MANAGEMENT BRIEFING

KAITLab

KAIT Accelerates Mathematical Mastery: Evidence and the Case for AI-Powered Math Intervention Across Every Level of Learner

AUTHORS	Dr. Joseph Davis Chief Mathematics & AI Learning Officer Beau Scott Executive Director of Client Success
Organization	KAITLab
Date	May 2026

EXECUTIVE SUMMARY

America's math performance crisis is not a knowledge problem; it is a visibility problem. New district-level data released in Sean Reardon's Educational Opportunity Project at Stanford makes the scale undeniable: **70% of school districts lost ground in math between 2015 and 2025. Only 30% improved.** In reading, the numbers are worse: **83% of districts declined.** A decade of curriculum investment, edtech purchasing, and instructional reform has produced a field where backward movement is the norm, not the exception. The tools in common use cannot see how students think. They record the final answer. They do not record the reasoning that produced it. Until that visibility gap is closed, no curriculum, tutor, or software will deliver a significant change in outcomes. We know what students need: personalized practice at the right level of challenge, immediate corrective feedback, and consistent engagement with their own work shown in their own hand. What has been missing is a scalable mechanism to deliver it.

KAIT (Key AI Training) is that mechanism. Developed by KAITLab, KAIT combines smartpen-captured handwriting with machine learning to create an AI-native training loop that surfaces each student's precise gap, adapts to their zone of proximal development, and gives teachers real-time insight they could not previously access.

This briefing presents the current evidence base, including a controlled classroom comparison in Poway Unified School District, emerging pilot data from the ReGeneration Schools network in Chicago, and the research literature that explains why KAIT's design works. In this report, you will find results that are directionally strong, mechanistically sound, and consistent with a growing body of educational science.

The headline findings are clear: **30 minutes of KAIT training per week, added to existing practice with no other changes, produced 2.5× the expected growth in math proficiency** in a third-grade classroom with a comparison group. Students arriving two or more grade levels behind dropped from 25% to 8% of the class. Students at or above grade level surged from 8% to 46%.

KAIT is not a tutoring program, and it is not remediation software. It is an intervention system designed to do three things simultaneously: catch up students who have fallen behind, cement and accelerate mastery for students who are on standard, and extend the performance of students who are already excelling. The data across every site where KAIT has been deployed confirms what the research predicts: **the students who need it most benefit the most, and the students who are already strong go further still.** In Ras Al Khaimah, a school of students entering deeply below grade level produced an IRT gain of +0.583; a school of students already in the Proficient band gained an additional +0.245. The intervention works across the full ability range because it is calibrated to each student's individual zone, not to a group average.

CONTENTS OF THIS DOCUMENT

I. THE PROBLEM WE ARE SOLVING

A National Math Emergency

The 2024 National Assessment of Educational Progress (NAEP) showed that only 26% of 4th graders and 28% of 8th graders performed at or above proficiency in mathematics. The post-pandemic recovery in reading has meaningfully outpaced recovery in math. Chronic underperformance in mathematics is not distributed evenly: it clusters in schools serving students in poverty, students of color, and students with interrupted learning.

26% / 28%	**National proficiency signal** Only 26% of 4th graders and 28% of 8th graders performed at or above proficiency in mathematics.

School systems have responded with increased instructional time, curricular investments, and a wave of edtech purchasing, much of it generic, most of it unproven. Spending has not equated to results. The gap between what teachers teach and what students actually learn remains largely invisible until a standardized test reveals it, often too late in the year for meaningful intervention.

The Invisible Gap

Most classroom instruction produces a fundamental information asymmetry: teachers deliver content to groups; learning happens (or doesn't) individually. A teacher with 25 students and 50 minutes per day cannot observe each child's mathematical reasoning in real time. They see outputs, right or wrong answers, but not the cognitive process behind them. KAIT was built to close that gap.

II. HOW KAIT WORKS: THE MECHANISM

Understanding why KAIT produces results requires understanding its design logic. KAIT is not a digital worksheet, a video tutorial, or a chatbot. It is a training system built on three interlocking mechanisms, each grounded in established educational research.

