The Impact of the Mindsets Mathematics Program on Student Engagement, Ability to Reason, and Repeated Reasoning

Nicholas A. Gage

November 2025

WestEd.org

Suggested citation: Gage, N. (2025). *The impact of the Mindsets mathematics program on student engagement, ability to reason, and repeated reasoning.* WestEd.

WestEd is a nonpartisan, nonprofit organization that aims to improve the lives of children and adults at all ages of learning and development. We do this by addressing challenges in education and human development, increasing opportunity, and helping build communities where all can thrive. WestEd staff conduct and apply research, provide technical assistance, and support professional learning. We work with early learning educators, classroom teachers, local and state leaders, and policymakers at all levels of government.

For more information, visit WestEd.org. For regular updates on research, free resources, solutions, and job postings from WestEd, subscribe to the E-Bulletin, our semimonthly e-newsletter, at WestEd.org/subscribe.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

Executive Summary

Mindsets is an educational company that developed the Mindsets mathematics program—a hands-on, challenge-based curriculum designed to boost engagement and equity. The program teaches grade-level mathematics through real-world, collaborative projects while developing students' proficiency with the Standards for Mathematical Practice.

Mindsets partnered with a large school district in California to implement a 5-week Mindsets mathematics program, specifically created to deepen middle school students' engagement and reasoning skills through collaborative, hands-on challenges rooted in real-world contexts. Grounded in the Standards for Mathematical Practice (SMP), particularly SMP 2: Reason Abstractly and Quantitatively and SMP 8: Repeated Reasoning, the program emphasizes mathematical thinking, communication, and relevance. Teachers receive professional development and real-time support to implement daily multi-hour challenges aligned with grade-level standards and the California Mathematics Framework.

The 5-week implementation for grades 6–8 featured themed weeks and included mixed-grade groupings to support equitable access to grade-level content and prior-year skill development. WestEd conducted an independent evaluation of the program's impact on student engagement and mathematical ability to reason.

Using cumulative link mixed models, the analysis revealed

- A **large, statistically significant increase in engagement** from the beginning to the end of the program ($d \approx 1.11$), with engagement peaking during the middle of the program.
- A **significant improvement in students' mathematical reasoning** (SMP 2), with a large pre–post effect ($d \approx 1.20$) and a moderate-to-large increase across all three time points ($d \approx 0.78$).
- A **significant improvement in students' repeated reasoning** (SMP 8), with a large pre–post effect ($d \approx 1.19$) and a moderate-to-large increase across all three time points ($d \approx 0.77$).

These findings suggest that Mindsets meaningfully improves both how students participate in mathematical learning and their ability to reason abstractly and quantitatively—key outcomes for long-term success in mathematics.

Program Overview

Mindsets is a mathematics program centered on collaborative, hands-on challenges that immerse students in real-world scenarios. Each challenge is designed to deepen students' understanding of the Standards for Mathematical Practice (SMPs) while making mathematics engaging, relevant, and meaningful. As part of the Mindsets program, teachers receive targeted professional development and real-time implementation support, equipping them to lead students through complex mathematical investigations rooted in authentic contexts.

The program includes a broad selection of multi-hour challenges, enabling students to engage in a new challenge each day across a 5-week experience. Each challenge begins by introducing a compelling context that situates mathematics in real life. Students then work collaboratively through a progression of increasingly complex problems that highlight core elements of a particular SMP. Each challenge concludes with a culminating hands-on activity—such as making pancakes, creating bracelets, or setting prices for a dog-walking business—that reinforces the mathematical concepts explored.

Mindsets' challenges are closely aligned with the vision of the California Mathematics Framework and particular SMPs; the program is not California-specific. Students engage with grade-level content standards through investigations built on big mathematical ideas.

Study Design

In the fall of 2025, WestEd evaluated the impact of the Mindsets math program on a group of middle school students in grades 6–8 from the San Diego Unified School District in California, following their participation in a five-week summer program earlier that year. In the implementation of the Mindsets program for the study sample, challenges were organized around the following weekly themes: Building Community, Food Week, Movement, Shark Tank (business and financial literacy), and The Amazing Race (competition). Most sessions were conducted in mixed-grade groupings, allowing older students to revisit foundational content while supporting younger students' preparation for middle school rigor. The contextual, problem-based design is intended to allow students to engage meaningfully and reduce the stigma often associated with remediation.

