

Flexibility Before Fluency: Exploring How Flexible Strategies Emerge through Collaborative Problem-Solving

Seiyon Lee, University of Florida, leeseiyon@ufl.edu

Shan Zhang, University of Florida, zhangshan@ufl.edu

Hongming Li, University of Florida, hli3@ufl.edu

Ji-Eun Lee, Singapore University of Technology and Design, jieun_lee@sutd.edu.sg

Avery Closser, University of Florida, avery.closser@ufl.edu

Anthony F. Botelho, University of Florida, abotelho@coe.ufl.edu

Abstract: In K-12 math education, it is important not only to foster *fluency* but also *flexibility* in the problem-solving process. Prior works have focused on measuring learning in terms of efficiency or productivity by examining the extent to which students can optimize their problem-solving paths to involve the fewest number of steps. Limited work has paid attention to the notion of flexibility, or the ability to understand and navigate different paths to the same problem. In this study, we examine how middle school students' use a dynamic notation tool to explore flexible solution paths during collaborative problem solving (CPS) activity. Results show that there are differences in patterns of CPS process between students who demonstrated greater flexibility and those who did not. Specifically, they differed in their engagement with higher-level operations. We discuss the implications for supporting students' flexible problem solving and teachers in orchestrating CPS for similar goals in classrooms.

Introduction

An important goal of mathematics learning is to promote both conceptual and procedural knowledge. Students must not only grasp the underlying concepts, but also apply the appropriate strategies. Prior work has examined the extent to which students can optimize this problem-solving process in an efficient (e.g., Newton et al., 2020) or productive manner (e.g., Lee et al., 2022), suggesting that students reach the goal state in the quickest and the most accurate way possible (Hickendorff et al., 2022). Achieving such *fluency*, however, requires students to have developed some level of *flexibility*, or the ability to understand the relative strengths and shortcomings of different strategies before they take the optimal path for a given context (Star & Rittle-Johnson, 2008). While the importance of flexibility has been well-established (e.g., Blöte et al., 2001; Siegler, 1996), little is known about how students demonstrate procedural flexibility while engaging in problem-solving processes.

Over the past decade, several studies have examined instructional approaches to foster flexibility. Some have focused on explicit instruction, directly teaching students multiple strategies (Rittle-Johnson et al., 2012). Others have used worked examples to help students compare and contrast alternative strategies (Newton et al., 2010; Rüede et al., 2023), or encouraged students to generate and evaluate their own solutions (Xu et al., 2017). However, most of this work has focused on individual learning contexts, overlooking how flexibility may emerge during collaborative problem solving (CPS).

In CPS, learners construct a shared cognitive space through social interactions (Miyake & Kirschner, 2014; Roschelle & Teasley, 1995; Swiecki et al., 2020), supporting one another's learning in ways that teachers may not always be able to do in real time. In this regard, the use of computer-supported collaborative learning (CSCL) environments are particularly relevant for facilitating *flexibility* through CPS as they allow students easily to *explore* a wide range of strategies and *reflect* on them in rapid cycles (Mavrikis et al., 2022), which could be difficult in traditional instruction. Yet, few studied how flexibility may be fostered by leveraging such CSCL affordances.

In this study, we investigate how CPS may support the development of flexibility in K-12 mathematics classrooms. Specifically, we orchestrated CPS for 40 sixth-grade middle school students using Graspable Math (GM), where they used dynamic notational tools to coordinate and jointly solve algebra problems on a shared digital whiteboard. We conducted a fine-grained analysis of students' process data, focusing on their use of different algebraic manipulations within GM (e.g., distributing, factoring, canceling terms), perception of collaboration, math beliefs as well as their overall performance. Specifically, our work is guided by the following research questions (RQs):

- RQ1. How do students' flexibility compare when using graded versus binary scoring approaches?
- RQ2. How do students' behavioral indicators of flexibility during CPS relate to their learning?
- RQ3. How do patterns of algebraic actions differ between higher- and lower-performing groups?

Related Work

Procedural flexibility through CPS

The development of flexibility is a critical yet challenging goal in K–12 mathematics classrooms. Research has shown that students who are flexible problem-solvers tend to have deeper conceptual knowledge and exhibit greater fluency in the problem-solving process (Blöte et al., 2001; Siegler, 1996). While teachers can provide instruction, the effectiveness of the intervention significantly varies within and across contexts, and importantly students often revert to a single, standard solution (Star et al., 2022). Prior work has primarily focused on the specific types of instruction that teachers can provide; however, even well-designed curricula offer limited opportunities for students to explore a range of strategies and solution paths (Hickendorff & van Zanten, 2024). Less attention has been given to the role that students themselves can play in the process (Munaji et al., 2025).

One of such under-explored possibilities is CPS, a socio-cognitive process in which learners take turns on the same task and build a shared understanding, which in turn shapes their subsequent interactions (Clark, 1996). CPS is recognized in K-12 math education as an opportunity for students to articulate mathematical ideas and reasoning (NCTM, 2024; Moses & Cobb, 2001; Mullins et al., 2011;). Studies also found that flexibility is improved when students engage in group work and math discussions (Durkin et al., 2021).