1 Handwriting as the Learning Act	**2** Zone of Proximal Development	**3** Showing the Work Creates the Data

Mechanism 1: Handwriting as the Learning Act

KAIT requires students to show their work physically, with a smartpen on paper. This is not an interface preference; it is a pedagogical imperative. Students are not staring at another screen. They are writing, thinking, and working with their hands on real paper, an experience that is both cognitively superior and a meaningful break from the screen saturation that defines most children's school days. A substantial body of cognitive science research demonstrates that handwriting engages deeper memory encoding than typing or tapping (Mueller & Oppenheimer, 2014). When students write out mathematical processes by hand, they activate the motor system, visual attention, and working memory simultaneously, a combination that strengthens the neural pathways associated with conceptual understanding (James & Englehardt, 2012). The smartpen captures every stroke invisibly, turning the oldest learning tool, pencil and paper, into a precision data instrument.

The superiority of handwriting for activating multiple neural systems is reinforced by new learning analytics studies using digital pens in Japanese classrooms. Tonosaki et al. (2024) examined 94 junior-high students and showed that handwriting-process features such as total strokes, pauses, and eraser use varied significantly by problem type and directly correlated with performance, evidence that pen-stroke data capture authentic cognitive engagement rather than surface speed. Similarly, Okayama et al. (2024) found that indicators derived from daily handwritten math logs (writing velocity, active-pen time, and erasure ratio) show statistically meaningful correlations with knowledge and comprehension-level proficiency. Together, these studies validate the measurement fidelity of digital handwriting logs in reflecting true mathematical understanding.

Applied to KAIT, this evidence confirms that smart-pen stroke data can serve as a diagnostic mirror of reasoning accuracy and effort, complementing computer-based assessments rather than replacing them. As larger populations of U.S. and international KAIT users contribute anonymized pencast data, the predictive precision of our handwriting analytics will strengthen further, allowing researchers and teachers to observe fine-grained learning trajectories that traditional keyboards and touch screens cannot reveal.

Mechanism 2: Zone of Proximal Development

Vygotsky's (1978) Zone of Proximal Development (ZPD) is one of the most validated frameworks in learning science: students learn best when challenged just beyond their current capability, with scaffolding available. The problem is that identifying each student's ZPD in a classroom of 25 has always required expert diagnosis and time. KAIT automates this.

As students work through training sessions, KAIT's AI engine, using Item Response Theory (IRT) scoring, continuously estimates each learner's current ability level and selects the next problem at precisely the right difficulty (Linden, 1997). No two students are on the same path. Every session is calibrated to the individual.

In simple terms, IRT constructs a probability curve describing the likelihood that a learner at ability θ will solve a specific item, i, correctly. This ensures that every new task a student receives is optimally matched to their current competence, just beyond their comfort zone but within reach for growth.

Recent research in Japan demonstrates the same principle extended to digital handwriting analytics. Nakamura et al. (2023) integrated pen-stroke data from note-taking within the STACK system, color-coding writing speed and stagnation points, and verified that difficulty and learner-ability parameters estimated with IRT closely mirrored students' actual written reasoning.

Mechanism 3: Showing the Work Creates the Data

When a student types a correct answer, the system knows the answer was correct. When a student writes out their work with a smartpen, the system captures the reasoning process, stroke patterns, corrections, sequence, and pacing. This is the data source that makes KAIT different from every other adaptive math platform. The AI is trained not just on outcomes but on process, allowing it to distinguish between a student who got the right answer correctly and one who got there through a flawed shortcut.

This process-level signal is what teachers receive through the KAIT dashboard: insight into mathematical thinking that observation alone cannot replicate at scale.

New empirical work in Japan using systems like STACK and the LEAF handwriting framework illustrates how stroke-level records can reveal a student's thinking process in ways that parallel KAIT's infrastructure. When pen-stroke data are analyzed chronologically, they encode tempo, pause lengths, and corrections that map into "productive struggle", the moment-by-moment cognitive pattern of effort that builds mastery.

