Metacognition and Learning (2025) 20:31
https://doi.org/10.1007/s11409-025-09437-x

Unpacking strategy efficiency: Examining the relations between pre-solving pause time and productivity in a digital mathematics game

Ji-Eun Lee[1] · Carly Siegel Thorp[1] · Arba Kamberi[2] · Erin Ottmar[1]

Received: 22 July 2024 / Accepted: 29 July 2025

Abstract
This study examined the relations among strategic planning, execution, and strategy efficiency during problem-solving in a digital algebra learning game with 7th-grade students. We used pre-solving pause time as a proxy indicator of strategic planning, and the productivity of the initial strategy as a measure of effective strategy execution. Additionally, we explored how these variables correlated with students' posttest scores assessing algebraic knowledge. Mediation analyses at both the problem and student levels indicated that longer pre-solving pause times were associated with greater strategy efficiency. When considering both the direct and indirect effects of pre-solving pause time on strategy efficiency, the results revealed a partial positive mediation through the productivity of the initial strategy. Lastly, the results of a path analysis showed that strategy efficiency significantly predicted algebraic knowledge with a positive effect. These findings suggest that longer pause times are associated with more efficient problem solving as they increase the likelihood of a productive initial step, highlighting a positive mediating role of execution in the relation between planning and strategy efficiency in algebraic problem solving.

Keywords Mathematics problem solving · Strategy efficiency · Pause time · Productivity · Mediation analysis

✉ Ji-Eun Lee
jlee13@wpi.edu

Carly Siegel Thorp
cthorp@wpi.edu

Arba Kamberi
akamberi@college.harvard.edu

Erin Ottmar
erottmar@wpi.edu

[1] Learning Sciences & Technologies, Social Science & Policy Studies Department, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609, USA

[2] Harvard University, Massachusetts Hall, Cambridge, MA 02138, USA

Published online: 23 August 2025

Introduction

Procedural fluency, defined as the ability to apply procedures and solve math problems flexibly, efficiently, and accurately, is one of the primary goals in mathematics education (National Council of Teachers of Mathematics [NCTM], 2023). Achieving procedural fluency requires *strategy efficiency*, in which students must identify when a particular strategy or procedure is more appropriate for a given problem than others. Additionally, to solve math problems efficiently, students must demonstrate procedural flexibility (Smedt et al., 2010; Verschaffel, 2024). They must recognize that there are multiple solution strategies, choose the appropriate one, and flexibly switch to an alternative if their initial choice does not work (Star & Rittle-Johnson, 2008).

Despite the importance of strategy efficiency and procedural flexibility in achieving mathematical proficiency, many studies have shown that students often struggle with using appropriate mathematical strategies to solve problems. In particular, students in the U.S. tend to apply procedures based solely on rote memorization, such as PEMDAS (Parentheses, Exponents, Multiplication/Division, Addition/Subtraction), or habitually solve problems from left to right, without noticing the relations between numbers (Gunnarsson et al., 2016; Hiebert, 2013; Schoenfeld, 1992). Furthermore, because conceptual and procedural knowledge precede the development of strategy efficiency and procedural flexibility (Rau et al., 2017; Schneider et al., 2011), math curricula and assessments have primarily focused on students' conceptual understanding and procedural knowledge rather than efficiency or flexibility (NCTM, 2023).

To address these issues, several studies have examined student-level (e.g., student behavior) or problem-level factors (e.g., problem type) influencing strategy efficiency or procedural flexibility to better support the development of mathematical proficiency (Newton et al., 2020; Ramirez et al., 2016; Schulz, 2023; Star & Rittle-Johnson, 2008). Among various variables, one important factor influencing students' strategy efficiency and procedural flexibility is students' use of *metacognitive* and *self-regulatory strategies* during their problem solving. Studies have revealed that students' *planning* before problem solving or *appropriate execution of strategy* is positively associated with their problem-solving performance (Coppersmith & Star, 2022; García et al., 2019). The Common Core State Standards (CCSS) also highlighted the importance of this planning phase in achieving procedural fluency. According to the CCSS, students need to "make conjectures about the form and meaning of the solution and plan a solution pathway rather than simply jumping into a solution attempt" (Common Core State Standards Initiative, 2022, p.6). Our earlier studies also found that middle school students who paused before solving problems as a method of planning (Chan et al., 2022) and appropriately executed the correct strategy (Lee et al., 2022, 2023) had greater strategy efficiency in a digital math learning game.

While numerous studies have explored students' planning and execution of strategies in relation to their math performance, the intricate relationship between these two phases (i.e., planning, execution), as well as their combined contribution to students' strategy efficiency, remains less understood (Coppersmith & Star, 2022). Thus, this study aims to explore the interplay between middle school students' planning, execution, and strategy efficiency in a digital mathematics learning game. Building upon our prior work (Chan et al., 2022), we use *pre-solving pause time* as a proxy measure of students' strategic thinking, *the productivity of their initial transformation* as a measure of the appropriate execution of strategy, and strategy efficiency scores in the game as a measure of strategy efficiency. Additionally, we examine how these factors affect students' algebraic knowledge after playing the game.

Mathematical problem solving as a process

Strategic planning and pre-solving pause time

To successfully solve math problems, students need to engage in a series of cognitive and metacognitive processes: they first need to read the problem and comprehend it, organize information, explore possible solutions, and plan a strategy to solve a problem (Carlson & Bloom, 2005; Coppersmith & Star, 2022; Garofalo & Lester, 1985). The importance of this *planning* phase in mathematical problem-solving has been demonstrated in numerous empirical studies (Callan & Cleary, 2019; Chang et al., 2006; Kramarski & Gutman, 2006; Simatupang et al., 2019; Vula et al., 2017). For instance, Callan and Clearly (2019) examined how 8th-grade students' strategic planning prior to solving the math problem and strategy use correlated with their math performance. Students' strategic planning was measured using a single item collected through the interviews, specifically, asking the students whether they had a plan to complete the problem or not. The results indicated that the students' strategic planning significantly and positively predicted their strategy use, which in turn, correlated with better math performance. Further, studies have shown that interventions prompting students to do planning (e.g., understanding the problem, thinking about possible strategies) prior to problem solving improved their mathematics performance (Kramarski & Gutman, 2006; Vula et al., 2017).

Traditionally, students' planning before problem-solving is assessed through self-reported measures or structured interview questions (e.g., Callan & Clearly, 2019). While these methods are widely used, they may produce incorrect information due to students' potential recall errors or response biases (García et al., 2016). One alternative way to gauge student strategic planning is by using the *time spent before executing an action*. For example, Ku and Ho (2010) conducted a qualitative study exploring how university students' metacognitive strategies related to their critical thinking performance. The students' use of metacognitive strategies was collected through think-aloud procedures. The results showed that the students who spent more time developing plans prior to the execution performed better on critical thinking tasks than those who spent less time planning.

The advancement in educational technology and data analytics enabled researchers to utilize log data from these technologies to measure students' cognitive and metacognitive processes. Several studies have used metrics such as the frequency or duration of pauses during problem solving as indicators of students' cognitive engagement in online learning environments (Gobert et al., 2015; Li et al., 2015; Paquette et al., 2014). For example, Paquette et al. (2014) identified students' disengagement (i.e., gaming behaviors) in an online intelligent tutoring system for algebraic learning. They posited that students' pauses before solving or requesting support after longer than six seconds were an indicator of thinking and a sincere attempt, whereas a pause time shorter than six seconds was an indicator of guessing or non-thinking. Using these identifiers, they successfully trained a model identifying the sequential patterns of disengagement (e.g., incorrect → guessing → attempt). One of our earlier works (Chan et al., 2022) also examined how middle school students' pause time before problem solving affected their strategy efficiency, using a small subset of problems (i.e., 20 problems) in a digital algebraic learning game. We posited that pause time, measured by the duration the students spent before making the initial transformation on each problem, could serve as a proxy measure of their strategic thinking and planning. The results indicated that students' longer pauses prior to problem solving were positively related to strategy efficiency in the game after controlling for prior knowledge.