In the 2025 summer session, Mindsets emphasized SMP 2: Ability to Reason and SMP 8: Repeated Reasoning (also known as SMP 8: Ability to Analyze) because they were observable and teachable within a short time period. Based on Mindsets' chosen SMPs, WestEd developed the following research questions:

1. What was the pre–post impact on student engagement?
2. How did student engagement change across the three time points?
3. What was the pre–post impact on students' ability to reason and repeated reasoning in mathematics?
4. How did students' ability to reason and repeated reasoning in mathematics change across the three time points?

Study Sample

Data were collected from middle school students in grades 6–8 attending a school-sponsored program in the San Diego Unified School District in California. All Mindsets mathematics program data for this study were collected by Mindsets and provided to the WestEd team in two datasets: an engagement-focused dataset and an SMP dataset with both abstract reasoning and a repeated reasoning dataset. Each dataset included student observations by time (beginning of the program [pre], middle of the program, and end of the program [post]) and by challenge name.

There were two challenges at the beginning, two challenges in the middle, and three challenges at the end. However, the SMP data was only available for one of the middle challenges and one of the end challenges. For the engagement dataset, there were 1,163 observations at the beginning, 1,091 observations at the middle, and 819 at the end. For the SMP dataset, there were 889 observations at the beginning, 380 at the middle, and 763 at the end. The engagement file also included a classroom indicator for our analyses. There were 232 unique classrooms in the engagement file.

Study Measures

Engagement Scale

The WestEd study team analyzed student engagement by developing and using an engagement scale that reflects students' progress and participation in each Mindsets challenge, captured along a 5-point ordinal scale. While specific definitions were not standardized across contexts, the progression represents increasing levels of task completion and cognitive engagement:

- **Started** (1) indicates students began the challenge but may not have sustained engagement or made significant progress beyond the initial steps.
- **Underway** (2) suggests students moved beyond the starting phase and engaged with core aspects of the challenge, possibly in a partially complete or fragmented way.
- **Nearly** (3) reflects substantial progress toward completing the challenge, with evidence of sustained effort and deeper involvement, but with some components remaining unfinished.
- **Completed** (4) represents full participation in the challenge, with students engaging in all required components and demonstrating persistence through to the intended learning outcomes.
- **Completed Bonus** (5) indicates students not only completed the core challenge but also extended their engagement through optional or enrichment tasks, suggesting high levels of motivation and self-directed learning.

WestEd used this scale to provide insights about students' task-level engagement and as a proxy for measuring depth of participation across time points.

Assessments of Student Reasoning

The WestEd study team analyzed data from Mindsets sources, which assessed student performance in relation to SMP 2: Ability to Reason and SMP 8: Repeated Reasoning, respectively. Specifically, Mindsets provided data profiles that assessed student SMP strengths and areas for growth, as well as reports on student engagement and SMP proficiency. All measures were developed by Mindsets and were completed by students within the Mindsets platform. Scoring was automated within the Mindsets platform.

Assessment of Standard for Mathematical Practice 2: Ability to Reason

This assessment measured students' ability to reason both abstractly and quantitatively—a core goal of SMP 2, part of the Common Core State Standards for Mathematics. Students begin with a real-world scenario and are asked to interpret the context mathematically, using quantities and relationships to formulate a solution pathway. Through hands-on, collaborative problem-solving, students move between representing the situation symbolically and interpreting what their mathematical results mean in context.

The structure of the assessment supports multiple entry points, allowing students at different grade levels and readiness levels to participate meaningfully. Students deconstruct the problem, identify relevant quantities, and explore how mathematical operations model a real-world situation. The culminating task requires students to justify their reasoning and critique the reasoning of others, reinforcing both conceptual understanding and mathematical communication.

The assessment activity is embedded in a theme relevant to students' lived experiences—such as financial literacy or community service—ensuring that reasoning is always anchored in purpose. This approach supports equitable access to rigorous mathematical thinking while cultivating students' confidence and identity as problem solvers.