While a substantial body of research has examined CPS, most studies examined students' discourse or other coarse-grained artifacts (e.g., Hu, 2021). These approaches are often limited, either by students' ability to verbalize or demonstrate their knowledge, or by the time and resources required for researchers to qualitatively code such data. Therefore, there is a need for novel tools that not only facilitate CPS activity, but also capture the process at a more granular level, for instance through the automatic tracking and tagging of students' behaviors in the CSCL environment.

Affordances of dynamic notation tools

The emergence of dynamic notation tools helped provide a much-needed affordance. Often embedded in digital whiteboards (e.g., MathSpace, FluidityMath, Graspable Math), students use these tools to interact with numbers and mathematical symbols through specified gesture-actions (e.g., dragging, and tapping). As they are largely analogous to the real dynamics of solving algebra problems, students can focus on identifying the visual cues in the mathematical notations, grasping their meaning, and navigating different variations in the equation (Lindgren & Johnson-Glenberg, 2013; Ottmar et al., 2015). Studies have shown that embodied experiences shape math thinking, and improve learning and engagement (Abrahamson et al., 2020; Goldstone et al., 2017; Landy & Goldstone, 2010; Ottmar & Landy, 2017; Weitnauer et al., 2016).

Dynamic notation tools are well-suited for studying flexibility or CPS more broadly for two reasons. First, they allow students to explore multiple approaches with minimal scaffolding to guide the process. For example, students' actions are immediately followed by real-time visual transformations that provide feedback on their operations. When students attempt to perform an invalid operation, the system simply does not execute it, often signaling this through subtle visual cues or responsive motion. This design helps avoid growing any misconceptions while reinforcing underlying mathematical principles. Combined with interactive features, these affordances enable rapid, low-risk experimentation from which students can learn to discern which strategies may be more appropriate under certain conditions (Komatsu & Jones, 2020). Importantly, all of this process is automatically recorded and stored in the form of log data, which can later be used for post-hoc analyses (Chan et al., 2023). Recent studies have begun using such process data to unpack CPS dynamics, though most have remained descriptive in nature (e.g., Li et al., 2024; Ma et al., 2023).

We leverage one of such tools in *Graspable Math* (GM), where students can encounter and discover algebraic principles by physically moving and rearranging numbers and symbols in a digital whiteboard, much like a mathematical playground. Their interactions are automatically logged as either algebraic or non-algebraic, or as sub-operations of algebraic actions (e.g., *commuting terms*, *canceled fractions*). Prior studies have shown that a gamified version of GM can enhance algebra learning outcomes (Chan et al., 2022; Decker-Woodrow et al., 2023). In this study, we explore the potential of GM to naturally create opportunities for cultivating flexibility through CPS activities.

Methodological approaches to modelling CPS

The CPS process is often complex and difficult to represent in ways that translate into actionable insights for teachers (e.g., determining whether students successfully executed a set of strategies). To model CPS effectively, it is important to account for how each event at a given moment is influenced by what has immediately preceded it or by recent contextual interactions (Ruis et al., 2019). Several methods have been used to reflect the temporal and interdependent nature of CPS processes, such as analyzing the co-occurrence of codes (Shaffer et al., 2017)

or the sequential order of events (Fogel et al., 2021; Swiecki et al., 2019). However, when used in isolation, these methodological approaches risk either overlooking the temporal dynamics of CPS or over-representing micro-sequences that may not be meaningful.

In this sense, Ordered Network Analysis (ONA; Tan et al., 2022) is a promising approach that accounts for the interactive, interdependent, and temporal aspects of CPS. ONA involves constructing *directed* network models, which visualize how patterns of interactions unfold over time. These models are relatively easy and intuitive to interpret and support both visual and statistical comparisons. For example, individual networks can be combined and compared across groups, such as between higher- and lower-performing students. ONA visualizations can also be incorporated into broader CSCL tools that help teachers monitor how students engage in CPS, as partially demonstrated in prior work (Echeverria et al., 2025). In this study, we use ONA for its dual benefits of modelling students' CPS processes in a way that also holds the potential to inform the future design of teacher-facing orchestration tools.

Methods

Study context and procedure

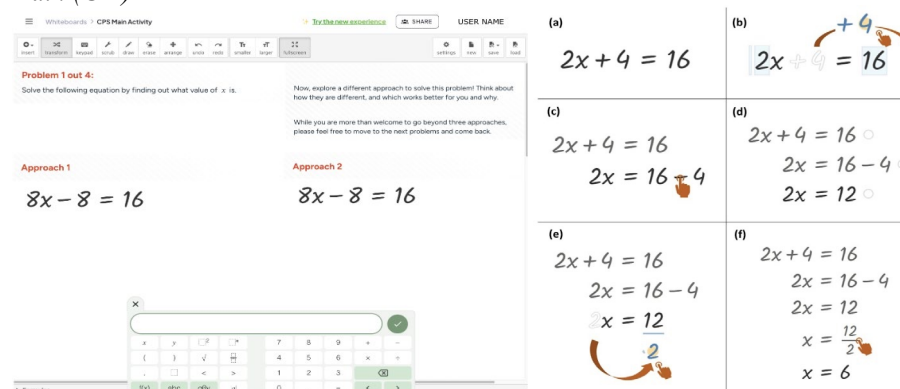
The current study used a recent prototype in Graspable Math (GM), a collaborative whiteboard that enables student pairs to work synchronously on shared problems within a common workspace. Students were seated face-to-face and coordinated their activity through the shared whiteboard, using drag-and-drop gestures to manipulate the digital mathematical representations. This design supports joint engagement, with both students contributing through shared interaction rather than individual control.

