxiety (M = 1.70) was descriptively higher than those of the final sample (M = 1.22). We discuss these differences in the Limitation and Future Directions section.

The analytic sample included 285 students (11- to 13-year-olds); the majority were in sixth grade (97%), and the remaining in seventh grade (3%). In terms of the mathematics instruction level, 86% were in advanced classes, 9% in on-level classes, and 5% in support classes. All participating seventh-graders were in support classes and learning sixth-grade content, suggesting that their mathematical knowledge might be comparable to that of sixth-graders. Therefore, the seventh-graders were included in the analyses. Within the analytic sample, we received demographic information on 278 students from the school district. Of the 278 students, 45% were female and 55% were male; 51% were above grade level in academic achievement. In terms of race/ethnicity, 52% were Asian, 38% were White, 5% were Hispanic, 2% were Black, 1% were Native American, and 2% were Multi-racial. The sample affords 95% power to detect a small to medium effect of $f^2 \geq .04$ at $p = .05$. The conventional cut-offs for small, medium, and large effects are .02, .15, and .35, respectively (Cohen, 1992).

This research was approved by the Institutional Review Board at a University in the Northeastern United States. This research involved typical educational practices and did not require parental consent. Parents received a letter about the research and were offered the opportunity to opt-out. No parents opted their child out of this study.

2.2 Procedure

The intervention study consisted of a 45-minute pretest, four 30-minute intervention sessions, and a 40-minute posttest within six weeks. All study assignments were administered online during mathematics instructional periods, and students worked individually using a device. The study instruction for all sessions was embedded in the assignments, therefore, teachers only needed to provide students the link and allocate time for students to complete the assignments. All students followed the same study protocol presented to them on their browsers. In FH2T, students could request hints to solve each problem; therefore, they could progress through the game without being stuck.

We used the mouse- and keyboard-action data recorded within FH2T and the pretest measures of algebraic knowledge, mathematics self-efficacy, and mathematics anxiety to address our research questions. The procedure of the RCT and results on the intervention effects are reported in Chan et al. (2021); thus, we only describe tasks relevant to the current study.

2.3 From Here to There! (FH2T)

In each problem of FH2T, students saw a starting expression and a mathematically equivalent goal. The objective was to transform the starting expression into the specified goal using gesture-actions, tapping or dragging that transforms expressions from one state to another. Students were encouraged to use efficient strategies by taking the fewest steps possible. They were not instructed to solve problems quickly or to take their time, providing an opportunity to examine individual differences in pause time and its relation with strategy efficiency.

Figure 2 illustrates a FH2T problem with a series of steps—gesture-actions that lead to valid transformation—to reach the goal. In this example, students were to transform the starting expression, 7+2+10+8, to match the goal, 5+2+5+15 (Figure 2a). The student first *added 7 to 8* by dragging 7 on top of 8 (Figure 2b). Next, they *decomposed 10 into 5 and 5* by using the keypad (Figure 2c). After, they *commuted 5* by dragging it to the left of the expression (Figure 2d). In this example, the student took three steps to reach the goal (Figure 2e): 1) dragging to add (2+10+15), 2) keypad to decompose (2+5+5+15), and 3) dragging to commute (5+2+5+15). They received a reward of three clovers for completing the problem using the minimum number of steps required to reach the goal (Figure 2f). On all problems, students could reset the expression and retry the problem multiple times; however, they were encouraged to use efficient strategies. Specifically, students could earn more clovers (i.e., a maximum of three) when they use more efficient strategies; further the step count (in the bottom right next to the goal; Figure 2) turns red when students' steps exceed the minimum number required for the problem.

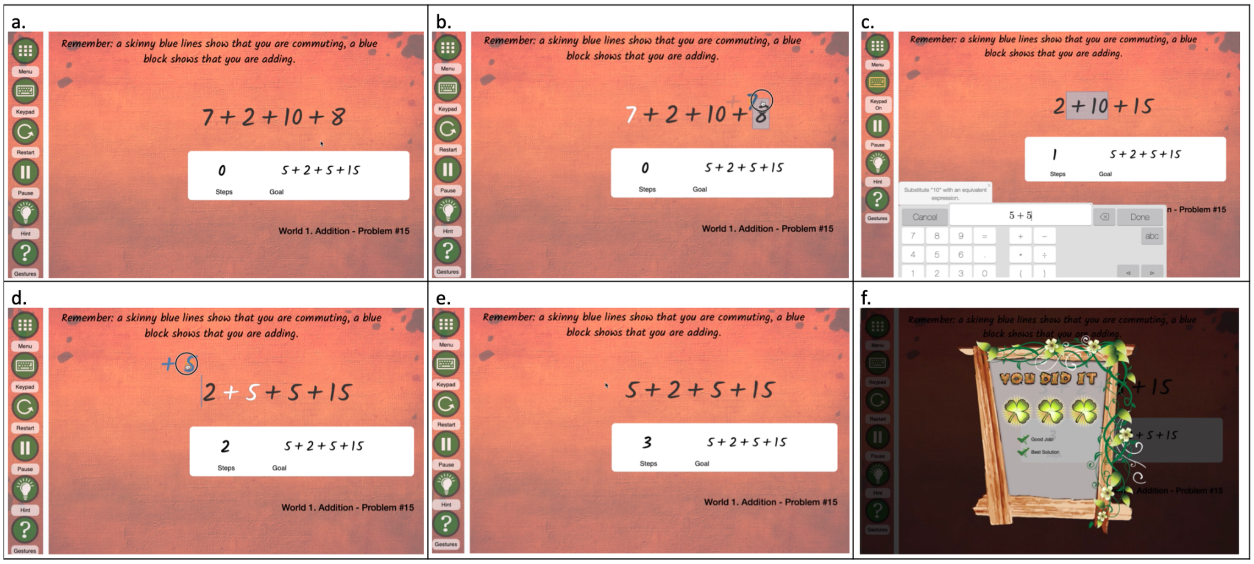

Figure 2. A sample problem in From Here to There! (a) and a potential transformation process involving three steps (b, c, d) to reach the goal state (e, f).