Together, previous research suggests that taking time to understand the problem and devise a plan before attempting to solve it contributes to efficient problem-solving. Building upon our earlier study (Chan et al., 2022), we posit that students' pauses prior to problem solving can serve as an indicator of their planning process as they take time to think through the problem and solution before attempting to solve it. Accordingly, we use pre-solving pause time as a proxy measure of strategic planning and hypothesize that students' pre-solving pause time is positively related to strategic efficiency in a digital algebra learning game.

Execution of strategy

After students pause to think through a math problem and devise a plan for a problem-solving strategy, they then need to turn their well-thought plans into appropriate actions, in other words, *executing an action* based on the strategy (Carlson & Bloom, 2005; Coppersmith & Star, 2022; Garofalo & Lester, 1985). While *planning* refers to identifying the problem and exploring possible strategies through reading a problem, recalling a similar problem, or thinking, *execution* indicates students regulating their behaviors or making an action that conforms to their plans, such as performing a calculation or writing a response (García et al., 2019; Garofalo & Lester, 1985).

Several studies have explored how students' execution of strategy is related to their performance in math learning contexts. For example, García et al. (2019) examined the relationship between upper elementary students' problem-solving process and performance on two mathematical problems. The findings showed that when students planned strategies before execution and revisited their strategies after execution, these were both positively related to math performance. Additionally, the students who spent more time executing the task than planning were more likely to solve the second problem correctly. Our previous work also found that students' productive or valid execution of strategy in their initial solution attempt is associated with better performance. For example, Lee et al. (2022) investigated how middle school students' productivity of their first strategy execution correlated with problem-solving efficiency on two problems within a digital algebraic learning game. Productivity was defined as whether a student executed an appropriate mathematical transformation towards the goal state of the problem on their first step. The results indicated that productivity was positively related to problem-solving efficiency in the game after accounting for students' prior knowledge. Our follow-up study (Lee et al., 2023) then explored how middle school students' in-game behaviors related to their math performance after playing the digital algebra learning game. The results showed that the mathematical validity of the first problem step (i.e., whether the students made a mathematically valid first transformation without making any errors) was the most influential predictor of the posttest scores out of 37 in-game behavioral features included in the prediction model.

Together, these findings suggest that students' correct, or valid, execution of strategy is positively associated with their math performance. However, although many studies have explored how execution is related to math performance, little is known about the complex relationship between planning and execution, as noted by Coppersmith and Star (2022). Thus, our study intends to explore the intricate relationship between planning, execution, and performance, as well as the mediating effect of execution between planning and performance.

Strategy efficiency in mathematics learning

In the context of mathematical problem-solving, *strategy efficiency* has been defined as using a strategy with the fewest steps and/or the computation that involves simple (e.g., small or whole numbers) rather than complex numbers (e.g., large numbers, fractions; Xu et al., 2017). For example, students can solve $3(4 + x) = 21$ using a two-step strategy by dividing both sides by 3 (i.e., [step 1] $4 + x = 7$, [step 2] $x = 3$) or using a three-step strategy (i.e., [step 1] $12 + 3x = 21$, [step 2] $3x = 9$, [step 3] $x = 3$). To solve the problem using the most efficient 2-step strategy, students need to recognize that the left side of the equation is in production form, identify the relationship between the factors 3 and 21, and understand that division by 3 will simplify the equation directly. Utilizing efficient strategies 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 more challenging content.

Students' strategy efficiency may influence their performance on algebra assessments (Ramirez et al., 2016), and algebra math performance may also affect how efficiently they apply problem-solving strategies. This potentially bidirectional association between knowledge and strategy use has been demonstrated across different age groups and mathematics topics (Star & Rittle-Johnson, 2008; Torbeyns et al., 2006). For instance, in algebraic problem-solving, middle school students with higher mathematics achievement are more likely to use a more efficient strategy compared to those with lower mathematics achievement (Newton et al., 2020; Wang et al., 2019). While the relationship between strategy efficiency and math knowledge may be bidirectional, this study focuses on whether students' strategy efficiency in the game predicts their posttest scores of algebraic knowledge after playing the game.

The current study

This study investigated how middle school students' strategic planning, execution, and strategy efficiency during algebraic problem solving were related to each other in a digital mathematics learning game. We used pre-solving pause time (hereafter, pause time) as a proxy measure of strategic planning, and productivity of the initial strategy (hereafter, productivity) as a measure of appropriate execution of strategy. We investigated these relationships not only at the student level (RQ2) but also at the problem level (RQ1) to examine how these variables interact when accounting for differences in problem structures and difficulty. Additionally, we examined how these variables correlated with students' posttest scores assessing algebraic knowledge after playing the game. We addressed the following three research questions, and Fig. 1 represents the conceptual model of the study.

RQ1 Does productivity mediate the relation between the pause time and strategy efficiency at the problem level?

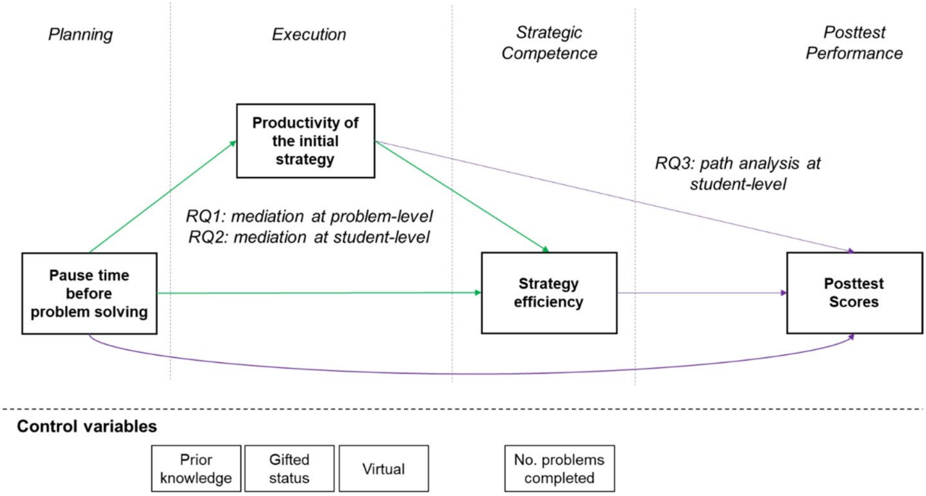

Fig. 1 A conceptual model linking predictors to the strategy efficiency and posttest scores

RQ2 Does productivity mediate the relation between the pause time and strategy efficiency at the student level?

RQ3 Do productivity, pause time, and strategy efficiency directly affect students' posttest scores assessing algebraic knowledge?

Method

Research context

From Here to There! (FH2T) is a digital mathematics learning game developed based on cognitive and learning sciences theories to improve students' conceptual understanding of algebra as well as procedural fluency (Ottmar et al., 2015). The goal of each problem in the game is to turn an algebraic expression (e.g., $10 \times 20a \times 10 \times 5$ in Fig. 2a) into a mathematically equivalent but perceptually different goal expression (e.g., $a \times 100 \times 100$ in a white box in Fig. 2a) through a series of transformations (Fig. 2b and e).