Studies by Tonosaki et al. (2024) and Okayama et al. (2024) confirmed that students with greater stroke persistence and longer on-task writing time consistently demonstrated higher math proficiency. These findings support KAIT's claim that pencast data are not simply digital traces, they are evidence of reasoning quality.

Integrating these principles with KAIT's IRT framework creates an unmatched visibility tool for teachers and families. Each stroke's timing and revision feeds the ability model, closing the long-standing "visibility gap" in mathematics education. Where traditional adaptive software sees only answers, KAIT sees the thinking that produced them. As deployment scales across districts and homes, the handwriting dataset will form a research-grade corpus for continuous validity testing, expanding support for schools and families and further narrowing achievement gaps that no other tool has been able to address to date.

III. THE EVIDENCE: WHAT WE KNOW NOW

Study A: Poway Unified School District, Grade 3 (2024–2026)

Design: One KAIT classroom (Class A) compared to three non-KAIT classrooms (Classes B, C, D) within the same school. No other instructional changes were made. KAIT was used twice per week, 15 minutes per session, a total of approximately 30 minutes per week.

Measurement: Third-party diagnostic assessment (i-Ready) administered in Fall 2025 and Winter 2026.

Results 2025–26 Cohort:

	Students At/Above Grade Level Fall 2025	Students At/Above Grade Level Winter 2026	Change	Students 2+ Levels Behind Fall 2025	Students 2+ Levels Behind Winter 2026	Change
Class A (KAIT)	8%	46%	+38 pts	25%	8%	−17 pts
Class B	17%	17%	0 pts	21%	8%	−13 pts
Class C	0%	12%	+12 pts	20%	8%	−12 pts
Class D	23%	38%	+15 pts	12%	4%	−8 pts

2.5×

Poway Unified growth multiple
Class A's 38-point gain in students at or above grade level was 2.5 times the i-Ready national norm for meaningful growth between diagnostic windows (approximately 15 points).

Key finding: Class A's 38-point gain in students at or above grade level was 2.5 times the i-Ready national norm for meaningful growth between diagnostic windows (approximately 15 points). The next highest gain in the school was 15 points. Class A also produced the largest absolute reduction in severely behind students.

Spring 2025 Cohort (same school, prior year):

The 2024–2025 KAIT pilot in the same Poway Unified school produced parallel findings: the proportion of KAIT students performing above grade level increased by 375%, while students performing two or more grade levels below standard decreased from 36% to 4%. Three comparison classrooms not using KAIT did not replicate these gains.

375% Above Grade Level: Poway Unified KAIT Class Source: Spring 2025 Cohort paragraph. Increase in students performing above grade level during Spring 2025 pilot (prior year cohort).	**36%→4%** Severely Behind: Poway Unified KAIT Class Students two or more grade levels below standard. No comparison class matched this reduction.

Study B: ReGeneration Schools network Pilot, Chicago (2025–2026)

KAITLab is currently running an active partnership with the ReGeneration Schools network in Chicago, spanning three CICS campuses: Washington Park (lead implementation site), Basil (in progress), and Avalon (onboarding). Implementation is staggered by design: lead sites generate the replication blueprint that accelerates activation across the network.

School	Status	Students Enrolled	Assessments Complete	AI Training Sessions
CICS Washington Park	Lead Site	57	57 / 57 (100%)	349
CICS Basil	In Progress	69	35 / 69 (50.7%)	0
CICS Avalon	Not Started	59	0 / 59 (0%)	0

At CICS Washington Park, where full implementation is active:

- 81.7% of enrolled students at the lead site are actively engaged in AI training sessions
- 349 cumulative training sessions completed across the cohort
- 7.1 average training sessions per active student
- 27 standards-level mastery milestones achieved
- +0.308 average IRT ability gain across the active cohort
- 8 students moved from the Developing band into Proficient; Emerging learners dropped from 13 to 5

Individual learner profiles illustrate the platform's precision. In one documented case, a student completed 30 training sessions, mastered 11 of 54 standards, and posted a +27.5 point gain in a single standards domain from 55.1% to 82.6% in whole number quotients.

The platform identified the productive intervention target; the teacher received the signal; the learning happened.