Because all student actions and the corresponding transformations of expressions were timestamped and recorded, we could systematically and quantitatively examine students' problem-solving processes in ways not accessible in answer-based learning systems or paper-and-pencil tasks. Furthermore, the problems in FH2T do not merely ask students to transform expressions into the simplest form. Instead, noticing the structures of the starting expression and the goal may help students efficiently reach the goal. Students could take any series of mathematically valid steps that link the starting expression and the goal (Figure 3). FH2T thus provided an ideal context for examining variation in strategy and process during algebraic problem-solving.

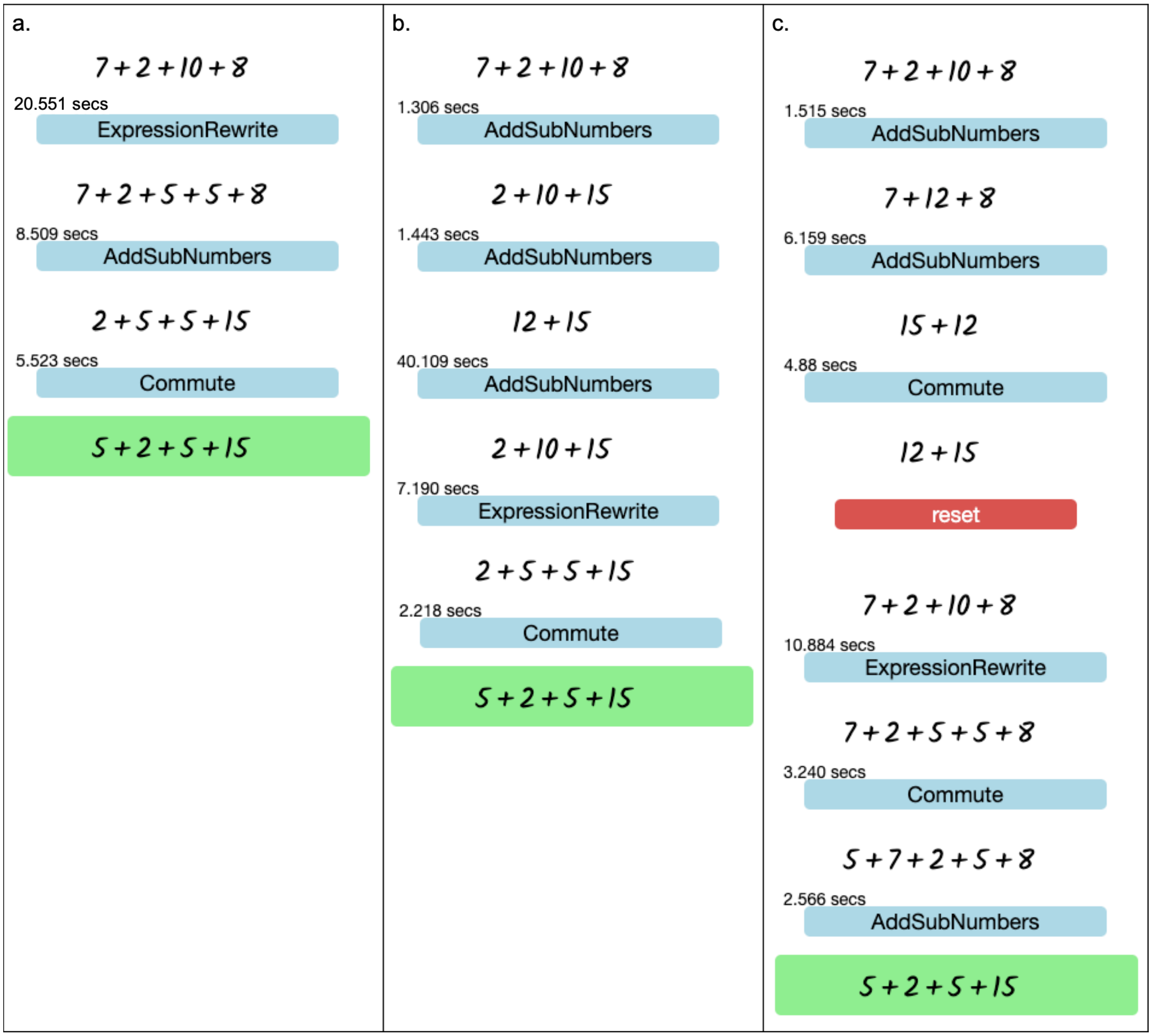

Figure 3. An illustration of three different series of steps between a starting expression (7+2+10+8) and its goal (5+2+5+15). Each column shows a chain of mathematically equivalent states produced serially by each gesture-action. (a) An efficient equation-transformation strategy with three solution steps. (b) An inefficient equation-transformation strategy with five solution steps. (c) A strategy with a reset and six solution steps.

FH2T was developed based on the Common Core State Standards for middle-school mathematics (2010) to improve students' algebraic understanding. Topics, such as arithmetic operations, fractions, factoring, and distribution, were organized into a series of worlds, each containing 18 problems. Students started from simple topics and built up their knowledge and skills throughout the game. Prior work has demonstrated positive effects of playing FH2T on elementary students' mathematical performance (Hulse et al., 2019) as well as middle-schoolers' equation-solving performance (Ottmar et al., 2015) and understanding of equivalence (Chan et al., 2021).

In this study, all students started from addition, and worked their way through the arithmetic operations, and inverse operations. To help students learn the gesture-actions in FH2T, we interspersed training problems to provide just-in-time instruction on the gesture-actions for arithmetic operations and mathematical properties. These training problems included videos demonstrating the gesture-actions and required only one or two steps to complete. In order to accurately measure students' problem-solving strategies, our analyses focused on the non-training problems—the problems that did not include the demonstration videos and required three or more steps to complete. Students worked through FH2T at their own pace, and they completed an average of 35 non-training problems (*SD* = 18) across the four intervention sessions.