In this game, the numbers and mathematical symbols are made into tactile objects, which allow students to dynamically manipulate them on a screen using various gesture actions, such as clicking, dragging, or typing on a keypad, to reach the goal expression. This system also provides a fluid visualization that shows students how their gesture actions change the algebraic expressions in real-time. The expression will change if an action (e.g., multiplying 10×10) is mathematically permissible in its resulting expression (i.e., 100). If the students attempt an action that is not mathematically permissible (e.g., trying to add $2 + 2x$), they will not be allowed to complete the action, and the expression will be unchanged, indicating they need to instead try a different action. When they reach the goal state using the most efficient solution pathway, with the fewest steps possible, students get rewards in the form of clovers (Fig. 2f). However, the number of clovers decreases if they use more steps than necessary to

Fig. 2 A sample problem in the game that shows the steps students take to transform the equation from the start to the goal state

be considered efficient. Notably, the system does not provide the optimal solution pathway to encourage students to retry the problem using a more efficient or different way. In this way, students can easily identify the underlying structure of algebraic expressions, think more flexibly, and realize that math problems can be solved in multiple different ways, rather than a single correct pathway.

The game consists of 14 worlds, each containing 18 problems, totaling 252 problems. Of these, 39 problems are tutorial problems that introduce new gesture actions in the game. To proceed to the next world, students must complete 11 consecutive problems in each world. The last four problems in each world are optional, allowing students to skip them if they want. Each world represents a different algebraic topic, such as addition, multiplication, division, and factoring, with an increased level of difficulty. Problems in the earlier stages of the game involve simple tasks such as simplifying expressions (e.g., transforming 24+y+6+13 into 13+y+30) or combining like terms (e.g., transforming 2y+6+3y into 5y+6). However, in later stages, problems include more advanced topics such as factoring (e.g., transforming 9×b+27×c into (3b+9c)×3) or the distributive property (e.g., transforming (b+9)×(c-4) into bc+9c-4b-36). Previous empirical studies have consistently shown that the game is effective in improving students' algebraic learning (Chan et al., 2022; Decker-Woodrow et al, 2023; Hulse et al., 2019).,

Participants and procedure

The sample was drawn from a larger Randomized Controlled Trial (RCT) conducted between the fall of 2020 and spring of 2021 across 11 schools in the Southeastern U.S., comprised of 10 in-person schools and one virtual academy (Ottmar et al., 2023). The RCT examined the impacts of three educational technologies on 7th-grade students' algebraic understanding (Decker-Woodrow et al., 2023). The students in the RCT were randomly assigned to one of three educational technologies across the four different conditions. Results revealed that students in the gamified learning conditions had significantly higher posttest scores than those in the business-as-usual condition with online math problem sets. Please see Decker-

Woodrow et al. (2023) for the details and the results of the RCT. For this study, we focused on students assigned to FH2T (N = 1,430), one of the gamified learning conditions.

Before the intervention, the students completed a pretest assessing their algebraic knowledge. Subsequently, they participated in nine intervention sessions throughout the school year, each lasting 30 min and implemented in their regular math classes. As the students played the game at their own pace using their devices, they completed a different number of problems in the game when the intervention ended. Following the intervention period, students completed a posttest assessing their algebraic knowledge, using problems isomorphic to those in the pretest.

Of the 1,430 students assigned to the gamified intervention, 754 students who completed the pretests and posttest assessing their algebraic knowledge were included in our analyses. Further, we eliminated the students who only completed the instruction or tutorial problems (n = 21), resulting in an analytic sample of 733 students. Student demographic information was provided by the school district. Of the 733 students, 48% (n = 354) were female, 53% were White, 23% were Asian, 16% were Hispanic/Latino, 5% were Black, and the remaining students were in other racial categories. Additionally, 17% (n = 126) were identified as gifted as determined by the school district based on a nationally normed test. As data was collected during the pandemic, the school district allowed the students and their families to choose between in-person or fully asynchronous virtual academy for the school year, and 33% (n = 238) opted for virtual classes. Students in both in-person and virtual learning covered the same curriculum throughout the school year; however, those who opted for the virtual academy participated in the intervention fully remotely.

Measures

Although the students had the opportunity to complete up to 157 problems, except for tutorial and optional problems, the students completed 60.67 problems on average (SD = 19.91), resulting in many null values for most measures. For this reason, we included up to World 7: Order of Operations (i.e., 77 problems except for tutorial and optional problems) in the game, which most students reached.

Strategy efficiency

As mentioned earlier, strategy efficiency was defined as selecting a strategy that involves the fewest number of steps to complete a problem. Strategy efficiency was calculated by dividing the optimal number of steps to solve the problem by the number of steps it took for the student to solve the problem. For example, the optimal number of steps to complete the problem in Fig. 2 (start state: $10 \times 20a \times 10 \times 5$, goal state: $a \times 100 \times 100$) is using two steps strategy ($10 \times 20a \times 10 \times 5 \rightarrow 20a \times 100 \times 5$ [step 1: commuting 10 and multiplying 10 and 10] $\rightarrow a \times 100 \times 100$ [step 2: commuting 20 and multiplying 20 and 5]). When a student solves the problem using two steps, the strategy efficiency score for the student is 1 (i.e., 2 divided by 2).

However, if a student completes the problem using three steps, the strategy efficiency score for the student is 0.67 (i.e., 2 divided by 3). We used the average strategy efficiency scores for each student across all problems attempted for the student-level analysis (RQ2 and RQ3).

Pre-solving pause time

Pre-solving pause time refers to the time spent before making an initial transformation on each problem. Using the timestamp data recorded in the database, we measured the time between the moment a student received the problem and the moment the student interacted with the technology to make an action within the game (e.g., clicking an operator, dragging a number). To account for the students spending longer on problems due to spending more time before they initiated an action, we used *percent pause time* (Chan et al., 2022; Li et al., 2015). The percent pause time was computed by dividing the pre-solving pause time by the total time spent on a problem. For example, if a student spends 6 s before making a first transformation on a problem and 30 s to complete that problem, the pause time for the student is 0.2 ($6 \div 30$). For the student-level analysis (RQ2 and RQ3), we used the average percent pause time across all problems attempted.

Productivity of solution attempt

The productivity of the solution attempt was measured in terms of whether a student made a productive step in transforming the expression towards the goal state expression. For example, for the problem with the start state of "$10 \times 20 \times a \times 10 \times 5$" and the goal state of "$a \times 100 \times 100$", the first step of transforming "$10 \times 20 \times a \times 10 \times 5$" into "$10 \times 100 \times a \times 10$" was coded as a productive solution attempt as this action resulted in creating the "100" found in the goal state expression and therefore brought the student closer to the goal state. Multiplying 10 and 5 to transform "$10 \times 20 \times a \times 10 \times 5$" into "$10 \times 20 \times a \times 50$" was coded as a non-productive attempt because "50" does not match any numbers in the goal state expression and therefore did not bring the student any closer to the goal state (See Table 1 for more examples).

Here, we focused on the productivity of the students' initial solution attempt, which is an action executed right after taking the time to pause. The productivity of each problem was hand-coded by two coders as productive (1) or nonproductive (0). Any discrepancies were discussed and resolved before the analyses. Note that the productivity of the students' initial solution attempt was not coded if a step was taken by fewer than 5% of the students who solved that problem.