Data were collected in a one-week classroom-based study with 43 sixth-grade students from two class periods at a developmental research school in the southern United States. The study was embedded in regular mathematics instruction and conducted over four consecutive days: Day 1 onboarding and a pre-survey on group work perceptions and mathematics beliefs; Day 2 an individual pre-test; Day 3 random pairing for collaborative problem-solving on four problems (Figure 1); and Day 4 an individual post-test and post-survey. All tasks and assessments were conducted in Graspable Math (GM), and surveys were administered via Qualtrics. The study was approved by the local Institutional Review Board.

For analysis, we retained 28 students (14 pairs) after excluding those (and their partners) absent from any pre-test, post-test, or CPS activity. With one missing pre-survey, the final sample included 14 female, 11 male, and 2 non-binary students; 45.8% White, 20.8% Hispanic/Latinx, 16.7% Black, 8.3% Asian, and 8.4% multiracial or other.

Figure 1

Example Problem Layout for CPS activity and Types of Manipulations Possible Through Graspable Math (GM)



Note. For this study, pairs of students solved a set of algebraic equations on the digital whiteboard (left) using gesture-based actions (right; a–f). These manipulations included (b, e) moving variables, coefficients, and constants; (c) tapping to perform basic arithmetic or cancel fractions; and (d) using hidden circles to reposition entire equations.

Measures and data preprocessing

Given the novelty of the study focus, the survey items were informed by and adapted from prior works, such as K-12 mathematics education (e.g., Duran et al., 2014), collaborative learning (e.g., Dewiyanti et al., 2007), and

engagement in digital environments (e.g., Xu & Correia, 2024). The pre-survey included 12 items measuring groupwork attitudes and 7 items assessing mathematics beliefs. The post-survey included 21 items focused on students' perceptions of the collaborative activity, including open-ended questions (e.g., "Can you describe how you and your partner worked together to discover or explore a different approach to solving the problem?"). All items measured on a 5-point Likert-scale. For analysis, composite scores were computed by taking the average from the items within each construct.

Students' flexible problem solving was evaluated through a pre-test, a CPS activity, and a post-test. All assessments were completed within the same Graspable Math (GM) environment to ensure consistency across conditions. Each assessment included three algebraic equation tasks and one ratio task. For this set of problems, students were encouraged to elicit at least two possible solution strategies. For the purposes of this study, we focused on the three algebra tasks, as they were most relevant for examining flexibility. Consistent with the survey measures, performance scores were averaged across the three items. Differences between the pre- and post-tests, as well as between the CPS activity and the post-test, were also computed.

As an alternative indicator of flexibility, we also considered students' behavioral variability during the CPS activity. Specifically, three variance measures were derived: (1) *Variance in Action Types* (VAT), capturing variability in the distribution of algebraic and non-algebraic manipulations (e.g., editing texts); (2) *Variance in Algebraic Actions* (VAA), capturing variability in the types of algebraic operations performed; and (3) *Variance in Key Strategies* (VKS), capturing variability in the use of algebraic strategies that are most relevant to the task (e.g., *AddSubInvert*, *MulDivInvert*, *FractionCancelTerms*; see full list in Table 1). Higher values from these measures indicate greater diversity in students' use of available manipulations, algebraic operations, or key strategies, and vice versa. All of these action and strategy types were automatically derived from GM log data, as the system classifies each user action into predefined categories or codes.

Table 1
Types of Algebraic Actions that are Most Relevant to the CPS Task

Algebraic Action	N (%)	Description
AddSubInvert	2,651 (15.2%)	Move the constant to the other side of the equation. (e.g., $8x + 16 = 0 \rightarrow 8x - 16 = 0$)
MulDivInvert	2,391 (13.7%)	Invert both sides to simplify expressions into fractions. (e.g., $8x = 24 \rightarrow x = \frac{24}{8}$)
Commute	2,118 (12.2%)	Reorder terms in an equation to change their sequence. (e.g., $x8 - 24 = 0 \rightarrow -24 + x8 = 0$)
FractionCancelTerms	1,289 (7.4%)	Cancel fractions to convert expressions into integers. (e.g., $x = \frac{24}{8} \rightarrow x = 3$)
AddSubNumbers	1,267 (7.3%)	Combine or simplify constant terms within the equation. (e.g., $-16 - 8 + 8x = 0 \rightarrow -24 + 8x = 0$)
DistributeIntoBrackets	1,139 (6.5%)	Apply the distributive property to remove brackets. (e.g., $8\text{left}(x - 3)\text{right} = 0 \rightarrow x8 - 24 = 0$)
FactorCommonTerms	1,010 (5.8%)	Factor out the common term shared across expressions. (e.g., $8x - 8 = 16 \rightarrow 8(x - 1) = 16$)
GreatestCommonFactor	773 (4.4%)	Factor out the greatest common factor. (e.g., $-16 + 8x - 8 = 0 \rightarrow (-2 + x)8 - 8 = 0$)
Total	12,638 (100%)	--