To compare the strategy efficiency across students, we focused our analyses on the first 20 non-training problems that were completed by 75% of the analytic sample (Table 1). The remaining 25% of the students completed at least 10 of these problems, providing some data to estimate the association between pause time and strategy efficiency. Instead of excluding 25% of the sample, we included all students in the analyses and used the number of non-training problems students completed as a covariate to account for the variation in students' FH2T progress.

Table 1.
The starting expression and the goal of the 20 problems included in the analyses, and minimum number of steps required to transform the starting expression into the goal.

Problem	Starting Expression	Goal	Min. Steps
1	$0 + 1 + 2 + 3$	$3 + 2 + 1 + 0$	3
2	$7 + 6 + b + 10$	$12 + 11 + b$	3
3	$3 + 3 + 3 + 3$	$4 + 4 + 4$	5
4	$9 \cdot 4$	$3 \cdot 6 \cdot 2$	3
5	$6 \cdot 10$	$2 \cdot 15 \cdot 2$	3
6	$3 \cdot 7 + 7 + 3$	31	3
7	$8 + 4 \cdot 4 \cdot 8$	136	3
8	$2 \cdot 3 + 3 \cdot 2$	12	3
9	$5 + 2 + 2 \cdot 5$	17	3
10	$4 + 1 \cdot 1 + 4$	9	3
11	$1 \cdot 6 \cdot 6 + 1$	37	3
12	$8 + 2 \ + 3 \cdot 5c$	$15c + 10$	3
13	$2 \cdot a + 2 \cdot b$	$a + b + a + b$	3
14	120	$5 \cdot 12+12 \cdot 5$	3
15	$-14 + (-5) - (-18)$	$-5 + 4$	4
16	$6 - 2 + 7$	$11 \cdot 1$	3
17	$4 + 6 - 3 - 7 + 0$	$-10 + 0 + 10$	3
18	$-14 + 16$	$-7 + 1 + 8$	3
19	$15 - 25 + 8 + 2$	$10 - 10$	3
20	$27 + 65 + 3 - 25 + 75 - 65$	$5 + 0 + 75$	3

Note: Min. Steps = the minimum number of steps required to transform the starting expression into the goal.

2.4 Measures

2.4.1 Focal Measures in From Here to There!

FH2T logged all student actions in the game. The following four variables were recorded for FH2T problems. For each variable, we averaged the values across the completed problems to obtain the student-level data for analyses.

2.4.1.1 Pause Time (in Percent)

We first computed the number of seconds students spent *before* taking their first step on each problem. This was the amount of time from when the problem first appeared on the screen to when students made their first expression transformation. Because the pause time varied across students and problems depending on problem complexities and we were interested in the amount of pause time relative to the total time on the problems, we followed Li and colleagues' (2015) procedure to compute the percent pause time (i.e., pause time ÷ total time). As an example, the percent pause time in Figure 3a is 59.4% (20.551sec ÷ 34.583 sec). We used the percent pause time, hereafter pause time, across completed problems as a predictor in our analyses.

2.4.1.2 Solution Steps

The solution steps comprised the sum of all the steps students took to reach the goal on a problem. As an example, the students in Figures 3a, 3b, and 3c took three, five, and six steps, respectively. We used the average number of solution steps across the completed problems as the dependent variable for RQ1.

2.4.1.3 Initial Completion Rate

Students could transform and reset their expressions using the "restart" button as many times as they wanted prior to completing the problem. On each problem, we recorded whether students successfully transformed the starting expression into the goal on their initial attempt (i.e., without resetting the problem). As an example, the students in Figure 3a and 3b completed the problem on their initial attempt, whereas the student in Figure 3c reset and did not complete the problem on their initial attempt. We computed the proportion of problems students completed on their initial attempt and used it as the dependent variable for RQ2.

2.4.1.4 Reset

Because students can reset the expressions as many times as needed during problem-solving, we also computed the number of times students reset the expression on each problem. For example, the students in Figure 3a and 3b did not reset the expression, whereas the student in Figure 3c reset the expression once. We used the average number of resets across completed problems as the dependent variable for RQ3.

2.4.2 Pretest Measures

Three pretest measures were included as predictors of strategy efficiency in FH2T.

2.4.2.1 Algebraic Knowledge

Students' algebraic knowledge was assessed with 11 items adapted from two validated measures (Rittle-Johnson, et al., 2011; Star et al., 2014). All problems were directly taken from the two measures. We edited the instructional texts of four items to improve the clarity of the instruction within the online platform. For example, the original instruction for an item (e.g., $8+\_\_\_=8+6+4$) from Rittle-Johnson et al. (2011) was, "Find the number that goes in each box." We edited the text to "Enter the number that goes in the box" so students know to input the answer into the textbox using their keyboard. The assessment measured various aspects of students' algebraic knowledge, including balancing equations (e.g., $898+13 = 896 +\_\_\_$.), solving for a variable (e.g., $5(y-2) = 3(y-2)+8$), and evaluating equation-solving strategies (e.g., *Which would be the best way to start the problem 3(x+2) = 14? A. distribute first, B. divide by 3 on both sides first, C. multiply by 3 on both sides first, or D. subtract 14 from both sides first*). Each item was scored as correct (1) or incorrect (0), and the reliability of the items was fair, $KR\text{-}20 = .68$. The average score on the assessment was included as a predictor in the analyses.

2.4.2.2 Mathematics Self-efficacy

Students' mathematics self-efficacy was measured with five items adapted from the Academic Efficacy subscale of the Patterns of Adaptive Learning Scales (Midgley et al., 2000). The original items were designed to measure students' general academic efficacy (e.g., "I can do even the hardest work in this class if I try"). We adapted the items to focus on mathematics (e.g., "I can do even the hardest work in my math class if I try"). Students rated how often they felt like the way described in the statements from *never* (0), *very rarely* (1), *rarely* (2), *often* (3), *very often* (4), to *always* (5). The reliability of the items was $\alpha = .86$, comparable to that reported by Midgley and colleagues ($\alpha = .82$). The average rating was included as a predictor in the analyses.