Algebraic knowledge

Before and after the intervention, students' algebraic knowledge was measured using 10 items adapted from a previously validated measure (Star et al., 2015). The assessment measured three different aspects: conceptual knowledge (4 items), procedural knowledge (3

Table 1 Examples of productivity of initial solution attempt

Start State	Goal State	Productive first steps	Non-productive first steps
$10 \times 20 \times a \times 10 \times 5$	$a \times 100 \times 100$	$10 \times 100 \times a \times 10$ $10 \times a \times 10 \times 100$ $100 \times 20 \times a \times 5$ $20 \times a \times 100 \times 5$ $10 \times 10 \times 20 \times a \times 5$*	$10 \times 20 \times a \times 50$ $200 \times a \times 10 \times 5$

*Moving numbers to be added adjacent to each other was considered productive in bringing the student closer to the goal state items), and mathematical flexibility (3 items). Each item was scored as correct (1) or incorrect (0), and the overall score was computed. The inter-item reliability of the items was KR-20 = 0.74 at the pretest and KR-20 = 0.72 at the posttest, respectively. The pretest scores were used as a covariate, and the posttest scores were used as an outcome variable in the path model for RQ3.

Covariates

Since the RCT was conducted during the COVID-19 pandemic, some students participated in the study virtually while others completed it in person. To account for this potential source of variation, we included the students' enrollment in a virtual classroom (virtual = 1, in-person = 0) as a covariate in the analyses. The status of students who were considered "gifted" (yes = 1, no = 0) was also included as a covariate to account for the possible differences in prior mathematics knowledge, in addition to accounting for students' pretest scores. Students' gender was not included in the model as there was no statistically significant difference in strategy efficiency by gender, $t(726) = -0.12$, $p = .91$. In addition, the total number of problems completed were included in the student-level analysis to control the potential influence of these variables on outcome variables.

Data pre-processing

The initial dataset included 44,471 problem-level observations from 733 students. However, 19,524 problem-level observations were excluded due to missing data on the productivity of the students' initial solution attempt, which were not coded if a step was taken by fewer than 5% of the students who solved that problem. An additional 135 problem-level observations were excluded due to missing pre-solving pause time data from incomplete attempts, and 509 observations were removed for a problem in which all students received a strategy efficiency score of 1. This process resulted in a final dataset of 24,303 problem-level observations from 733 students. For problem-level analyses (RQ1), we used 24,303 problem-level observations, and for student-level analyses (RQ2, RQ3), we used aggregated averages of the problem-level data across the 733 students.

Analytical approach

For RQ1, we performed a mediation analysis in combination with mixed-effects models at the problem level. Specifically, we included the pause time as an independent variable (IV), strategy efficiency as a dependent variable (DV), and productivity of the initial solution attempt as a mediator, with the student modeled as a random effect. As the mediator was a binary variable and the DV was a continuous variable, we conducted a mixed-effects logistic regression analysis for the mediator (Model 1.2) and mixed-effects linear regression analyses for the DV (Models 1.1 and 1.3) using the "glmer" and "lmer" functions from the lme4 package in R (Bates et al., 2015), respectively. Then, we took the output from the two models (Models 1.2 and 1.3) and used the "mediate" function from the mediation package in R (Tingley et al., 2014) to estimate the indirect effect, direct effect, and total effect of the mediation model. Pretest scores and two demographic variables (i.e., gifted status, virtual status) were included as covariates.

For RQ2, we performed a mediation analysis at the student level. We used the same IV, mediator, and DV as RQ1, but used the aggregated average across all problems attempted. As all outcome variables were continuous variables, we conducted multiple linear regression analyses for the mediator (Model 2.2) and the DV (Models 2.1 and 2.3) using the "lm" function in R. We then took the output from the two models (Models 2.2 and 2.3) and used the "mediate" function from mediation package in R to estimate the indirect effect, direct effect, and total effect of the mediation model. As with RQ1, pretest scores and two demographic variables (i.e., gifted status, virtual status) were included as covariates. Additionally, the number of problems completed was included as a covariate for the model predicting strategy efficiency.

Lastly, we conducted a path analysis using the "sem" function from the lavaan package in R (Rosseel, 2012) to investigate the relationships among pause time, productivity, strategy efficiency, and posttest scores (RQ3). Pretest scores and two demographic variables were included as covariates in the model.

Results

RQ 1: Relationship between pause time and strategy efficiency as mediated by productivity at the problem level

We first conducted descriptive statistics and correlation analyses using the raw problem-level variables (see Table 2). Descriptive statistics for the student-level variables (i.e., pretest scores, demographic variables) are presented in Table 4. All variables were significantly associated with strategy efficiency. Productivity showed the strongest positive association with strategy efficiency (r_{pb} = .29, p<.001), followed by pause time (r=.20, p<.001).

Before conducting analyses, strategy efficiency was standardized (z-score transformed) at the problem level to account for variability in difficulty across problems. Additionally, continuous variables were grand-mean centered, and binary variables were recoded to −0.5 or 0.5 to improve the interpretability of model parameters.

We first estimated a null model with no predictors to examine the proportion of variance in strategy efficiency attributable to between-student differences. The intraclass correlation coefficient (ICC) was 0.053, indicating that 5.3% of the variance in strategy efficiency was due to between-student differences. To investigate the role of productivity in the relationship between pause time and strategy efficiency at the problem level, we first conducted a mixed-effects linear regression analysis of pause time predicting strategy efficiency, while

Table 2 Means, standard deviations, and correlations at problem level (N=24,303)

	1	2	3
1. Pause time	-	-	-
2. Productivity	0.14***	-	-
3. Strategy efficiency	0.20***	0.29***	-
M	0.51	0.84	0.92
SD	0.22	0.37	0.16
Min.	0.00	0.00	0.11
Max.	1.00	1.00	1.00
Skewness	0.15	−1.81	−1.80
Kurtosis	−0.69	1.26	2.05

***p<.001

controlling for the effects of gifted status, virtual status, and pretest scores (Model 1.1 in Table 3). The pause time significantly predicted the strategy efficiency scores with a positive effect (β = 0.150, p < .001), meeting the first requirement of the mediation analysis.

Next, Model 1.2 tested the relationship between pause time and productivity (i.e., mediator) using a mixed-effects logistic regression model, while the effects of gifted status, virtual status, and prior knowledge (i.e., pretest scores) were controlled (see Table 3). Pause time significantly predicted the productivity of the first step with a positive effect (β = 0.379, p < .001). In other words, the students who exhibited a longer pause time before solving a problem were more likely to make a productive first step.

Third, we conducted a mixed-effects linear regression analysis with pause time and productivity predicting strategy efficiency, while the effects of gifted status, virtual status, and pretest scores were controlled (Model 1.3 in Table 3). Productivity significantly predicted strategy efficiency with a positive effect (β = 0.081, p < .001), suggesting that the students who made a productive first step had higher strategy efficiency scores at the problem level. With the productivity in the model, pause time still significantly predicted strategy efficiency with a positive effect (β = 0.139, p < .001). Notably, the positive correlation between pause time and strategy efficiency decreased when productivity was added (Model 1.3) compared to the model without productivity (Model 1.1). This indicates that the effect of pause time on strategy efficiency reduces when productivity is considered. All three covariates were significantly related to strategy efficiency: prior knowledge (β = 0.092, p < .001), gifted status (β = 0.037, p = .001), and virtual status (β = −0.029, p = .009).