The WestEd study team created a 4-point ordinal scale from the following benchmark categories: Not Met Benchmark = 1, Nearly Met Benchmark = 2, Met Benchmark = 3, Exceeded Benchmark = 4. Benchmark categories were assigned within the Mindsets platform. Benchmarks were created by Mindsets following psychometric analyses of their internal data.

Assessment of Standard for Mathematical Practices 8: Repeated Reasoning

This assessment measured students' ability to look for and express regularity in repeated reasoning—a core goal of SMP 8, part of the Common Core State Standards for Mathematics. Students engage with patterns in procedures or problem contexts and are asked to generalize those patterns to develop efficient strategies and make predictions. As students iterate through calculations or analyze models, they begin to notice consistencies, articulate general rules, and evaluate the reasonableness of intermediate steps.

The structure of the assessment supports multiple entry points, allowing students at different grade levels and readiness levels to engage meaningfully. Students are encouraged to identify repeated steps, analyze emerging patterns, and formulate conjectures that can be applied to new situations. The culminating task requires students to articulate these generalizations and justify their applicability, reinforcing both mathematical fluency and metacognitive reflection.

The assessment activity is completed on the Mindsets platform and embedded in a theme connected to students' lived experiences—such as planning a savings goal or tracking data trends over time—ensuring that repeated reasoning is purposeful and grounded in authentic contexts. This approach fosters equity by inviting all students to participate in rigorous mathematical thinking and builds their capacity to reason flexibly and efficiently.

The study team created a 4-point ordinal scale from the following benchmark categories: Not Met Benchmark = 1, Nearly Met Benchmark = 2, Met Benchmark = 3, Exceeded Benchmark = 4.

Data Analysis

As the first step of our analysis, we estimated a cumulative link mixed model (CLMM) to evaluate whether student engagement and students' ability to reason and repeated reasoning increased from pre- to post-Mindsets implementation. All three outcome variables, Engagement Scale, Ability to Reason Benchmarks, and Repeated Reasoning Benchmarks, were treated as an ordinal factor reflecting increasing levels of engagement (1 = Started to 5 = Completed Bonus) or reflecting improvement in performance related to the respective SMP (1 = Not Met Benchmark to 4 = Exceeds Benchmark). The predictor of interest was time, which was coded as a binary factor indicating whether a score was recorded during the beginning or end time periods. To account for the nested structure of the data, we included random intercepts for the classroom and challenge.[1] The models were fit using the CLMM function from the *ordinal* R package (Christensen, 2023) with a logit link.

To examine linear growth in student engagement, ability to reason, and repeated reasoning across the study periods, we again estimated a series of CLMM with both outcomes treated as ordinal. The time variable was treated as an ordered factor with three levels: beginning, middle, and end. Polynomial contrasts were automatically applied, producing orthogonal linear (*time.L*) and quadratic (*time.Q*) terms that test for overall trends. Random intercepts were included for both the classroom and the challenge to account for nesting.

[1] The Ability to Reason and the Repeated Reasoning datasets did not include the classroom indicator; therefore, these models only included a random intercept for challenge.

Results

Increase in Student Engagement From Pre- to Post-Intervention

Results indicated a statistically significant increase in engagement from pre- to post-intervention, $b = 2.01$, $SE = 0.26$, $z = 7.74$, $p < .001$ (see Table 1). This positive effect suggests that students were substantially more likely to report higher levels of engagement following the intervention. **The odds of being in a higher engagement category at post-intervention were approximately 7.47 times greater than at pre-intervention** (odds ratio = exp(2.011) ≈ 7.47). Random intercept variance was substantial across classrooms ($\sigma^2 = 2.06$), but negligible across challenges, suggesting that most between-group variability was due to classroom-level effects.

To facilitate interpretation of the intervention's magnitude, we converted the estimated ordinal regression coefficient for the pre–post effect into an approximate standardized mean difference (*d*). While engagement was modeled as an ordinal outcome, we used a well-established method to approximate the *d* from the cumulative logit coefficient by dividing the log-odds estimate by 1.81, the approximate standard deviation of the standard logistic distribution (Chinn, 2000). This method is appropriate when the ordinal scale has multiple levels and approximately equal intervals.