2.4.2.3 Mathematics Anxiety

Students' mathematics anxiety was measured with 13 items from the Math Anxiety Scale for Young Children–Revised (Ganley & McGraw, 2016). The items were designed to measure student perceptions of their own anxieties towards mathematics. A sample item was "I get worried before I take a math test." Students rated how much they felt like the way described in the statements from *No* (0), *Not really* (1), *Kind of* (2), to *Yes* (3). The reliability of the items was $\alpha = .88$, comparable to that reported in Ganley and McGraw ($\alpha = .87$). The average rating was included as a predictor in the analyses.

2.5 Analysis Plan

Prior to addressing our research questions, we conducted descriptive and correlation analyses to examine the distribution of, and relations between, each variable at the student-level. To examine the effects of pause time, algebraic knowledge, mathematics self-efficacy, and mathematics anxiety on students' solution steps (RQ1), we conducted an OLS regression model with pause time, algebraic knowledge, mathematics self-efficacy, and mathematics anxiety as the predictors, solution steps as the dependent variable, and completed problems as the covariate. To further investigate these effects on students' strategy efficiency, we repeated the regression model with students' initial completion rate (RQ2) and reset frequency (RQ3) as the dependent variable in each model, respectively (Figure 1). All analyses were conducted with the *lme4* package (Bates et al., 2015) in R (R Core Team, 2020).

3 Results

3.1 Descriptive and Correlation Analysis

The descriptive analysis revealed that the values for all measures were widely distributed, indicating a wide range of performance levels in the sample (Table 2). The skewness and Kurtosis for each variable were well within the normality cutoffs, i.e., ±2 for skewness and ±7 for Kurtosis (Byrne, 2010). On average, students paused for 26% (SD = 10%) of their total problem-solving time prior to taking their first steps; this is an average of 16 seconds (SD = 12) of pause time within 63 seconds (SD = 35) of total problem-solving time. The correlation analysis demonstrated that students' algebraic knowledge was significantly correlated with all three aspects of their strategy efficiency—solution steps ($p < .001$), initial completion rate ($p < .001$), and reset ($p < .001$). These correlations were somewhat weaker but still significant after partialling out the number of FH2T problems students completed, ps < .001. Similarly, students' mathematics anxiety was weakly correlated with solution steps ($p = .046$) and initial completion rate ($p = .030$), but these correlations were not significant when controlling for completed problems. Pause time was moderately correlated with the three aspects of strategy efficiency (ps < .001), and these correlations remained after controlling the number of completed problems (ps < .001).

Table 2.
Descriptive statistics and correlations between measures at the student level (N = 285)

Variable	1	2	3	4	5	6	7	8
1. Completed Problems	-	-	-	-	-	-	-	-
2. Pause Time (Percent)	$.13^{*}$	-	$-.53^{***}$	$.31^{***}$	$-.45^{***}$	$.34^{***}$	.06	$-.16^{**}$
3. Solution Steps	-.11	$-.54^{***}$	-	$-.45^{***}$	$.75^{***}$	$-.26^{***}$	-.03	.10
4. Initial Completion Rate	$.67^{***}$	$.32^{***}$	$-.41^{***}$	-	$-.67^{***}$	$.23^{***}$	.02	-.05
5. Resets	$-.15^{*}$	$-.46^{***}$	$.76^{***}$	$-.59^{***}$	-	$-.28^{***}$	-.08	.06
6. Algebraic Knowledge	$.28^{***}$	$.36^{***}$	$-.28^{***}$	$.35^{***}$	$-.31^{***}$	-	$.27^{***}$	$-.36^{***}$
7. Math Self-Efficacy	$.12^{*}$	.08	-.04	.09	-.10	$.29^{***}$	-	$-.53^{***}$
8. Math Anxiety	$-.14^{*}$	$-.17^{**}$	$.12^{*}$	$-.13^{*}$	.08	$-.38^{***}$	$-.54^{***}$	-
Mean	18.73	26%	5.41	0.70	0.60	0.57	3.71	1.22
Standard Deviation	2.41	10%	1.46	0.14	0.41	0.22	0.87	0.61
Minimum	10.00	5%	3.00	0.30	0.00	0.09	1.00	0.00
Maximum	20.00	74%	11.83	1.00	3.00	1.00	5.00	3.00
Skewness	-1.63	0.89	1.68	-0.33	1.82	0.07	-0.41	0.39
Kurtosis	1.34	2.02	3.69	-0.17	5.78	-0.95	-0.41	-0.52

Note: The values in the lower triangle represent the zero-order correlations. The values in the upper triangle represent the partial correlations controlling for completed problems. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table 3.

Regression models predicting students' solution steps, initial completion rate, and reset in FH2T.

	Solution Steps		Initial Completion Rate		Resets	
Predictors	β (SE)	95%CI	β (SE)	95%CI	β (SE)	95%CI
Completed Problems	-.02 (.05)	[-.13, .08]	.62 (.04)***	[.53, .70]	-.06 (.05)	[-.16, .05]
Pause Time	-.50 (.05)***	[-.61, -.39]	.20 (.05)***	[.11, .28]	-.40 (.06)***	[-.51, -.29]
Algebraic Knowledge	-.10 (.06)	[-.22, .02]	.13 (.05)*	[.03, .22]	-.17 (.06)**	[-.29, -.05]
Math Self-Efficacy	.03 (.06)	[-.09, .15]	-.02 (.05)	[-.12, .08]	-.06 (.06)	[-.18, .07]
Math Anxiety	.01 (.06)	[-.11, .13]	.03 (.05)	[-.07, .13]	-.09 (.06)	[-.21, .04]
R^2	.299		.515		.245	

* $p < .05$; ** $p < .01$; *** $p < .001$

3.2 Longer Pause Time was Associated with Fewer Solution Steps

An OLS regression revealed that students with longer pause time completed problems with more efficient strategies that involved fewer solution steps, $p < .001$ (Table 3). In particular, a one standard deviation increase in pause time was associated with 0.50 standard deviation decrease in solution steps, 95%CI: [-0.61, -0.39]. Students' algebraic knowledge ($p = .092$), mathematics self-efficacy ($p = .586$), and mathematics anxiety ($p = .865$) were not significantly associated with their solution steps. The model accounted for 29.9% of the variance in students' solution steps.