Lastly, we estimated the indirect effect of the mediation model. The results indicated a statistically significant indirect effect of pause time on strategy efficiency through the

Table 3 Standardized direct and indirect effects on strategy efficiency at the problem-level

Variable	Model 1.1: Effect of IV on DV (Strategy efficiency)			Model 1.2: Effect of IV on Mediator (Productivity)			Model 1.3: Effect of IV and Mediator on DV (Strategy efficiency)		
	β	SE	p	β	SE	p	β	SE	p
Fixed effects									
Intercept	−0.007	0.013	0.335	1.684***	0.027	< 0.001	−0.007***	0.014	< 0.001
Pause time (IV)	0.150***	0.029	< 0.001	0.379***	0.084	< 0.001	0.139***	0.029	< 0.001
Productivity (M)	-	-	-	-	-	-	0.081***	0.017	< 0.001
Prior knowledge	0.096***	0.004	< 0.001	0.131***	0.008	< 0.001	0.092***	0.004	< 0.001
Gifted status	0.039***	0.028	< 0.001	0.070**	0.055	0.001	0.037**	0.027	0.001
Virtual status	−0.029**	0.024	0.008	−0.017	0.045	0.431	−0.029**	0.023	0.009
Random effects	Variance		*SD*	Variance		*SD*	Variance		*SD*
Student (Intercept)	0.038		0.194	0.022		0.149	0.036		0.191
Residual	0.926		0.962	-		-	0.920		0.959
Indirect effect at the problem level									
Pause time → Productivity → Strategy efficiency					0.004***			-	< 0.001

Note: **p < .01, ***p < .001

mediator of productivity (β = 0.004, p < .001) while controlling for students' gifted status, virtual status, and pretest scores. In other words, the students who took longer pre-solving pause times and made productive first transformations on each problem were more likely to achieve high strategy efficiency scores. Figure 3 illustrates the relationship between pause time and strategy efficiency as mediated by productivity at the problem level.

RQ 2: Relationship between pause time and strategy efficiency as mediated by productivity at the student level

We conducted descriptive statistics and correlation analyses at the student level (see Table 4). All variables were significantly correlated with strategy efficiency. Specifically, productivity showed the strongest positive association with strategy efficiency among all variables (r = .42, p < .001), and pause time also had a positive association with strategy efficiency (r = .28, p < .001).

We first tested the relationship between pause time and strategy efficiency at the student level using a multiple linear regression model, controlling for prior knowledge, gifted status, virtual status, and number of problems completed (see Model 2.1 in Table 5). The pause time significantly predicted the efficiency scores with a positive effect (β = 0.237, p < .001), meeting the first requirement of the mediation analysis.

Second, we tested the relationship between pause time and productivity (i.e., mediator) at the student level using a multiple linear regression model while the effects of prior knowledge, gifted status, virtual status, and number of problems completed were controlled (see Model 2.2 in Table 5). The pause time was positively related to the productivity of the

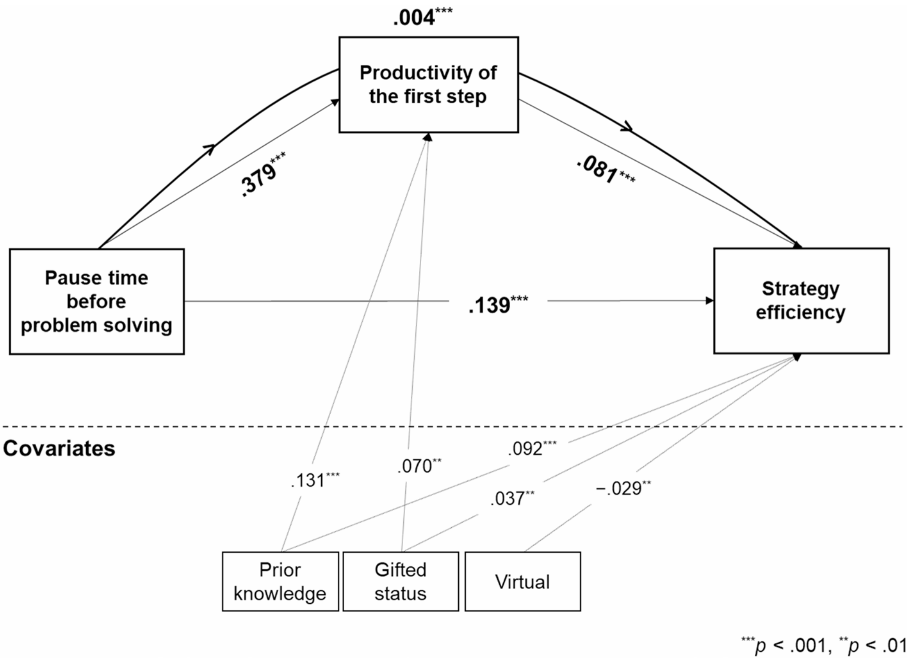

Fig. 3 Standardized coefficients for the relationship between pause time and strategy efficiency as mediated by productivity at the problem level

Table 4 Means, standard deviations, and correlations at the student level (N=733)

	1	2	3	4	5	6	7
1. Pause time	-	-	-	-	-	-	-
2. Productivity	0.12^{***}	-	-	-	-	-	-
3. Strategy efficiency	0.28^{***}	0.42^{***}	-	-	-	-	-
4. Problems completed	0.03	0.36^{***}	0.35^{***}	-	-	-	-
5. Pretest scores	0.16^{***}	0.30^{***}	0.34^{***}	0.30^{***}	-	-	-
6. Gifted status	0.06	0.20^{***}	0.21^{***}	0.16^{***}	0.43^{***}	-	-
7. Virtual status	0.10^{*}	0.11^{***}	0.12^{***}	0.04	0.47^{***}	0.25^{***}	-
M	0.52	0.83	0.92	33.85	4.80	0.17	0.32
SD	0.06	0.08	0.05	11.19	2.69	0.38	0.47
Min.	0.31	0.50	0.67	3.00	0.00	0.00	0.00
Max.	0.77	1.00	1.00	51.00	10.00	1.00	1.00
Skewness	0.14	−0.80	−1.14	−0.72	0.32	1.74	0.75
Kurtosis	0.95	0.74	2.22	−0.42	−0.93	1.04	−1.44

$^{*}p<.05$, $^{**}p<.01$, $^{***}p<.001$

Table 5 Standardized direct and indirect effects on strategy efficiency at the student-level

Variable	Model 2.1: Effect of IV on DV (Strategy efficiency)			Model 2.2: Effect of IV on Mediator (Productivity)			Model 2.3: Effect of IV and Mediator on DV		
	β	SE	*p*	β	SE	*p*	β	SE	*p*
Direct effect at the student level									
Intercept	−0.015	0.044	0.723	−0.032	0.045	0.478	−0.007	0.042	0.872
Pause time (IV)	0.237^{***}	0.033	<0.001	0.080^{*}	0.034	0.019	0.215^{***}	0.032	<0.001
Productivity (M)	-	-	-	-	-	-	0.269^{***}	0.034	<0.001
Prob. completed	0.278^{***}	0.034	<0.001	0.292^{***}	0.036	<0.001	0.199^{***}	0.034	<0.001
Prior knowledge	0.200^{***}	0.041	<0.001	0.162^{***}	0.043	<0.001	0.156^{***}	0.040	<0.001
Gifted status	0.190^{*}	0.095	0.045	0.230^{*}	0.099	0.020	0.129	0.091	0.159
Virtual status	−0.053	0.079	0.499	−0.022	0.082	0.788	−0.047	0.075	0.532
Indirect effect at the student level									
Pause time → Productivity → Strategy efficiency							0.022^{*}	-	0.024

$^{*}p<.05$, $^{***}p<.001$

first step (β=0.080, p=.019), indicating that the students who exhibited a longer pause time before solving a problem were more likely to make a productive first step at the student level.

Third, we conducted a multiple linear regression model with pause time and productivity predicting strategy efficiency at the student level, while controlling for students' prior knowledge, gifted status, virtual status, and number of problems completed (see Model 2.3 in Table 5). Productivity significantly predicted strategy efficiency with a positive effect (β=0.269, p<.001), suggesting that students who had a productive first step had higher strategy efficiency scores. With the productivity in the model, pause time still significantly predicted strategy efficiency with a positive effect (β=0.215, p<.001). Like the problem-level analysis, the positive correlation between pause time and strategy efficiency decreased when productivity was added (Model 2.3) compared to the model without productivity (Model 2.1), indicating that the effect of pause time on strategy efficiency reduced when productivity was considered. Among the covariates, two of them were significantly related to strategy efficiency: prior knowledge ($\beta=0.156$, $p<.001$) and the number of problems completed ($\beta=0.199$, $p<.001$).