+0.308	**Cohort IRT Ability Gain: CICS Washington Park** Average IRT ability gain across the active student cohort. IRT score of 0 represents grade-level expectation.

Study B2: Academy of Saint Elizabeth, New Jersey

The pattern documented at CICS Washington Park **replicates** at the Academy of Saint Elizabeth in New Jersey. KAIT training produced **a +0.150 average IRT ability gain across the student cohort,** statistically comparable to a 3-to-5 point RIT gain on NWEA MAP, a meaningful shift in a single implementation window. More striking: **the share of students in the lowest "Emerging" proficiency band fell by 67%.**

A parallel result appeared at CICS Irving Park (Chicago), where Emerging learners dropped by 50% over the same implementation window. Across every site where KAIT has been deployed with consistent usage, the pattern is the same: students in the lowest band move up, and measurable gains appear on third-party assessments.

+0.150 Academy of Saint Elizabeth average IRT ability gain	**67%** Academy of Saint Elizabeth Emerging proficiency band reduction	**50%** CICS Irving Park Emerging learners reduction

Study C: International Validation: Ras Al Khaimah, United Arab Emirates (2026)

The most detailed international evidence comes from a district-level deployment in Ras Al Khaimah, UAE where two schools, with two very different student populations, carried out a 12-week pilot of KAIT. The results are the strongest cross-population validation KAITLab has produced to date.

Design and setting: International Secondary Khuzam and British School Khuzam, both serving 8th-grade mathematics. 116 students were tracked from the first diagnostic to the latest training. A comparison group of 50 students without KAIT training ran alongside 20 students with KAIT training for the within-school performance comparison. No curriculum changes were introduced.

Headline results:

- +21.3% average performance increase for AI-trained students over the 12-week window, compared to peers without training
- +0.385 IRT ability shift across the full 116-student cohort, from IRT 0.281 to 0.666, crossing the Developing-to-Proficient boundary. Comparable to a 7–10 point RIT gain on NWEA MAP.
- +40.5% diagnostic-to-exam improvement at International Secondary Khuzam, from offline diagnostic to the latest exam
- −21 Emerging learners across the cohort (48 down to 27, a 44% reduction), with 20 students advancing into the Developing band
- KAIT-trained students: +26% above baseline at test versus +18% for the comparison group, both starting from the same baseline

The replication signal: International Secondary Khuzam's students entered well below grade level (IRT −0.825) and produced the larger absolute gain (+0.583). British School Khuzam's students entered already in the Proficient band (IRT 1.062) and still climbed further into Advanced (+0.245). The platform is not a low-floor intervention. It produces measurable growth across the full ability range. That cross-population replication in an international setting, across two distinct school cultures, with no curriculum changes, is the most compelling structural evidence in KAITLab's current evidence base.

Resilience, not just averages: The RAK data also surfaces a finding that aggregate statistics typically obscure. When new chapter content was introduced mid-window, both trained and untrained students experienced a performance dip. The critical difference: KAIT-trained students dropped less and recovered faster. At test, the trained group finished at +26% above baseline; the comparison group finished at +18%. The AI training pathway is consolidating durable skill, not producing short-lived gains that dissolve under instructional disruption.

Standards-level gains: The five highest-growth standards in the **RAK cohort posted IRT gains between +1.571 and +1.943, roughly equivalent to a student moving up an entire proficiency band on a single standard.** The broadest reach was in Geometric Transformations (+1.608, 26 students), a notoriously difficult 8th-grade concept. Decomposing Fractions and Multiplying Multiples of 10 also showed gains in the foundational fluency skills that predict long-term algebraic readiness.

+0.385	**IRT Ability Shift: Ras Al Khaimah, 116 Students** Cohort moved from IRT 0.281 to 0.666, crossing the Developing-to-Proficient boundary. Comparable to a 7–10 point RIT gain on NWEA MAP.

IV. THE RESEARCH FOUNDATION

KAIT's design did not emerge from a technology roadmap. It emerged from the research literature on how humans learn mathematics. The following evidence base informs the platform's core architecture.