We checked the model assumptions by examining plots for linearity of the residuals versus fitted values, normality of residuals, homogeneity of the residual variance, and potential influential cases. A visual inspection of the plots suggested that the first three assumptions might not be satisfied; therefore, we log-transformed solution steps and repeated the analyses. Because the pattern of results remained after the transformation (i.e., only pause time significantly predicted solution steps), we reported the non-transformed results above to aid the interpretation.

3.3 Longer Pause Time was Associated with Higher Initial Completion Rate

An OLS regression revealed that students with longer pause time had a higher completion rate on their initial attempt, $p < .001$ (Table 3). A one standard deviation increase in pause time was associated with 0.20 standard deviation increase in initial completion rate, 95%CI: [0.11, 0.28]. Students' algebraic knowledge was positively associated with their initial completion rate ($\beta = 0.13$, 95%CI: [0.03, 0.22], $p = .011$), whereas students' mathematics self-efficacy ($p = .740$) and mathematics anxiety ($p = .606$) were not. The model accounted for 51.5% of the variance in students' initial completion rate. An examination of the model assumptions suggested that the model met the assumptions.

3.4 Longer Pause Time was Associated with Fewer Resets

An OLS regression revealed that students with longer pause time reset problems less frequently, $p < .001$ (Table 3). In particular, a one standard deviation increase in pause time was associated with 0.40 standard deviation decrease in resets, 95%CI: [-0.51, -0.29]. Students' algebraic knowledge was negatively associated with students' reset frequency ($\beta = -0.17$, 95%CI: [-0.05, -0.29], $p = .006$), whereas students' mathematics self-efficacy ($p = .362$) and mathematics anxiety ($p = .176$) were not. The model accounted for 24.5% of the variance in students' reset frequency. An examination of the model assumptions suggested that the normality of residuals and homogeneity of residuals variance might not be satisfied; thus we log-transformed students' reset frequency and repeated the analyses. Because the pattern of results remained after the transformation (i.e., both pause time and algebraic knowledge significantly predicted students' reset frequency), we reported the non-transformed results above to aid the interpretation.

4 Discussion

We set out to investigate the relations between students' pause time as a proxy indicator of thinking before problem-solving and strategy efficiency in an algebra learning system. We found that longer pause time was associated with more efficient strategy use as indicated by fewer solution steps, higher initial completion rate, and fewer resets. Students' algebraic knowledge was only weakly associated with higher initial completion rate and fewer resets; students' mathematics self-efficacy, and mathematics anxiety were not significantly associated with their strategy efficiency above and beyond pause time and algebraic knowledge. The findings suggest that the relation between pause time and strategy efficiency may be robust and independent of factors previously identified in the literature.

4.1 Pause Time is Associated with Strategy Efficiency in Problem-Solving

The current findings add to the literature by demonstrating that pause time is associated with strategy efficiency above and beyond knowledge and affective factors. In prior work, researchers have used pause time as a proxy indicator of thinking and planning (Gobert et al., 2015; Paquette et al., 2014), and it significantly predicted strategy efficiency in problem-solving (Li et al., 2015). However, it remained unclear whether pause time was a *unique* and *independent* predictor of strategy efficiency or whether this association was influenced by students' knowledge or affective dispositions. By testing the effects of pause time, algebraic knowledge, mathematics self-efficacy, and mathematics anxiety simultaneously, we found that students' pause time was the only consistent predictor of strategy efficiency. Specifically, longer pause time was associated with fewer solution steps, higher initial completion rate, and fewer resets. The standardized beta coefficients and confidence intervals further revealed that these effects of pause time on strategy efficiency were moderate (Table 3), suggesting its potential importance in students' problem-solving processes.

Extending previous studies (e.g., Hoffman, 2010; Newton et al., 2020; Passolunghi et al., 2016; Ramirez et al., 2016), we found that students' algebraic knowledge was only weakly associated with their initial completion rate and reset frequency, whereas students' mathematics self-efficacy, and mathematics anxiety were not associated with the three measures of strategy efficiency in FH2T. Students' algebraic knowledge was moderately correlated with the three aspects of strategy efficiency, and mathematics anxiety was weakly correlated with solution steps and initial completion rate (Table 2); however, when all the predictors were included in the models, these relations were attenuated. One potential explanation is that students' knowledge and affective factors are moderately correlated with each other and they may have some shared variance in students' strategy efficiency, leading to the non-significant results when they are tested simultaneously in one model. Further, these factors may potentially influence students' strategy efficiency through pause time as students with higher algebraic knowledge or lower mathematics anxiety tend to pause longer prior to solving. The current data do not allow us to explore the potential pathways through which students' knowledge, affects, and pause time influence strategy efficiency; however, they do demonstrate the importance of considering these factors simultaneously in one study.

Overall, our results extend prior research in two important ways. First, above and beyond students' knowledge and affective dispositions, pause time is significantly associated with strategy efficiency in mathematics problem-solving. The relation between pause time and strategy efficiency may be partly driven by the fact that both variables were measured in the same task context—FH2T. However, context alone cannot fully explain this relation as we also included the total number of completed FH2T problems in the analyses yet pause time remained a significant predictor of strategy efficiency. Second, even after accounting for the potential effects of student knowledge and affective dispositions, pause time is still moderately associated with aspects of students' strategy efficiency, indicating its potential significance in students' mathematics problem-solving. Together, the findings suggest that the relation between pause time and strategy efficiency may be meaningful and has implications for research and practices.

4.2 Plausible Mechanisms for Pause Time and Strategy Efficiency

While the correlational findings do not allow causal inferences, they do suggest a link between longer pause times and more efficient strategies. Several plausible cognitive, metacognitive, and affective mechanisms could be at play (Cleary & Chen, 2009; Dinsmore et al., 2008; Montague et al., 2011). For instance, pause time may provide a window of opportunity for students to carefully read, understand, and identify the type of problem, and together these behaviors may contribute to devising more viable strategies, and consequently increase students' strategy efficiency. This hypothesis aligns with the findings that interventions on metacognitive skills improve students' mathematics performance (Perels et al., 2005; Vula et al., 2017), providing students with opportunities to process and reflect on the problems improves their conceptual knowledge (González-Cabañes, et al., 2020), and adults report planning during pause time which leads to higher efficiency in problem-solving (Welsh et al., 1995). Our analyses also align with other researchers' approach of using pause time as a proxy indicator of thinking and planning in online learning environments (Gobert et al., 2015; Li et al., 2015; Paquette et al., 2014).