Analytic approach

Data analyses addressed the three RQs using complementary quantitative methods. To address RQ1, we examined whether students' flexibility improved after the CPS activity. Specifically, we used two scoring approaches. The first was a graded scale (0–2) that indicated the number of distinct correct solutions students generated (0 = no correct solution, 1 = one correct solution, 2 = two distinct correct solutions). Because the transition from producing one correct solution to generating two represents a qualitatively greater cognitive shift, we also employed a binary measure that distinguished between students who generated two distinct solutions and those who did not (1 = flexible, 0 = not flexible). This binary measure captures conceptual gains that might be underestimated by a linear scale assuming incremental improvement. Differences between pre- and post-tests, as well as between the CPS activity and the post-test, were then computed. Prior to the analyses, the normality of difference scores was assessed using the Shapiro–Wilk test. Based on the results, the appropriate parametric (paired-samples t-test) or

non-parametric (Wilcoxon signed-rank test) was applied. Cohen's d was also calculated to estimate effect sizes, and statistical significance was evaluated at $\alpha = .05$. Afterwards, a graded scale was used for the learning measure in subsequent analyses.

To address RQ2, we conducted Pearson's correlation to examine the relationships among students' learning (e.g., pre-, CPS activity, post- scores), survey measures (i.e. math beliefs, groupwork attitudes, and perceptions of Collaboration), and indicators of behavioral variability (e.g., VAT, VAA, and VKS, described above). Then, Benjamini–Hochberg (BH) correction was applied to control the false discovery rate at $\alpha = .05$.

To address RQ3, we conducted ONA to examine patterns of students' algebraic strategy use during the CPS activity. The ONA algorithm identifies and accumulates connections between the current line of data (in this case, a specific strategy students demonstrated) and preceding lines within a defined window. ONA is particularly useful in contexts where the precise sequence of actions is less important than their temporal overlap, such as when students deviate from a prescribed order of operations. The moving window (or stanza) length was set to 14, meaning that actions within this range were considered temporally related. Alternative window sizes had been tested, and the final selection yielded the most moderate goodness of fit.

ONA was implemented with individual students as the unit of analysis, generating a separate network graph for each participant's sequence of actions within their collaborative whiteboard. This resulted in 28 individual-level networks across two groups (Improved vs Not Improved) based on learning gains (i.e., the difference between pre- and post-test scores). Students with positive gain values were classified as the higher-performing group, comprising 15 of the 28 students, while the remaining 13 students were classified as lower-performing. The overall sample had a mean learning gain of 0.44 ($SD = 0.72$).

Results

RQ1. Effect of CPS on flexibility

We compared students' pre- and post-test flexibility scores using both the graded scale and binary measures as described above. A Shapiro–Wilk test indicated that the graded scores did not significantly deviate from normality ($W = 0.94, p = .051$), whereas the binary scores showed a significant deviation ($W = 0.59, p < .001$). Accordingly, both parametric (t -test) and non-parametric (Wilcoxon W signed-rank) analyses were conducted to ensure robustness.

For the graded scale, both analyses revealed a significant improvement from pre- to post-test, $t(n-1) = 3.09, p = .004$; $W = 42.50, p = .004$, with a medium effect size ($d = 0.54$). In contrast, for the binary measure of flexibility (coded as 1 for flexible and 0 otherwise), the t -test suggested a marginally significant increase, $t(n-1) = 1.97, p = .058$, and the Wilcoxon test yielded a similar trend, $W = 4.00, p = .059$, with a small-to-medium effect size ($d = 0.34$). This finding is consistent with our hypothesis that it is more difficult for students to come up with a second/alternative approach to solve the same problem.

Overall, results suggest that students became more flexible in their problem-solving after engaging in the CPS activity. Improvements were most evident when flexibility was measured by the linear scale, indicating that many students were able to apply multiple solutions more effectively after CPS. The smaller, marginal change in the binary measure aligns with the hypothesis that producing two distinctly different solutions requires a higher level of cognitive demand.

RQ2. Correlations within CPS processes

In the correlation analysis, three associations remained statistically significant after Benjamini–Hochberg (BH) correction (see Table 2). The first was a strong negative correlation between pre-test score and gain from the pre-test to the CPS activity ($r = -0.76, p < .001$). This finding indicates that students with lower initial flexibility tended to show greater relative improvement during collaboration. Similarly, pre-test score was negatively correlated with overall learning gain ($r = -0.69, p < .01$), reflecting the same pattern across a broader timeframe. Finally, overall learning gain was positively correlated with gain during the CPS activity ($r = 0.69, p < .01$), suggesting that students who showed greater improvement during CPS also tended to exhibit larger overall gains. These relationships suggest that changes in flexibility were related to students' starting points and that progress observed during CPS was closely aligned with cumulative learning across the intervention.