The estimated coefficient for the effect of time (post vs. pre) was β = 2.01 ($SE = 0.26$), indicating a statistically significant increase in engagement following the intervention. Using the conversion formula:

$$SMD \approx \frac{2.01}{1.81} \approx 1.11$$

This result corresponds to a very large effect size (Cohen, 1988), suggesting that students' engagement significantly increased following the intervention. This standardized estimate provides a useful common metric for comparing results across studies and outcome types, particularly when engagement outcomes are reported using different measurement scales.

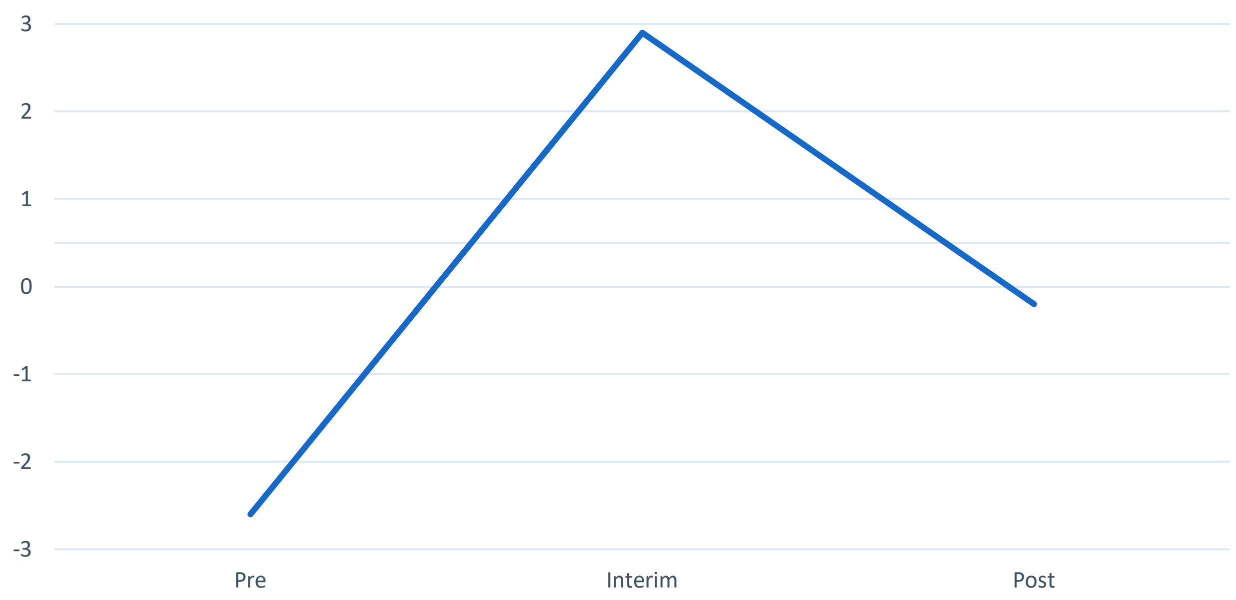
Table 1. Cumulative Link Mixed Model Predicting Engagement Stage

Predictor	Estimate	SE	z	p-value	Odds Ratio
Intercept 1|2	–2.89	0.19	–14.90	< .001	–
Intercept 2|3	1.24	0.18	7.02	< .001	–
Intercept 3|4	1.83	0.18	10.15	< .001	–
Intercept 4|5	2.75	0.19	14.78	< .001	–
Post (vs. pre)	**2.01**	**0.26**	**7.74**	**< .001**	**7.47**

***Note*. Random intercept *SD* for classroom = 1.435; for challenge = 0.000, *n* = 1,982.**

Next, we explored engagement change across time. This model revealed a **significant positive linear trend across time** ($b = 1.24$, $SE = 0.17$, $z = 7.33$, $p < .001$), **indicating that engagement generally increased from pre to post**. However, the significant negative quadratic effect ($b = -1.46$, $SE = 0.17$, $z = -8.59$, $p < .001$) suggests that **engagement peaked at Interim and declined slightly by post** (see Figure 1). This nonlinear pattern is visualized in the plot, which shows the highest latent engagement at the Interim phase, followed by a moderate drop at post. Random intercept variance was substantial across classrooms ($SD = 1.32$) but negligible across challenges.