Under this hypothesis, pause time may be a proxy of students' metacognitive skills, and the variation in pause time may indicate individual differences in students' thinking of their strategies, behaviors, and learning processes. Solving problems with efficient strategies requires students to monitor their knowledge and regulate their behavior (Caviola et al., 2017), and previous studies have reported consistent correlations between metacognitive skills and mathematics performance (e.g., García et al., 2019; Losenno et al., 2020). Further, prior research with a diverse sample of participants between the age of 10 and 30 indicated that participants' age positively predicted their performance on the Tower of London (a variation of the Tower of Hanoi task), and the age-related gains were associated with maturational improvement on participants ability to control their impulses (Albert & Steinberg, 2011). If pause time reflects metacognitive skills, it may serve as a potential behavioral mechanism through which these skills influence mathematics problem-solving.

Alternatively, pause time may not be an indicator of students' metacognitive skills, but instead, it may reflect the influences of cognitive (e.g., knowledge, distractedness) and/or affective (e.g., self-efficacy, anxiety) factors. For instance, students with higher algebraic knowledge may take their first step quickly, whereas students with higher mathematics anxiety may freeze up when they see the problem. Although possible, the data provided little support for these accounts. Specifically, students with higher algebraic knowledge or lower math anxiety tended to have longer pause time, and these correlations were weak to moderate (Table 2). These findings suggest that pause time may reflect skills or constructs other than those measured in the current study. Further, students' average pause time was 16 seconds, which was longer than that of guessing or careless responding (Kong et al., 2007; Meade & Craig, 2012; Wise et al., 2009) yet shorter than that if students left their seat for breaks. Although the results suggest that students may be using pause time to understand the problem and plan their solutions instead of being distracted, we cannot rule out the possibility that students' minds wandered briefly during pause time. However, if students were daydreaming or mind-wandering during pause time, we would potentially see a negative or no relation between pause time and strategy efficiency instead of a positive relation as reported in our study.

The current study has implications for research and practices in mathematics teaching and learning. Our findings suggest that pause time may support uses of efficient strategies, and contribute to the efforts in delineating the relation between metacognitive skills and mathematics problem-solving. Future studies can incorporate assessments of metacognitive skills to directly measure their association with pause time. Providing explicit instructions that minimize mind-wandering, video-recording students' behaviors during the study sessions, and including talk-aloud protocols that ask students to explain their problem-solving processes may reveal the extent to which students plan for, think about, and engage with the problems during pause time. Further, examining students' expression transformation processes more deeply by coding the types of steps taken and the mathematical properties applied may reveal potential pathways between pause time and strategy efficiency.

4.3 Limitation and Future Directions

This study had several limitations. First, the analyses focused on a subset of students within a larger RCT and a subset of FH2T problems that required three or more steps. Although only 20 problems were included in the analyses, they covered a range of difficulties and were designed to prompt variations in students' strategy efficiency. Further, due to the specific teachers who responded to our recruitment and dedicated instructional periods for the study, the sample comprised a higher number of Asian students and students in advanced math classes. In fact, students who made little progress in FH2T and thus were excluded from the analyses have descriptively lower algebraic knowledge and mathematics self-efficacy, and higher mathematics anxiety compared to the analytic sample. Therefore, the final analytic sample might not be representative of other students in the district or the populations in the State or the U.S. Still, among our relatively high-performing and homogenous sample, we observed large variations in students' behaviors in FH2T, algebraic knowledge, mathematics self-efficacy, and mathematics anxiety, and found significant associations between these measures. Future work should replicate these findings in other learning platforms, and with more problems, additional covariates (e.g., processing speed, general reasoning ability), and more diverse samples.

Next, since the RCT was conducted in the classroom, the fidelity of implementation of the students relied on teachers allocating classroom time for the study. However, because the study took place between October and December of 2019 (right before the holiday break), several teachers did not provide adequate class time for students to complete all the sessions. This resulted in a portion of students who did not engage with the game for very much time, contributing to variation in student progress within FH2T. To account for this variation, we included the number of completed problems as a covariate in all analyses. Although the sample size was constrained by the fidelity of implementation, the analytic sample comprised 82% of the larger sample and had adequate power to detect effects. Further, the study was conducted in a more ecologically valid context, providing some evidence for the association between pause time and strategy efficiency in a classroom setting.

The results presented here, although significant, may not generalize to problem-solving in other contexts and do not provide insights into *how* longer pause time related to higher strategy efficiency. Future directions include replicating these findings with other platforms, and implementing other research methods, such as observations and talk-aloud protocols, to delineate plausible mechanisms underlying this association. The log data within FH2T also provides opportunities to explore the relations between students' pause time and other aspects of their problem-solving behaviors. For example, one potential mechanism through which pause time impacts strategy efficiency is through students' first step of problem-solving. Specifically, longer pause time may be associated with a more productive first step that leads students closer to the goal, and a productive first step may set students up for using a more efficient strategy.

Finally, another future direction is to experimentally test the effects of enforcing pause time by prompting students to plan out their strategy and think about their solutions before problem-solving. If the experimental findings align with the current results, it may be beneficial to encourage students to pause and think before solving problems. Teachers may also consider encouraging students to take time planning out their strategy to improve strategy efficiency.

4.4 Conclusion

An identified challenge in mathematical cognition research is how to adequately measure students' thinking process and behaviors during problem-solving (García, Betts, et al., 2016). As Wilson et al. (1993) wrote, "Technology can be used to enhance or make possible exploration of conceptual or problem situations." Technologies like FH2T allow students to explore mathematics problem-solving and researchers to examine students' strategy efficiency. Implications for research and practice are to move beyond the correctness of student answers, and leverage data within technologies to examine and promote the flexibility and efficiency in students' problem-solving. The current study is our step towards investigating students' problem-solving process beyond correctness.