Several associations emerged prior to applying the BH correction, which provides additional insight into potential relationships among behavioral and learning measures. Variance in the types of actions (including both algebraic and non-algebraic manipulations) was positively associated with CPS activity score ($r = .45, p < .05$) and with learning gain between pre-test and the CPS activity ($r = .52, p < .01$). These findings suggest that students who engaged in a wider range of interaction types tended to perform better during collaboration and showed

greater improvement from their initial levels of flexibility. Similarly, variance in algebraic actions was positively correlated with post-test performance ($r = .41, p < .05$), indicating that students who exhibited greater variability in their algebraic operations also tended to perform better individually after collaboration. Considering that students engaged in only one day of CPS activity, this pattern supports the idea that behavioral diversity, reflected in variability across action types, may serve as an indicator of a flexible problem-solving process during CPS. Importantly, the diversity of interactions appeared to matter more than whether students explored specific algebraic strategies or whether those actions were directly relevant to the task.

It is also worth noting that students' math beliefs were positively correlated with both their perceptions of collaboration ($r = .49, p < .01$) and their post-test performance ($r = .41, p < .01$), highlighting a potential link between positive mathematical dispositions, collaborative attitudes, and achievement. Moreover, post-test scores were positively associated with CPS activity performance ($r = .40, p < .01$), and CPS activity performance itself was strongly correlated with immediate learning gain from pre-test to CPS activity ($r = .69, p < .01$). Overall, these patterns suggest that students who engaged more flexibly and held stronger beliefs about mathematics and collaboration tended to achieve higher levels of performance both during and after the CPS activity.

Table 2
Correlations Among CPS Measures

Measures	1	2	3	4	5	6	7	8	9	10
1 Variance in Action Types	-									
2 Variance in Algebra Actions	.35	-								
3 Variance in Key Strategies	-.13	-.05	-							
4 Math Belief	-.07	-.14	.12	-						
5 Perceptions of Collaboration	-.10	-.21	.03	.49	-					
6 Pre-Test Score	-.25	.01	.07	.17	.24	-				
7 CPS Activity Score	.45	.32	.03	.25	.09	.21	-			
8 Post-Test Score	.07	.41	-.02	.41	.14	.36	.40	-		
9 Learning Gain (Post – Pre)	.29	.31	-.08	.16	-.12	-.69	.11	.43	-	
10 CPS Activity – Pre	.52	.20	-.04	.01	-.15	-.76	.48	-.06	.69	-

Note. Significant correlations are in bold; bh-corrected ones are highlighted in grey.

RQ3. CPS patterns between groups

To better understand how students engaged in CPS at a finer-grained level, we compared the aggregated mean ONA networks between groups of students who showed improvement in flexibility and those who did not. (Figure 2). In ONA, the node size is proportional to the number of occurrences of that code as a response to other codes in the data, while the color and saturation of the circle within each node are proportional to the number of self-connections for that code. In this regard, we found that the similar sets of codes (or algebraic actions) were large in both groups. For instance, both groups responded most frequently with *MulDivInvert* and *Other*, which were followed by *AddSubInvert*. These reflect codes with more frequent self-connections, or where a code is both what students responded to and what they responded with.

At a high level, the high-performing group exhibited fewer salient connections (with line weight $LW > 0.1$) compared to the low-performing group. This suggests that successful students engaged in a more strategic process, relying on a smaller set of carefully sequenced moves rather than exploring a wider variety of actions. In contrast, the low-performing group's more complex network of transitions indicates exploratory behavior, which may be less consistent in terms of control. These findings imply that effective flexibility in algebraic problem solving may depend less on the number of strategies employed and more on the purposeful coordination of strategies in particular sequences.

More importantly, these groups also exhibited notably distinct patterns in terms of the directionality between key algebraic actions. Given the ways tasks are designed, certain sequences of algebra actions are more logical and sensible than others. For instance, if Invert both sides to simplify expressions into fractions remaining x in one side and constant in others (*MulDivInvert*) and then it naturally follows Cancel fractions to convert expressions into integers (*FractionCancelTerms*).

Results showed that patterns of key algebraic actions mostly overlapped between the higher- and lower-performing groups. For example, both groups demonstrated transitions such as *AddSubInvert* $\rightarrow$ *MulDivInvert*, *AddSubNumbers* $\rightarrow$ *MulDivInvert*, and *MulDivInvert* $\rightarrow$ *FractionCancelTerm*. This suggests that students engaged

in lower-level alternation, such as between addition–subtraction and multiplication–division operations, which are fundamental to the process of systematically restructuring the algebraic equation.