The Handwriting Advantage in Mathematics

The act of writing out mathematical work, rather than typing or selecting answers, activates more cognitive resources and produces stronger retention. Research by James & Engelhardt (2012) and Mueller & Oppenheimer (2014) established that the motor engagement of handwriting creates deeper memory traces than typing. In mathematics specifically, writing out procedures forces students to sequence steps, which builds procedural fluency alongside conceptual understanding.

Adaptive Difficulty and Motivation

Students disengage when work is too easy or too hard. Self-determination theory (Deci & Ryan, 2000) and flow research (Csikszentmihalyi, 1988) converge on the same point: optimal challenge is the engine of intrinsic motivation. By keeping each student in their productive zone, KAIT sustains engagement without the frustration that drives avoidance.

Spaced Practice and Retrieval

KAIT sessions are structured to revisit prior standards at intervals aligned with spaced retrieval research. Students are not asked to master a topic in a single sitting; they return to it across sessions, each time with slightly greater complexity. The spacing effect, one of the most replicated findings in cognitive psychology, ensures practice converts to durable long-term memory rather than surface-level performance (Cepeda et al., 2006).

Immediate Feedback and Error Detection

A student who practices a procedure incorrectly and receives no feedback practices the error to fluency. KAIT surfaces errors at the point of production, not days later on a graded paper. This aligns with decades of feedback research showing that immediate, specific, and corrective feedback accelerates learning dramatically over delayed feedback (Wisniewski et al., 2020).

The Power of Showing Your Work

“Show your work” is one of the oldest instructions in mathematics education and one of the most cognitively important. When students externalize their thinking on paper, they are not merely recording a procedure; they are constructing understanding (Bransford, 2000). Writing out each step forces a student to sequence, monitor, and self-correct in real time. The act of stepping in and out of mental math, holding a value in working memory while writing an intermediate result, then returning to the mental calculation, is a high-order cognitive exercise that builds both procedural fluency and conceptual flexibility simultaneously (Sfard, 1991).

KAIT makes this visible. Because every stroke of the smartpen is captured, the platform can observe not just whether the answer is right but whether the student’s written process reflects sound mathematical reasoning. A student who skips steps or writes them in the wrong order is flagged, not for punishment, but for targeted re-engagement at the precise conceptual moment where the thinking broke down. This is what “show your work” has always promised. KAIT is the first platform capable of delivering on it at scale.

Procedural Practice, Mastery, and Automaticity

Mathematics education research has long wrestled with the relationship between conceptual understanding and procedural fluency. The most rigorous synthesis of this evidence arrives at the same conclusion: they are not in competition (National Mathematics Advisory Panel, 2008). Procedural practice, rules-based, structured, repetitive at the right level of challenge, builds the automaticity that frees working memory for higher-order reasoning. A student who must consciously retrieve the multiplication algorithm cannot simultaneously attend to the logic of a multi-step word problem. Automaticity is not rote memorization for its own sake; it is the cognitive foundation that makes complex mathematical thinking possible (Ericsson et al., 1993).

KAIT’s training structure is explicitly designed around this principle. Each session presents problems that are rules-based and procedurally grounded at the student’s current level, building toward mastery through deliberate repetition before advancing difficulty. The IRT engine ensures that students are not pushed to the next standard before consolidating the current one. The result is not just improved scores, it is faster learning, deeper retention, and the kind of mathematical confidence that transfers across domains and grades.

V. EPISTEMIC HONESTY: WHAT THIS EVIDENCE DOES AND DOES NOT SHOW

We are committed to honest representation of the evidence base. The following is what we can and cannot claim at this stage of KAITLab's development.

What the evidence supports:

- In a controlled classroom comparison at Poway Unified, the KAIT classroom outperformed three non-KAIT comparison classrooms on a third-party diagnostic by a substantial margin, producing 2.5 times expected growth in proficiency gains.
- A prior-year cohort at the same school showed a 375% increase in above-grade-level performance and a drop from 36% to 4% in students severely behind.
- At CICS Washington Park, students in active KAIT training gained an average of +0.308 IRT ability points, with measurable band-level movement from Emerging to Developing to Proficient.
- KAIT's design is aligned with well-established learning science: handwriting advantage, zone of proximal development, spaced retrieval, and immediate feedback.