This study contributes to research on the individual differences of learning dispositions and behavioral processes as well as how these differences influence mathematics problem-solving strategies. If pause time predicts strategy efficiency, it warrants further investigation to identify the underlying mechanisms of this relation. Educators and developers of instructional materials would benefit from a deeper understanding of the effect of pause time on students' thinking, learning, and performance. The findings from the current study suggest that longer pause time before solving problems is related to the use of more efficient strategies, providing some validity to the paradox that sometimes it is better to "slow down to speed up."

References

Albert, D., & Steinberg, L. (2011). Age differences in strategic planning as indexed by the Tower of London. *Child Development*, *82*(5), 1501–1517. https://doi.org/10.1111/j.1467-8624.2011.01613.x

Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science*, *11*(5), 181–185. https://doi.org/10.1111/1467-8721.00196

Ashkenazi, S., & Najjar, D. (2018). Non-adaptive strategy selection in adults with high mathematical anxiety. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-27763-w

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J. (2015). Package "lme4." *Convergence*, *12*(1), 2.

Bernacki, M. L., Nokes-Malach, T. J., & Aleven, V. (2015). Examining self-efficacy during learning: variability and relations to behavior, performance, and learning. *Metacognition and Learning*, *10*(1), 99–117. https://doi.org/10.1007/s11409-014-9127-x

Byrne, B. M. (2010). Structural equation modeling with AMOS: Basic concepts, applications, and programming. New York: Routledge.

Carpenter, T. P., Kepner, H., Corbitt, M. K., Montgomery, M., & Reys, R. E. (1980). Results and implications of the second NAEP mathematics assessments: Elementary school. *The Arithmetic Teacher*, *27*(8), 10–12, 44–47.

Caviola, S., Carey, E., Mammarella, I. C., & Szucs, D. (2017). Stress, time pressure, strategy selection and math anxiety in mathematics: A review of the literature. *Frontiers in Psychology*, *8*(SEP), 1–13. https://doi.org/10.3389/fpsyg.2017.01488

Chan, J. Y.-C., Lee, J.-E., Mason, C. A., Sawrey, K., & Ottmar, E. (2021). From Here to There! A dynamic algebraic notation system improves understanding of equivalence in middle-school students. *Journal of Educational Psychology*. https://doi.org/http://dx.doi.org/10.1037/edu0000596

Cleary, T. J., & Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology*, *47*(5), 291–314. https://doi.org/10.1016/j.jsp.2009.04.002

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155.

Common Core State Standards (2010). Common Core State Standards for Mathematics (CCSSM). *Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers*.

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review*, *20*(4), 391–409. https://doi.org/10.1007/s10648-008-9083-6

Dreger, R. M., & Aiken, L. R., Jr. (1957). The identification of number anxiety in a college population. *Journal of Educational Psychology*, *48*(6), 344–351. https://doi.org/10.1037/h0045894

Fast, L. A., Lewis, J. L., Bryant, M. J., Bocian, K. A., Cardullo, R. A., Rettig, M., & Hammond, K. A. (2010). Does math self-efficacy mediate the effect of the perceived classroom environment on standardized math test performance? *Journal of Educational Psychology*, *102*(3), 729–740. https://doi.org/10.1037/a0018863

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The Nature of Intelligence* (pp. 231–236). Hilledale, NJ: Earlbaum.

Ganley, C. M., & McGraw, A. L. (2016). The development and validation of a revised version of the math anxiety scale for young children. *Frontiers in Psychology*, *7*(AUG). https://doi.org/10.3389/fpsyg.2016.01181

García, T., Betts, L., González-Castro, P., González-Pienda, J. A., & Rodríguez, C. (2016). On-line assessment of the process involved in maths problem-solving in fifth and sixth grade students: self regulation and achievement. *Revista Latinoamericana de Investigacion En Matematica Educativa*, *19*(2), 165–186. https://doi.org/10.12802/relime.13.1922

García, T., Boom, J., Kroesbergen, E. H., Núñez, J. C., & Rodríguez, C. (2019). Planning, execution, and revision in mathematics problem solving: Does the order of the phases matter? *Studies in Educational Evaluation*, *61*(March), 83–93. https://doi.org/10.1016/j.stueduc.2019.03.001

García, T., Rodríguez, C., Betts, L., Areces, D., & González-Castro, P. (2016). How affective-motivational variables and approaches to learning predict mathematics achievement in upper elementary levels. *Learning and Individual Differences*, *49*, 25–31. https://doi.org/10.1016/j.lindif.2016.05.021

García, T., Rodríguez, C., González-Castro, P., González-Pienda, J. A., & Torrance, M. (2016). Elementary students’ metacognitive processes and post-performance calibration on mathematical problem-solving tasks. *Metacognition and Learning*, *11*(2), 139-170. https://doi.org/10.1007/s11409-015-9139-1

Garofalo, J., & Lester, F. K., Jr. (1985). Metacognition , cognitive monitoring , and mathematical performance. *Journal for Research in Mathematics Education*, *16*(3), 163–176.

Gobert, J. D., Baker, R. S., & Wixon, M. B. (2015). Operationalizing and detecting disengagement within online science microworlds. *Educational Psychologist*, *50*(1), 43–57. https://doi.org/10.1080/00461520.2014.999919

González-Cabañes, E., García, T., Rodríguez, C., Cuesta, M., & Núñez, J. C. (2020). Learning and emotional outcomes after the application of invention activities in a sample of university students. *Sustainability (Switzerland)*, *12*(18), 7306. https://doi.org/10.3390/SU12187306

Hoffman, B. (2010). "I think I can, but I'm afraid to try": The role of self-efficacy beliefs and mathematics anxiety in mathematics problem-solving efficiency. *Learning and Individual Differences*, *20*(3), 276–283. https://doi.org/10.1016/j.lindif.2010.02.001

Hoffman, B., & Schraw, G. (2010). Conceptions of efficiency: Applications in learning and problem solving. *Educational Psychologist*, *45*(1), 1–14. https://doi.org/10.1080/00461520903213618

Hulse, T., Daigle, M., Manzo, D., Braith, L., Harrison, A., & Ottmar, E. R. (2019). From Here to There! Elementary: A game-based approach to developing number sense and early algebraic understanding. *Education Tech Research Development*, *67*, 423–441.