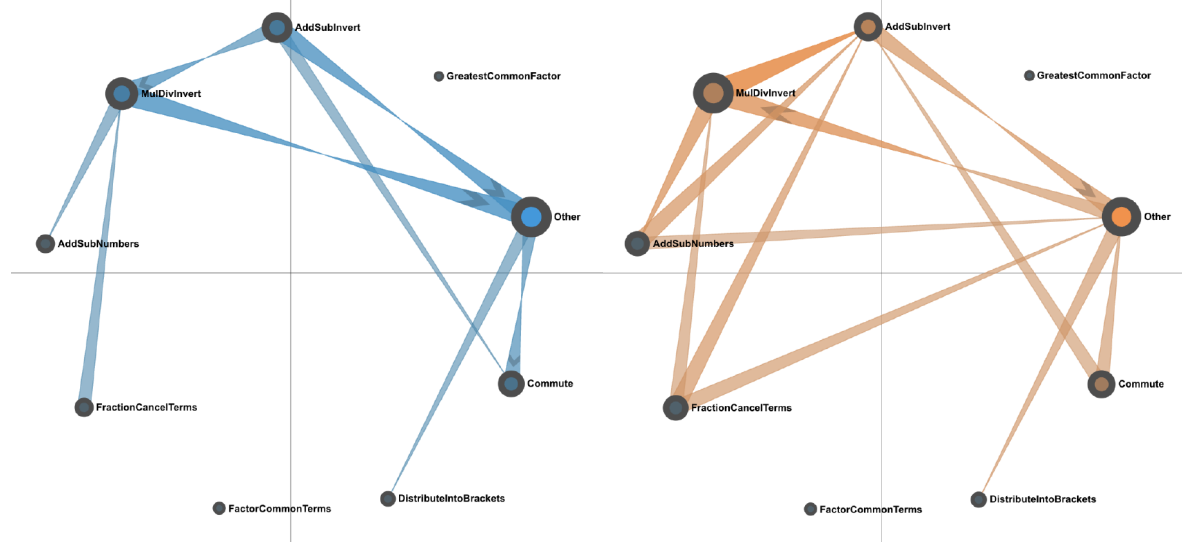
However, the two groups differed in their involvement of *Other*, which represents algebraic actions less directly relevant to solving the given problem. The higher-performing group showed more frequent transitions from *MulDivInvert* → *Other* (LW = .17), indicating that these students often engaged in higher-level operations before exploring alternative approaches. In contrast, the lower-performing group exhibited the opposite transition, *Other* → *MulDivInvert* (LW = .16), suggesting a possible reliance on inversion operations that could complicate the equations, particularly when executed before higher-level transformations had been applied.

As mentioned earlier, a wider range of action sequences was observed for the lower-performing group. For example, the higher-performing group displayed unique connections from *Commute* → *AddSubInvert*, which were absent in the lower-performing group. In contrast, the lower-performing group demonstrated transitions such as *AddSubInvert* → *AddSubNumbers* and *AddSubInvert* → *FractionCancelTerms*, and *Other* → *FractionCancelTerms*. These patterns suggest a tendency to move quickly among related but less coordinated operations.

Overall, the broader and less focused set of connections in the lower-performing group points to a more exploratory, trial-and-error approach to problem solving, whereas the higher-performing group exhibited more deliberate and strategically organized transitions among the same set of algebraic operations.

Figure 2

ONA Networks for a Group of Students Who Improved (Left; Blue) and Those Who Did Not (Right; Orange)



Note. Only the most salient connections (line weights LW > 0.1) are displayed to facilitate interpretability.

Table 3

Most Salient Patterns between High- and Low- Performing Groups (with Line Weight of 0.10 or Above)

No	High-Performing Group			Low-Performing Group		
	From	To	LW	From	To	LW
1	MulDivInvert	Other	0.17	AddSubInvert	MulDivInvert	0.20
2	AddSubInvert	Other	0.16	Other	MulDivInvert	0.16
3	AddSubInvert	MulDivInvert	0.15	AddSubNumbers	MulDivInvert	0.15
4	Other	Commute	0.14	AddSubInvert	Other	0.14
5	AddSubNumbers	MulDivInvert	0.11	AddSubInvert	AddSubNumbers	0.14
6	MulDivInvert	FractionCancelTerms	0.11	AddSubInvert	FractionCancelTerms	0.13
7	Commute	AddSubInvert	0.11	Distrib...Bracket	Other	0.12
7	Distrib...Bracket	Other	0.11	Other	FractionCancelTerms	0.12
8	-	-	-	MulDivInvert	FractionCancelTerms	0.11
9	-	-	-	Other	Commute	0.11

Note. Contrasting patterns between two groups are in bold.

Discussion

This study examined how students' flexibility unfolds during CPS in algebra. Results showed that participation in the CPS activity was associated with improvements in mathematical flexibility, particularly among students who began with lower initial knowledge. Although variance measures and mathematics beliefs did not reach statistical significance within this single instance of CPS, the findings suggest that these factors may still play a meaningful role. Furthermore, the observed gains appeared to depend less on the number of operations performed than on the purposeful coordination of strategies that distinguished higher- from lower-performing students. This pattern aligns with prior research emphasizing that flexible problem solvers are not only capable of executing multiple strategies but can also select and apply them adaptively based on problem structure (Heinze et al., 2009; Krems, 2014; Star & Rittle-Johnson, 2008).

Moreover, high-performing students in this study appeared to engage in high-level thinking, monitoring and regulation of their problem-solving processes, using the digital whiteboard environment to test and refine different strategies, as opposed to playing around with drag-and-drop interactions. In contrast, low-performing students exhibited more reactive and locally driven transitions, either reflecting procedural rigidity or lacking deliberate coordination among algebraic operations. One possible explanation for these differences may relate to affective factors such as mathematics anxiety, which has been shown to impede reasoning and reduce students' willingness to explore alternative strategies despite challenges or unexpected turns of events (Jiang et al., 2020). These differences highlight a competence–performance gap well-documented in the literature in that knowing multiple strategies does not necessarily entail using them flexibly in practice (Hickendorff, 2022).

The observed patterns highlight the importance of designing adaptive supports that can help teachers recognize and respond to students' differing levels of flexibility in real time. Future work could explore the integration of orchestration tools or dashboards that visualize students' problem-solving trajectories and alert teachers to moments of unproductive repetition or disengagement. Such tools could enable timely interventions, particularly for students who may struggle to coordinate strategies or persist through uncertainty.