Finally, there was a significant positive indirect effect of pause time on strategy efficiency through the productivity of the first step ($\beta=0.022$, $p=.024$). In other words, students who exhibited longer pause times before solving a problem and made productive first steps tended to have higher strategy efficiency scores. Figure 4 illustrates the relationship between pause time and strategy efficiency as mediated by productivity at the student level.

RQ 3: Relationships among pause time, productivity, strategy efficiency, and posttest scores

We conducted a path model predicting posttest scores of students' algebraic knowledge using the student-level data. As virtual status and gifted status were not significant predictors of strategy efficiency in Model 2.3, these covariates were included only in Model 3.3, predicting posttest scores, to improve the model fit. The path model showed a good fit to the data, $\chi^2(4)=7.74$, $p=.10$, CFI = 0.996, RMSEA = 0.036. Table 6 shows the results of the standardized coefficients for posttest scores after controlling for covariates.

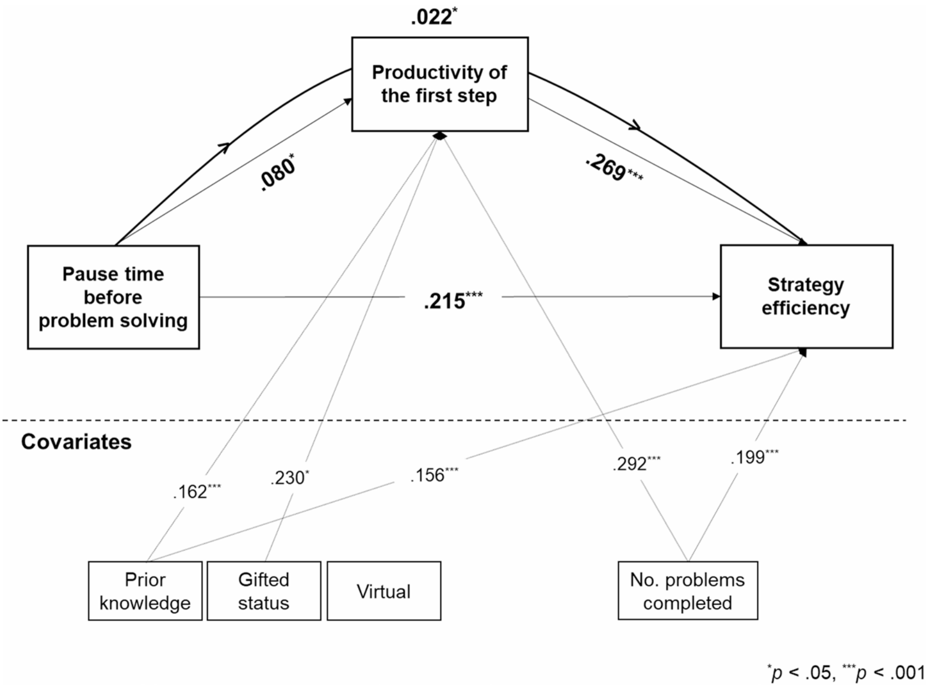

Fig. 4 Standardized coefficients for the relationship between pause time and strategy efficiency as mediated by productivity at the student level

Posttest scores were significantly predicted by strategy efficiency (β=0.097, p=.001) while controlling for the four covariates. Pause time (p=.766) and productivity did not significantly predict the posttest scores (p=.092). Regarding the association between covariates and posttest scores, all covariates (i.e., the number of problems completed, prior knowledge, gifted status, virtual status) significantly and positively predicted posttest scores. The path model (Model 3.3 in Table 6) accounted for 52.5% of the variance in students' algebraic posttest scores. Figure 5 represents the results of the full path analysis predicting posttest scores, with significant standardized path coefficients.

Discussion

In this study, we investigated the relation between students' pause time before problem-solving, their choice of a productive first step, and strategy efficiency in a digital algebraic learning game. Specifically, we used students' pause time before problem-solving as a proxy indicator of strategic planning, their choice of a productive first step as an indicator of appropriate execution of strategy, and strategy efficiency as an outcome variable.

Positive direct effect of pre-solving pause time on strategy efficiency

In our first finding, students who paused longer before problem-solving exhibited more efficient strategies by using fewer steps when solving an algebraic problem. In addition, the results were consistent at both the problem level and student levels, indicating that the relation between pre-solving pause time and strategy efficiency holds across levels of analysis.

Our results replicate the findings of our prior study (Chan et al., 2022) and align with previous research (Li et al., 2015), demonstrating that students' pre-solving pause time is positively associated with strategy efficiency. In other words, students who take time to plan, by reading the problem and comprehending it, organizing information, exploring possible solutions, and planning a strategy to solve a problem, tend to solve problems more efficiently than those who immediately jump into a solution attempt (Callan & Cleary, 2019; Chang et al., 2006; Kramarski & Gutman, 2006; Simatupang et al., 2019; Vula et al., 2017).

Table 6 Path analysis results

Variable	Model 3.1: Productivity			Model 3.2: Strategy Efficiency			Model 3.3: Posttest scores		
	β	SE	p	β	SE	p	β	SE	p
Variables of Interest									
Pause time	0.079^{*}	0.046	0.020	0.214^{***}	0.024	<0.001	0.008	1.304	0.766
Productivity	-	-	-	0.274^{***}	0.019	<0.001	0.049	1.048	0.092
Strategy efficiency	-	-		-	-	-	0.097^{*}	1.907	0.001
Covariates									
Prob. completed	0.296^{**}	<0.001	<0.001	0.202^{***}	<0.001	<0.001	0.132^{***}	0.008	<0.001
Prior knowledge	0.194^{***}	0.001	<0.001	0.164^{***}	0.001	<0.001	0.469^{***}	0.037	<0.001
Gifted status	-	–	-	-	-	-	0.133^{***}	0.220	<0.001
Virtual status	-	-	-	-	-	-	0.133^{***}	0.183	<0.001
R^2	0.172			0.302			0.525		

$^{*}p$<.05, $^{***}p$<.001

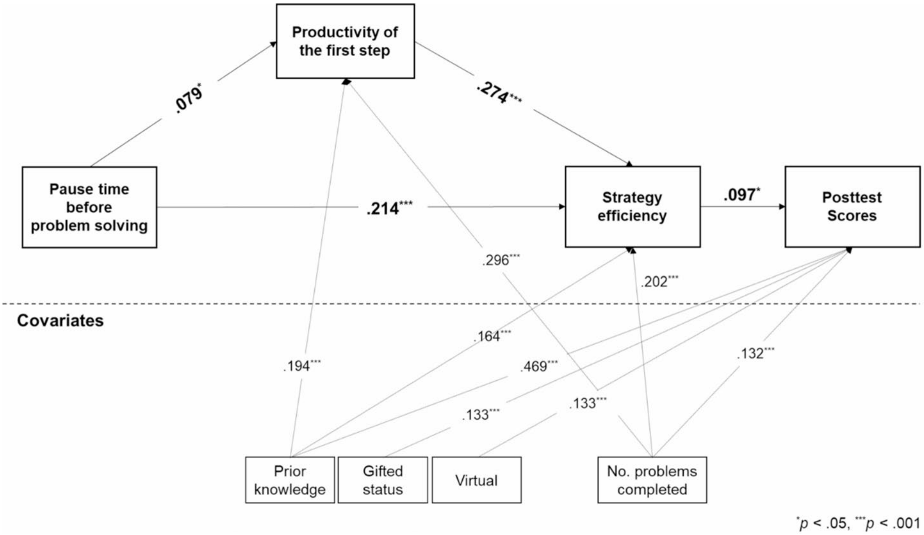

Fig. 5 Results of the path analysis for posttest scores

These findings highlight the critical role of strategic planning prior to problem solving in enhancing mathematical efficiency.