Figure 1. Engagement Changes Across All Three Time Points



Increase in Students’ Ability to Reason From Pre- to Post-Intervention

The CLMM revealed a significant effect of time, $b = 2.16$, $SE = 0.86$, $z = 2.51$, $p = .012$, indicating that **students were more likely to score in a higher Ability to Reason benchmark category at post compared to pre** (see Table 2). The estimated standard deviation of the random intercept for challenge was 0.69, suggesting moderate variability across different challenges. These findings suggest a statistically significant and practically meaningful improvement in student performance following program implementation. Again, to aid interpretation, the log-odds effect was converted to an approximate *d* using Chinn’s (2000) method, yielding an estimated *d* of 1.20, representing a large effect.

Table 2. Ordinal Regression Results for Achievement Level (Pre–Post)

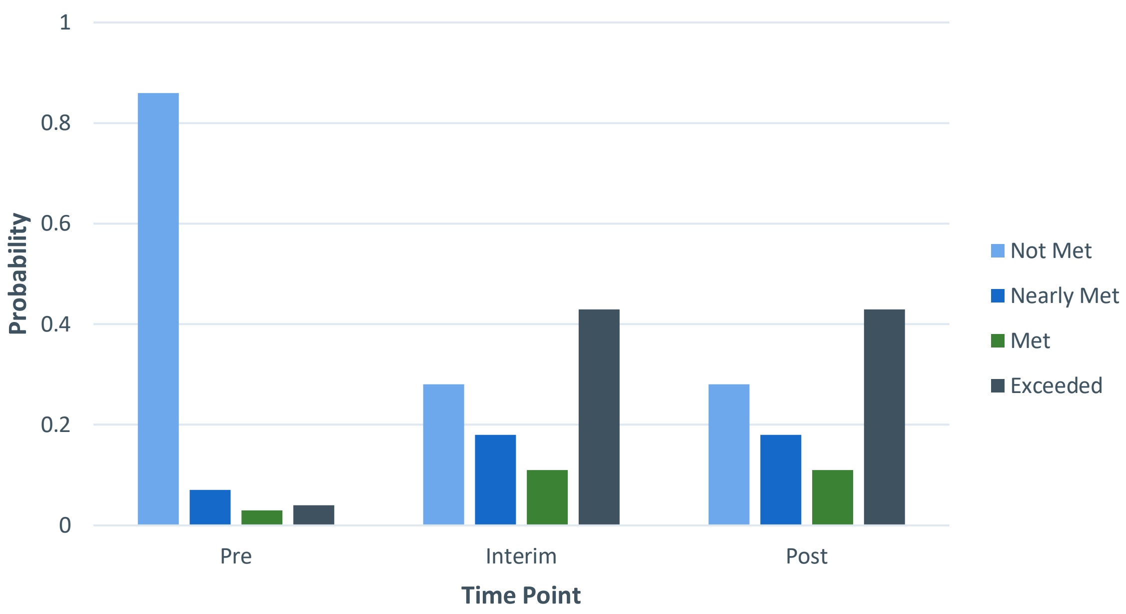
Predictor	Estimate	Standard Error	*z* value	*p* value
Fixed Effects				
timePost	2.164	0.862	2.51	.012 *
Random Effects				
challenge_name (Intercept SD)	0.695	—	—	—

Next, we explored the Ability to Reason benchmark change across time. The CLMM revealed a **significant linear effect of time on achievement level**, $b = 1.41$, $SE = 0.51$, $z = 2.77$, $p = .006$, **indicating that students’ Ability to Reason benchmark levels improved across the time points** (See Table 3). The quadratic term was non-significant ($b = –0.47$, $SE = 0.56$, $z = –0.84$, $p = .40$), suggesting no strong evidence of a non-linear trend (e.g., leveling off or decline at post). This is visualized in Figure 2. The estimated variance in random intercepts for challenge was modest ($SD = 0.58$), indicating **some variability in achievement levels across different activities**. Once again, we used the Chinn (2000) approximation for cumulative logit models and found an estimated $d \approx 0.78$. This suggests a moderate-to-large effect of time on achievement level, consistent with Cohen’s interpretation guidelines.