Limitations and future directions

Several limitations should be noted, which can inform directions for future research. First, this study focused on the design and development of three sets of algebra problems that built upon one another while introducing alternative ways to solve equations. Although the tasks explicitly prompted students to generate alternative approaches, the representational formats used in this study may have constrained the range of strategies students could explore. Relatedly, prior research has shown that the presentation of problems can influence students' strategy use and flexibility (Chan et al., 2022), suggesting that our findings may be somewhat conditional on the representational formats employed in this study. Future work can broaden the range of problem representations so that students are encouraged to explore different strategies in more flexible ways.

Second, this study was limited to observable actions as recorded by the platform. While this feature makes the data well suited for developing tools that can be used in real time, it does not capture students' cognitive and metacognitive processes, which are often latent and invisible but may play a critical role in shaping how students strategically engage in CPS—for example, through self-correction or reflection between moves. In this regard, prior work has combined transmodal analysis with ONA to provide a more contextualized understanding of learning processes (e.g., Borchers et al., 2024). However, further research is needed to examine how students coordinate cognitive and embodied aspects of problem solving within such integrated frameworks.

Last, this study did not examine individual-level ONA networks. Aggregating networks at the group level may have obscured within-group variation in how CPS differentially supported or hindered the development of procedural flexibility. Future research could analyze unit- or participant-level differences to explore how specific peer interactions contribute to individual learning trajectories and the development of flexibility over time.

Conclusion

This study highlights flexibility as a distinct dimension of mathematical problem solving, complementing traditional measures of fluency and efficiency. By examining students' interactions with a dynamic notation tool in a CSCL environment, we found that greater flexibility was linked to higher-level algebraic reasoning during CPS. These findings demonstrate the potential of such tools to both reveal and foster flexible problem-solving processes in K–12 mathematics classrooms. To better support flexibility across different levels of proficiency, it is important to design opportunities for students to explore multiple strategies with appropriate structure and guidance. Equipping teachers with tools to monitor and provide timely, targeted interventions can further promote equitable and inclusive opportunities for all learners to engage in meaningful mathematical exploration.

References

- Abrahamson, D., Nathan, M. J., Williams-Pierce, C., Walkington, C., Ottmar, E. R., Soto, H., & Alibali, M. W. (2020, August). The future of embodied design for mathematics teaching and learning. In *Frontiers in Education* (Vol. 5, p. 147). Frontiers Media SA.
- Blöte, A. W., Van der Burg, E., & Klein, A. S. (2001). Students' flexibility in solving two-digit addition and subtraction problems: Instruction effects. *Journal of educational psychology*, 93(3), 627.
- Borchers, C., Wang, Y., Karumbaiah, S., Ashiq, M., Shaffer, D. W., & Aleven, V. (2024, March). Revealing networks: Understanding effective teacher practices in AI-supported classrooms using transmodal ordered network analysis. In *Proceedings of the 14th Learning Analytics and Knowledge Conference* (pp. 371-381).
- Chan, J. Y. C., Closser, A. H., Smith, H., Lee, J. E., Drzewiecki, K. C., & Ottmar, E. (2023). Grasping patterns of algebraic understanding: Dynamic technology facilitates learning, research, and teaching in mathematics education. In *Mathematical Teaching and Learning: Perspectives on Mathematical Minds in the Elementary and Middle School Years* (pp. 207-225). Cham: Springer International Publishing.
- Chan, J. Y. C., Ottmar, E. R., Smith, H., & Closser, A. H. (2022). Variables versus numbers: Effects of symbols and algebraic knowledge on students' problem-solving strategies. *Contemporary Educational Psychology*, 71, 102114.
- 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.
- Clark, H. H. (1996). *Using language*. Cambridge university press.
- 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, 23328584231165919.
- Durkin, K., Rittle-Johnson, B., Star, J. R., & Loehr, A. (2023). Comparing and discussing multiple strategies: An approach to improving algebra instruction. *The Journal of Experimental Education*, 91(1), 1-19.
- Echeverria, V., Zhao, L., Alfredo, R., Milesi, M. E., Jin, Y., Abel, S., ... & Martinez-Maldonado, R. (2025, April). TeamVision: An AI-powered Learning Analytics System for Supporting Reflection in Team-based Healthcare Simulation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1-22).
- Fogel, A., Swiecki, Z., Marquart, C., Cai, Z., Wang, Y., Brohinsky, J., ... & Shaffer, D. W. (2021, January). Directed epistemic network analysis. In *International Conference on Quantitative Ethnography* (pp. 122-136). Cham: Springer International Publishing.
- Goldstone, R. L., Marghetis, T., Weitnauer, E., Ottmar, E. R., & Landy, D. (2017). Adapting perception, action, and technology for mathematical reasoning. *Current Directions in Psychological Science*, 26(5), 434-441.
- Heinze, A., Star, J.R. & Verschaffel, L. Flexible and adaptive use of strategies and representations in mathematics education. *ZDM Mathematics Education* 41, 535–540 (2009).
- Hickendorff, M., & van Zanten, M. (2024). Opportunity to learn flexible and adaptive strategy use in current and past Dutch mathematics textbooks. *Mathematical Thinking and Learning*, 1-18.
- Hickendorff, M., McMullen, J., & Verschaffel, L. (2022). Mathematical flexibility: Theoretical, methodological, and educational considerations. *Journal of Numerical Cognition*, 8(3), 326-334.
- Hickendorff, M. (2022). Flexibility and adaptivity in arithmetic strategy use: What children know and what they show. *Journal of Numerical Cognition*, 8(3), 367-381.
- Hu, L. (2021). Turn-usurping in dialogic collaborative problem solving. In *Proceedings of the 15th International Conference of the Learning Sciences-ICLS 2021*. International Society of the Learning Sciences.
- Jiang, R., Liu, R. D., Star, J., Zhen, R., Wang, J., Hong, W., ... & Fu, X. (2021). How mathematics anxiety affects students' inflexible perseverance in mathematics problem-solving: Examining the mediating role of cognitive reflection. *British Journal of Educational Psychology*, 91(1), 237-260.
- Komatsu, K., & Jones, K. (2020). Interplay between paper-and-pencil activity and dynamic-geometry-environment use during generalisation and proving. *Digital Experiences in Mathematics Education*, 6(2), 123-143.
- Krems, J. F. (2014). Cognitive flexibility and complex problem solving. In *Complex problem solving* (pp. 201-218). Psychology Press.
- Landy, D., & Goldstone, R. L. (2010). Proximity and precedence in arithmetic. *Quarterly journal of experimental psychology*, 63(10), 1953-1968.