Traditionally, researchers have relied on self-reported measures or interviews to assess students' planning before problem solving; however, these methods may be prone to inaccurate information due to students' potential recall errors or response biases (García et al., 2016). Alternatively, particularly in digital learning environments, researchers have used the frequency or duration of pauses during their problem-solving as a proxy indicator of students' thinking and planning processes (Gobert et al., 2015; Li et al., 2015; Paquette et al., 2014). Our findings support this approach, suggesting that the time spent before executing an action may serve as a meaningful indicator of cognitive engagement in math problem-solving.

Positive indirect effect of pause time on strategy efficiency through the productivity of the initial strategy

When considering *both* the direct and indirect effects of pause time on strategy efficiency, the results indicated a partial mediation through the productivity of the initial strategy. Specifically, the positive direct effect of pre-solving pause time on strategy efficiency was reduced, but not eliminated, when accounting for its indirect effect through productivity. This pattern was found at both the problem and student levels, indicating a consistent mediation effect across levels of analysis.

More specifically, students who paused longer before problem-solving were more likely to make a productive initial step than those who paused for a shorter time. In turn, students who made a productive initial step were more likely to have higher strategy efficiency compared to those who made a non-productive initial step. In other words, the productivity of the initial step served as a mediating pathway through which pre-solving pause time affected strategy efficiency.

These findings suggest that students who pause before problem-solving *to plan their first step* are more likely to execute a productive first step, which in turn leads to more efficient problem-solving (Chan et al., 2022; Li et al., 2015). This result also aligns with previous research indicating that students' appropriate execution of strategy, such as regulating their behaviors or taking action to conform to plans, is positively related to problem-solving efficiency or math performance (García et al., 2019; Lee et al., 2022, 2023).

Although both pre-solving pause time and productivity of the initial step had significant direct effects on strategy efficiency, the indirect effect of pre-solving pause time through productivity was relatively small (β ranging from 0.004 to 0.02). This suggests that while productivity partially mediates the relation between pre-solving pause time and strategy efficiency, much of the effect may operate through other cognitive or behavioral processes beyond initial strategy selection. Alternatively, this may also be explained by the flexibility in problem-solving strategies among students who exhibited longer pause times. Even though they began with a nonproductive initial step, they might have recovered and ultimately solved the problem using a less efficient but still effective approach.

Taken together, our findings suggest that longer pre-solving pause times are associated with more efficient math problem solving. These findings also highlight the importance of accounting for the execution of strategy (e.g., the productivity of the initial strategy) when examining how students' planning (e.g., pre-solving pause time) affects strategy efficiency in math problem solving.

Positive effects of productivity and strategy efficiency on post-assessment of algebraic knowledge

Lastly, we examined whether pause time, productivity, and strategy efficiency significantly predicted students' algebraic knowledge, measured by posttest scores. Our findings indicated that students' strategy efficiency in the game was a significant positive predictor of their algebraic knowledge, above and beyond their prior knowledge, virtual status, and gifted status. These results are consistent with other studies that found a positive association between strategy efficiency and performance (Newton et al., 2020; Star & Rittle-Johnson, 2008; Torbeyns et al., 2006; Wang et al., 2019). Together, this indicates that students who executed a more efficient strategy during problem-solving in a digital learning environment were more likely to have better math performance, as demonstrated by higher posttest scores.

However, pre-solving pause time and the productivity of the initial step were not found to be significant predictors of algebraic knowledge, while controlling for prior knowledge, gifted status, and virtual status. There are several possible explanations for these results. First, since strategy efficiency was included in the model, it is plausible that the direct effects of pre-solving pause time and productivity on algebraic knowledge were diminished. These factors may still matter, but primarily through their contribution to strategy efficiency rather than directly affecting learning outcomes. Another possibility is that these measures may capture different aspects of student performance. While posttest scores reflect students' generalized algebraic knowledge, including conceptual understanding and procedural knowledge, pre-solving pause time and initial productivity represent moment-by-moment behaviors or cognitive processes during problem solving, which may not necessarily translate into learning or long-term outcomes.

Limitations and future directions

This study had several limitations. First, our analyses included only a subset of students who completed both the pretest and posttest. While further exploration indicated no significant demographic differences between these students and those who did not complete both tests, the students who did not complete both tests had significantly lower pretest scores than those of the subset included in our study. Therefore, it is important to acknowledge that our findings may not fully represent the larger student body in other school districts throughout the represented state and the U.S., nor comprehensively capture the nuanced relationship between pre-solving pause time, productivity, and strategy efficiency in the game.

Additionally, while the indirect effect of pre-solving pause time on strategy efficiency through productivity was statistically significant, it was relatively small. This suggests that much of the effect may operate through other cognitive or behavioral processes beyond initial strategy selection. Future research could consider incorporating other metacognitive measures or behavioral indicators to identify mechanisms that better explain why longer pause times are associated with higher strategy efficiency.

Lastly, while we found that pre-solving pause time can serve as an indicator of their planning process, unlike Paquette et al.'s (2014) study, we did not differentiate between purposeful pauses for planning and pauses due to confusion or uncertainty. Some students may have paused because they were unsure how to solve the problem, experiencing mathematics-related frustration or anxiety, or simply mind-wandering, rather than actively engaging in planning (Walczyk et al.,2006). For example, in our data, we found that some students made a mathematically invalid action (e.g., trying to add 7 and 2 before multiplying in $7+2\times5$) as their first step. This warrants further investigation to better understand the cognitive and emotional processes that may underlie students' pause behaviors.

Another possible direction is to experimentally test the effects of pre-solving pause time on productivity and strategy efficiency by encouraging students in a classroom setting, to plan out their steps prior to solving an algebra problem. This approach may be particularly beneficial for students who simply mind-wander during pauses, by encouraging them to use that time to purposefully plan their problem-solving strategy.

Educational relevance and implications statement

The present study examined the relationships among planning, execution, strategy efficiency, and mathematics performance in a digital algebraic learning environment. Findings replicate prior research showing that students who pause longer before solving are more likely to appropriately execute an action based on their planning, which in turn leads to more efficient problem solving and higher mathematical achievement. These findings suggest that students should be given ample time for planning during classroom instruction. Teachers can apply metacognitive strategies like self-reflection to help students become more aware of their thinking processes and improve procedural flexibility. In addition, emphasizing flexibility in teaching problem-solving strategies can provide multiple opportunities for students to learn and enhance their procedural fluency.

Regarding implications for research, although many studies use total reaction time (or response time) alongside accuracy to measure student performance, total reaction time is often skewed, which can lead to incorrect interpretations (Berger & Kiefer, 2021). As an alternative, researchers may consider using percent pre-solving pause time as a more targeted indicator of student cognitive engagement during problem solving.

Conclusions

Procedural fluency is a primary goal in mathematics education, yet many students struggle with selecting appropriate strategies. This study examined how strategic planning and execution related to strategy efficiency in algebraic problem-solving among 7th-grade students using a digital algebraic learning game. Findings revealed that longer pause times before problem-solving were associated with more appropriate execution as well as more efficient strategies, indicating stronger procedural fluency. Moreover, strategy efficiency was positively associated with students' posttest scores assessing algebraic knowledge. These findings highlight the importance of effective strategic planning and execution in enhancing students' mathematical proficiency.