Table 3. Ordinal Regression Results for Achievement Level

Predictor	Estimate	Standard Error	*z* value	*p* value
Fixed Effects				
time.L	1.411	0.510	2.77	.006 **
time.Q	–0.474	0.562	–0.84	.399
Random Effects				
challenge (Intercept SD)	0.579	—	—	—

Figure 2. Ability to Reason Benchmark Level Change Across All Three Time Points



Increase in Students' Repeated Reasoning from Pre- to Post-Intervention

The CLMM revealed a significant effect of time, $b = 2.16$, $SE = 0.86$, $z = 2.51$, $p = .012$, indicating that **students were more likely to score in a higher Repeated Reasoning benchmark category at post compared to pre** (see Table 4). The estimated standard deviation of the random intercept for challenge was 0.69, suggesting moderate variability across different challenges. **These findings suggest a statistically significant and practically meaningful improvement in**

student performance following program implementation. To aid interpretation, the log-odds effect was converted to an approximate *d* using Chinn's (2000) method, yielding an estimated *d* of 1.19, representing a large effect.

Table 4. Ordinal Regression Results for Repeated Reasoning Benchmark Level (Pre–Post)

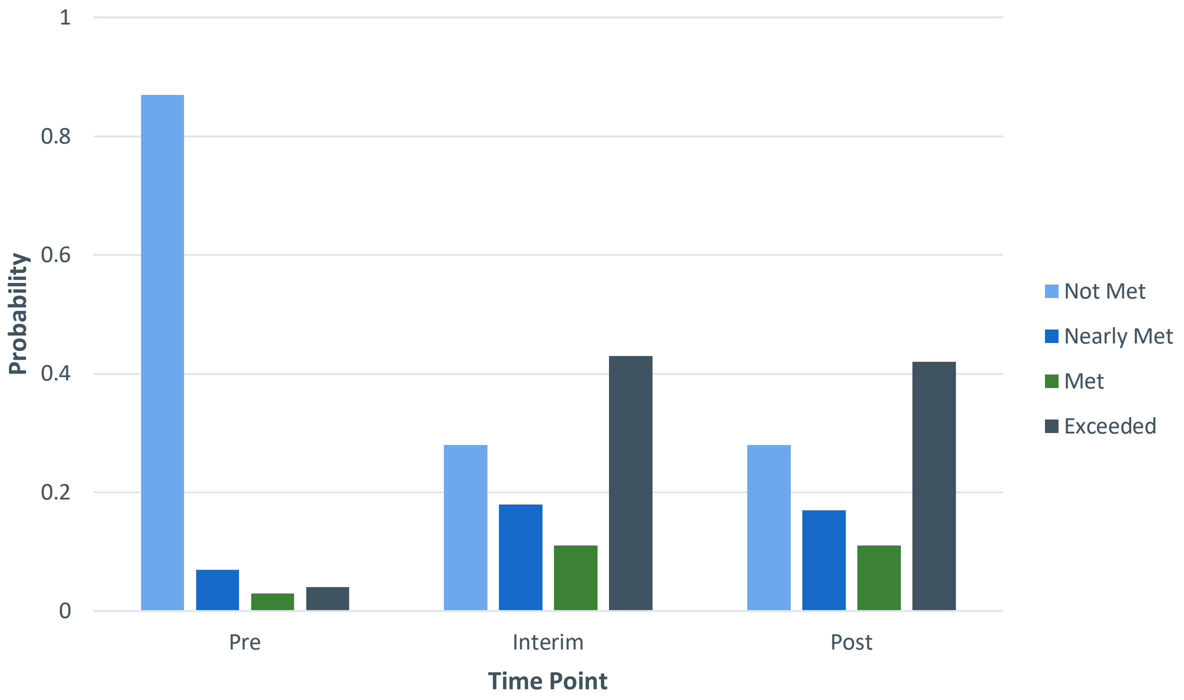
Predictor	Estimate	Std. Error	*z* value	*p* value
Fixed Effects				
timePost	2.164	0.862	2.51	.012 *
Random Effects				
challenge_name (Intercept SD)	0.695	—	—	—

Finally, we explored the Repeated Reasoning benchmark change across time. The CLMM revealed a **significant linear effect of time on achievement level**, $b = 1.41$, $SE = 0.51$, $z = 2.77$, $p = .006$, indicating that **students' Repeated Reasoning benchmark levels improved across the time points** (see Table 5). The quadratic term was non-significant ($b = –0.47$, $SE = 0.56$, $z = –0.84$,
$p = .399$), suggesting no strong evidence of a non-linear trend (e.g., leveling off or decline at post). This is visualized in Figure 3. The estimated variance in random intercepts for challenge was modest ($SD = 0.58$), indicating **some variability in achievement levels across different activities**. Once again, we used the Chinn (2000) approximation for cumulative logit models and found an estimated $d \approx 0.77$. This suggests a moderate-to-large effect of time on achievement level, consistent with Cohen's interpretation guidelines.