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

Kong, X. J., Wise, S. L., & Bhola, D. S. (2007). Setting the response time threshold parameter to differentiate solution behavior from rapid-guessing behavior. *Educational and Psychological Measurement*, *67*(4), 606–619. https://doi.org/10.1177/0013164406294779

Kramarski, B., & Gutman, M. (2006). How can self-regulated learning be supported in mathematical E-learning environments? *Journal of Computer Assisted Learning*, *22*(1), 24–33. https://doi.org/10.1111/j.1365-2729.2006.00157.x

Ku, K. Y. L., & Ho, I. T. (2010). Metacognitive strategies that enhance critical thinking. *Metacognition and Learning*, *5*(3), 251–267. https://doi.org/10.1007/s11409-010-9060-6

Li, J., Zhang, B., Du, H., Zhu, Z., & Li, Y. M. (2015). Metacognitive planning: Development and validation of an online measure. *Psychological Assessment*, *27*(1), 260–271. https://doi.org/10.1037/pas0000019

Losenno, K. M., Muis, K. R., Munzar, B., Denton, C. A., & Perry, N. E. (2020). The dynamic roles of cognitive reappraisal and self-regulated learning during mathematics problem solving: A mixed methods investigation. *Contemporary Educational Psychology*, *61*(April), 101869. https://doi.org/10.1016/j.cedpsych.2020.101869

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science*, *26*, 49–63. https://doi.org/10.1007/978-94-017-2243-8_5

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, *17*(3), 437–455. https://doi.org/10.1037/a0028085

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., … Roeser, R. (2000). Manual for the Patterns of Adaptive Learning Sciences (PALS). *University of Michigan - Ann Arbor*.

Montague, M., Enders, C., & Dietz, S. (2011). Effects of cognitive strategy instruction on math problem solving of middle school students with learning disabilities. *Learning Disability Quarterly*, *34*(4), 262–272. https://doi.org/10.1177/0731948711421762

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

Ottmar, E. R., Landy, D., Goldstone, R., & Weitnauer, E. (2015). Getting From Here to There! : Testing the effectiveness of an interactive mathematics intervention embedding perceptual learning. *Proceedings of the Annual Conference of the Cognitive Science Society*, 1793–1798.

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary Educational Psychology*, *21*(4), 325–344. https://doi.org/10.1006/ceps.1996.0025

Paquette, L., de Carvalho, A., & Baker, R. S. (2014). Towards understanding expert coding of student disengagement in online learning. In *36th Annual Cognitive Science Conference* (pp. 1126–1131). Quebec City, Canada.

Passolunghi, M. C., Caviola, S., De Agostini, R., Perin, C., & Mammarella, I. C. (2016). Mathematics anxiety, working memory, and mathematics performance in secondary-school children. *Frontiers in Psychology*, *7*(42), 1–8. https://doi.org/10.3389/fpsyg.2016.00042

Perels, F., Gürtler, T., & Schmitz, B. (2005). Training of self-regulatory and problem-solving competence. *Learning and Instruction*, *15*, 123–139. https://doi.org/10.1016/j.learninstruc.2005.04.010

R Core Team (2020). R: A language and environment for statistical computing.

Ramirez, G., Chang, H., Maloney, E. A., Levine, S. C., & Beilock, S. L. (2016). On the relationship between math anxiety and math achievement in early elementary school: The role of problem solving strategies. *Journal of Experimental Child Psychology*, *141*, 83–100. https://doi.org/10.1016/j.jecp.2015.07.014

Ramirez, G., Shaw, S. T., & Maloney, E. A. (2018). Math anxiety: Past research, promising interventions, and a new interpretation framework. *Educational Psychologist*, *53*(3), 145–164. https://doi.org/10.1080/00461520.2018.1447384

Rittle-Johnson, B., Matthews, P. G., Taylor, R. S., & McEldoon, K. L. (2011). Assessing knowledge of mathematical equivalence: A construct-modeling approach. *Journal of Educational Psychology*, *103*(1), 85–104. https://doi.org/10.1037/a0021334

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

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

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

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

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

Venkat, H., Askew, M., Watson, A., & Mason, J. (2019). Architecture of mathematical structure.

For the Learning of Mathematics, *39*(1), 13–17. Retrieved from https://flm-journal.org/Articles/15CDA7D47D380050F5587BF8C2BCB8.pdf

Vula, E., Avdyli, R., Berisha, V., Saqipi, B., & Elezi, S. (2017). The impact of metacognitive strategies and self-regulating processes of solving math word problems. *International Electronic Journal of Elementary Education*, *10*(1), 49–59. https://doi.org/10.26822/iejee.2017131886

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

Welsh, M., Cicerello, A., Cuneo, K., & Brennan, M. (1995). Error and temporal patterns in Tower of Hanoi performance: Cognitive mechanisms and individual differences. *Journal of General Psychology*, *122*(1), 69–81.

Wilson, J., Fernandez, M., & Hadaway, M. (1993). Mathematical problem solving. In P. S. Wilson (Ed.), *Research Ideas for the Classroom: High School Mathematics* (pp. 57–78). New York: MacMillan. https://doi.org/10.2307/2323271

Wise, S. L., Pastor, D. A., & Kong, X. J. (2009). Correlates of rapid-guessing behavior in low-stakes testing: implications for test development and measurement practice. *Applied Measurement in Education*, *22*(2), 185–205. https://doi.org/10.1080/08957340902754650

Wu, S. S., Barth, M., Amin, H., Malcarne, V., & Menon, V. (2012). Math anxiety in second and third graders and its relation to mathematics achievement. *Frontiers in Psychology*, *3*(JUN), 1–11. https://doi.org/10.3389/fpsyg.2012.00162

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