Acknowledgements We would like to thank the teachers and students for their participation; members of the Math, Abstraction, Play, Learning, and Embodiment (MAPLE) Lab for their work; Erik Weitnauer and members of Graspable Math Inc. for programming support; and Neil Heffernan and the ASSISTments Team for their support.

Funding The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education [Grant Number R305A180401 and National Science Foundation [Grant Number 2142984].

Data availability The data are located on the OSF (Open Science Framework) website and can be accessed by filling out the data sharing agreement found on the From Here To There Efficacy Study repository (https://doi.org/10.17605/OSF.IO/R3NF2; https://osf.io/r3nf2).

Declarations

Ethics approval All procedures performed in the study were approved by the Institutional Review Board (IRB). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Conflict of interest Erin Ottmar was a designer and co-developer of From Here to There! The other authors declare no conflicts of interest associated with the publication of this manuscript.

References

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review*, *14*, 243–248. https://doi.org/10.3758/BF03194059

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Berger, A., & Kiefer, M. (2021). Comparison of different response time outlier exclusion methods: A simulation study. *Frontiers in Psychology*, *12*, 675558.

Callan, G. L., & Cleary, T. J. (2019). Examining cyclical phase relations and predictive influences of self-regulated learning processes on mathematics task performance. *Metacognition and Learning*, *14*, 43–63. https://doi.org/10.1007/s11409-019-09191-x

Carlson, M. P., & Bloom, I. (2005). The Cyclic nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics*, *58*, 45–75. https://doi.org/10.1007/s10649-005-0808-x

Chang, K. E., Sung, Y. T., & Lin, S. F. (2006). Computer-assisted learning for mathematical problem solving. *Computers & Education*, *46*(2), 140–151. https://doi.org/10.1016/j.compedu.2004.08.002

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

Chan, J. Y. C., Ottmar, E. R., & Lee, J. E. (2022). Slow down to speed up: Longer pause time before solving problems relates to higher strategy efficiency. *Learning and Individual Differences, 93*, 102109. https://doi.org/10.1016/j.lindif.2021.102109

Common Core State Standards Initiative. (2022). *Common Core State Standards for Mathematics*. https://learning.ccsso.org/wp-content/uploads/2022/11/ADA-Compliant-Math-Standards.pdf

Coppersmith, J. G., & Star, J. R. (2022). A complicated relationship: Examining the relationship between flexible strategy use and accuracy. *Journal of Numerical Cognition*, *8*(3), 382–397. https://doi.org/10.5964/jnc.7601

Decker-Woodrow, L. E., Mason, C. A., Lee, J.-E., Chan, J. Y.-C., Sales, A., Liu, A., & Tu, S. (2023). The impacts of three educational technologies on algebraic understanding in the context of COVID-19. *AERA Open, 9*. https://doi.org/10.1177/23328584231165919

De Smedt, B., Torbeyns, J., Stassens, N., Ghesquiere, P., & Verschaffel, L. (2010). Frequency, efficiency and flexibility of indirect addition in two learning environments. *Learning and Instruction, 20*(3), 205–215. https://doi.org/10.1016/j.learninstruc.2009.02.020

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*, 83–93. https://doi.org/10.1016/j.stueduc.2019.03.001

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*, 139–170. https://doi.org/10.1007/s11409-015-9139-1

Garofalo, J., & Lester, F. K. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for Research in Mathematics Education*, *16*(3), 163–176. https://doi.org/10.2307/748391

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

Gunnarsson, R., Sönnerhed, W. W., & Hernell, B. (2016). Does it help to use mathematically superfluous brackets when teaching the rules for the order of operations? *Educational Studies in Mathematics*, *92*(1), 91–105. https://doi.org/10.1007/S10649-015-9667-2

Hiebert, J. (2013). Conceptual and procedural knowledge: The case of mathematics. *Conceptual and procedural knowledge: The case of mathematics*. Taylor and Francis. https://doi.org/10.4324/9780203063538

Hulse, T., Daigle, M., Manzo, D., Braith, L., Harrison, A., & Ottmar, E. (2019). From Here to There! Elementary: A game-based approach to developing number sense and early algebraic understanding. *Educational Technology Research and Development, 67*(2), 423-441. https://doi.org/10.1007/s11423-019-09653-8

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

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

Lee, J. E., Jindal, A., Patki, S. N., Gurung, A., Norum, R., & Ottmar, E. (2023). A comparison of machine learning algorithms for predicting student performance in an online mathematics game. *Interactive Learning Environments, 32*(9), 5302–5316. https://doi.org/10.1080/10494820.2023.2212726

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

National Council of Teachers of Mathematics [NCTM] (2023). *Procedural* fluency in mathematics. https://www.nctm.org/uploadedFiles/Standards_and_Positions/Position_Statements/PROCEDURAL_FLUENCY.pdf

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

Ottmar, E., Lee, J. E., Vanacore, K., Pradhan, S., Decker-Woodrow, L., & Mason, C. A. (2023). Data from the efficacy study of From Here to There! A dynamic technology for improving algebraic understanding. *Journal of Open Psychology Data, 11*(5), 1–15. https://doi.org/10.5334/jopd.87

Ottmar, E. R, Landy, D., Goldstone, R., & Weitnauer, E. (2015). Getting From Here to There!: Testing the effectiveness of an interactive mathematics intervention embedding perceptual learning. In D. C. Noelle, T. Matlock, R. Dale, C. Jennings, A. Warlaumont, P. P. Maglio, & J. Yoshimi (Eds.), *Proceedings of the annual meeting of the cognitive science society* (pp. 1793–1798). Cognitive Science Society. https://escholarship.org/uc/item/4wj428kw#article_main

Paquette, L., de Carvalho, A. M. J. A., & Baker, R. S. (2014). Towards understanding expert coding of student disengagement in online learning. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society, CogSci 2014* (pp. 1126–1131). The Cognitive Science Society. https://cognitivesciencesociety.org/past-conferences/

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

Rau, M. A., Aleven, V., & Rummel, N. (2017). Making connections among multiple graphical representations of fractions: Sense-making competencies enhance perceptual fluency, but not vice versa. *Instructional Science*, *45*, 331–357.

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

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

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

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

Schulz, A. (2023). Assessing student teachers' procedural fluency and strategic competence in operating and mathematizing with natural and rational numbers. *Journal of Mathematics Teacher Education*, 1–28. https://doi.org/10.1007/s10857-023-09590-7

Simatupang, R., Elvis Napitupulu, E., & Syahputra, E. (2019). Analysis of mathematical problem-solving abilities taught using problem-based learning. *American Journal of Educational Research*, *7*(11), 794–799. https://doi.org/10.12691/education-7-11-6

Star, J. R., Pollack, C., Durkin, K., Rittle-Johnson, B., Lynch, K., Newton, K., & Gogolen, C. (2015). 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

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis. *Journal of Statistical Software*, *59*(5), 1–38.

Torbeyns, J., Verschaffel, L., & Ghesqui`ere, 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. *Learning of Mathematics*, *39*(1), 13–17. https://www.jstor.org/stable/26742004

Verschaffel, L. (2024). Strategy flexibility in mathematics. *ZDM–Mathematics Education, 56*(1), 115-126. https://doi.org/10.1007/s11858-023-01491-6

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

Walczyk, J. J., & Griffith-Ross, D. A. (2006). Time restriction and the linkage between subcomponent efficiency and algebraic inequality success. *Journal of Educational Psychology*, *98*(3), 617. https://doi.org/10.1037/0022-0663.98.3.617

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

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