Table 5. Ordinal Regression Results for Repeated Reasoning Benchmark Level

Predictor	Estimate	Standard Error	*z* value	*p* value
Fixed Effects				
time.L	1.411	0.510	2.77	.006 **
time.Q	–0.474	0.562	–0.84	.399
Random Effects				
challenge_name (Intercept SD)	0.579	—	—	—

Figure 3. Repeated Reasoning Benchmark Level Change Across All Three Time Points



Discussion and Conclusion

The findings from this study provide strong evidence that the Mindsets program supports measurable improvements in student engagement and mathematical reasoning over a 5-week intervention. The program's emphasis on collaborative, hands-on challenges grounded in real-world contexts appears to be particularly effective in fostering both cognitive and motivational dimensions of mathematics learning.

First, student engagement increased significantly over the course of the program. The effect from pre to post was large ($d \approx 1.11$), and growth patterns revealed a peak at the program's midpoint, followed by a slight decline toward the end. This nonlinear pattern is common in short-term interventions, where novelty and momentum may taper slightly as students approach completion (Fredricks et al., 2004). Nonetheless, the sustained improvement across time highlights the importance of designing learning environments that are both structured and responsive to students' interests and lived experiences (Eccles & Roeser, 2011).

Second, students' ability to reason abstractly and quantitatively—captured through the SMP 2-aligned assessment—also improved meaningfully. A large pre–post effect ($d \approx 1.20$) and a moderate-to-large linear increase across all three time points ($d \approx 0.78$) suggest that students not only participated more actively but also developed deeper conceptual understanding. This finding reinforces the value of embedding mathematical reasoning in authentic tasks, where students construct, apply, and communicate mathematical ideas (Boaler, 2016; National Council of Teachers of Mathematics [NCTM], 2014).

Finally, students' ability to recognize and express regularity in repeated reasoning—measured through the SMP 8-aligned assessment—showed substantial growth. A large pre–post effect ($d \approx 1.19$) and a moderate-to-large linear gain across all three time points ($d \approx 0.77$) indicate that students increasingly generalized mathematical patterns and refined their strategies over time. This finding underscores the importance of designing tasks that allow students to identify structure, reflect on processes, and build mathematical fluency through repetition and pattern recognition (NCTM, 2014; Sfard, 2008).

The Mindsets model demonstrates that when students are provided with access to meaningful tasks, high expectations, and supportive teacher facilitation, they are capable of engaging deeply with mathematical content regardless of their prior achievement levels. The program's success with mixed-grade groupings further supports the argument that equitable mathematics instruction must prioritize context-rich learning opportunities and reject deficit-oriented remediation models (National Academies of Sciences, Engineering, and Medicine, 2018).

In conclusion, this study offers promising evidence that programs like Mindsets can significantly enhance students' engagement and reasoning—two critical drivers of long-term mathematics achievement. As education systems seek to recover from the lingering effects of the pandemic and support students' mathematical growth, especially in the middle grades, Mindsets offers a model worthy of broader consideration and replication.

References

Boaler, J. (2016). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages, and innovative teaching*. Jossey-Bass.

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine, 19*(22), 3127–3131. https://doi.org/10.1002/1097-0258(20001130)19:22<3127::AID-SIM784>3.0.CO;2-M

Christensen, R. (2023). *Ordinal: Regression models for ordinal data*. R package version 2023.12-4.1, https://CRAN.R-project.org/package=ordinal

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.

Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence, 21*(1), 225–241.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109.

National Academies of Sciences, Engineering, and Medicine. (2018). *English learners in STEM subjects: Transforming classrooms, schools, and lives*. The National Academies Press.

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*.

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing.* Cambridge University Press. https://doi.org/10.1017/CBO9780511499944

WestEd.org