- Li, H., Zhang, S., Lee, S., Lee, J. E., Zhong, Z., Weitnauer, E., & Botelho, A. F. (2024). Math in motion: Analyzing real-time student collaboration in computer-supported learning environments. In *Proceedings of the 17th International Conference on Educational Data Mining* (pp. 533-541).
- Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational researcher*, 42(8), 445-452.
- Loes, C. N., & Pascarella, E. T. (2017). Collaborative learning and critical thinking: Testing the link. *The Journal of Higher Education*, 88(5), 726-753.
- Ma, Y., Zhang, H., Ni, L., & Zhou, D. (2023). Identifying collaborative problem-solver profiles based on collaborative processing time, actions and skills on a computer-based task. *International Journal of Computer-Supported Collaborative Learning*, 18(4), 465-488.
- Mavrikis, M., Rummel, N., Wiedmann, M., Loibl, K., & Holmes, W. (2022). Combining exploratory learning with structured practice educational technologies to foster both conceptual and procedural fractions knowledge. *Educational technology research and development*, 70(3), 691-712.
- Melcer, E. F., & Isbister, K. (2016, May). Bridging the physical divide: a design framework for embodied learning games and simulations. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2225-2233).
- Miyake, N., & Kirschner, P. A. (2014). The social and interactive dimensions of collaborative learning.
- Munaji, M., Rohaeti, T., Mutadi, M., Sumliyah, S., & Kodirun, K. (2025). A literature review of flexibility in interactive mathematics classrooms: the role of teachers and students. *Journal of Education and Learning (EduLearn)*, 19(2), 597-605.
- National Council of Teachers of Mathematics (2024). Principles to Actions: Ensuring Mathematical Success for All. 10th Anniversary Edition. NCTM.
- 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.
- Ottmar, E., & Landy, D. (2017). Concreteness fading of algebraic instruction: Effects on learning. *Journal of the Learning Sciences*, 26(1), 51-78.
- 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 *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 37).
- Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology*, 101(3), 529.
- Roschelle, J., & Teasley, S. D. (1995, August). The construction of shared knowledge in collaborative problem solving. In *Computer supported collaborative learning* (pp. 69-97). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Ruis, A., Siebert-Evenstone, A., Pozen, R., Eagan, B., & Shaffer, D. W. (2019). Finding common ground: A method for measuring recent temporal context in analyses of complex, collaborative thinking.
- Star, J. R., & Newton, K. J. (2009). The nature and development of experts' strategy flexibility for solving equations. *ZDM*, 41(5), 557-567.
- Swiecki, Z., Lian, Z., Ruis, A., & Shaffer, D. W. (2019). Does order matter? Investigating sequential and cotermporal models of collaboration.
- Swiecki, Z., Ruis, A. R., Farrell, C., & Shaffer, D. W. (2020). Assessing individual contributions to collaborative problem solving: A network analysis approach. *Computers in Human Behavior*, 104, 105876.
- Tan, Y., Ruis, A. R., Marquart, C., Cai, Z., Knowles, M. A., & Shaffer, D. W. (2022, October). Ordered network analysis. In *International Conference on Quantitative Ethnography* (pp. 101-116). Cham: Springer Nature Switzerland.
- Tan, Y., Swiecki, Z., Ruis, A. R., & Shaffer, D. (2024). Epistemic network analysis and ordered network analysis in learning analytics. In *Learning Analytics Methods and Tutorials: A Practical Guide Using R* (pp. 569-636). Cham: Springer Nature Switzerland.
- Weitnauer, E., Landy, D., & Ottmar, E. (2016, December). Graspable math: Towards dynamic algebra notations that support learners better than paper. In *2016 Future Technologies Conference (FTC)* (pp. 406-414). IEEE.