What we are not yet claiming:

- We do not have a randomized controlled trial. Student assignment to classrooms in these settings was not random.
- We are not asserting that KAIT alone is responsible for all observed gains. Instructional delivery, classroom climate, and prior student preparation are not controlled variables.
- We are a growing company with a limited number of partner schools. We are building the evidence base, not resting on it.

We believe transparency about these limitations is itself a signal of integrity. We are not positioning this as a finished proof; we are inviting partners into the evidence-building process itself, because that is how trust between schools and educational technology companies should be built.

VI. IMPLICATIONS FOR SCHOOLS AND DISTRICTS

What Implementation Looks Like

The Poway Unified results were achieved with 30 minutes of KAIT training per week, two 15-minute sessions, with no changes to existing curriculum, pedagogy, or teacher assignment. KAIT is not a replacement for great teaching. It is a targeted intervention that runs alongside the teacher's work, catching up students who have fallen behind, deepening and cementing mastery for students who are on standard, building the speed and automaticity that make higher-level mathematics accessible, and accelerating the performance of students who are already high-achieving. This is what makes KAIT different from a tutoring program or a practice platform: it is designed to move every student forward, not just the ones who are struggling.

30 min	**Implementation load** The Poway Unified results were achieved with 30 minutes of KAIT training per week, two 15-minute sessions, with no changes to existing curriculum, pedagogy, or teacher assignment.

The CICS Washington Park pilot adds a further dimension: the educator-adoption data shows that when teachers engage with KAIT's data layer, averaging 87.3 training reviews per active teacher, the platform's impact on student learning is amplified. The insight KAIT surfaces is only as valuable as the teacher's ability to act on it. This is why KAITLab invests in educator onboarding, not just student deployment.

The Equity Dimension

Gains were not concentrated among students near the proficiency threshold. The Poway Unified data shows movement at both ends of the distribution simultaneously: **the highest-performing students advanced further while the lowest-performing students closed the largest absolute gaps.** This pattern, described in the intervention literature as *equitable acceleration*, is among the most difficult outcomes to produce through conventional instructional approaches and is rarely observed in single-variable comparisons of this duration.

What Scales

The ReGeneration Schools network partnership is already demonstrating what network-level scaling looks like: one lead site reaches full implementation; the data from that site becomes the blueprint for the next. This is the flywheel. As each school completes baseline assessment and activates training, the network's aggregate evidence base grows, making the case for continued investment stronger with every semester.

KAIT Is Complementary, Not Competing

A district evaluating KAIT is not being asked to rip and replace. The mathematics landscape is full of genuinely useful products: iReady, Magma Math, Eureka, Illustrative Mathematics, Khan Academy, and many others. Teachers have built routines around them. Districts have invested in them. KAIT does not compete with any of them. It is the pencast layer those products have never been able to provide, because that layer requires a sensor they do not have. A teacher using iReady still teaches with paper and pencil. KAIT is what finally lets teachers, parents, and AI engines see what happens there. The budgetary conversation is additive: KAIT sits alongside the curriculum line and the adaptive-practice line, and it compounds the return on both.

VII. THE INVITATION

The mathematics performance crisis is not waiting for a perfect study. The students currently one, two, and three grade levels behind in math are in classrooms right now. The evidence we have is early, honest, and encouraging. The research that explains why KAIT works has been accumulating for decades.

We are not asking for faith in a promise. We are looking for real partners, schools, districts, charter networks, private schools, and families who want more for their children, who are ready to bring KAIT into their classrooms, homes, and communities and build the evidence together.

KAITLab is actively partnering with schools, districts, charter networks, private schools, and families committed to solving the math performance crisis. Our partners span traditional public schools, urban charter networks, suburban districts, independent schools, and parents who want a serious learning tool for their children at home. We bring the platform, the onboarding, the data infrastructure, and the evidence to date. Our partners bring the students, the teachers, and the community. That is what a real partnership looks like, not a vendor relationship, not a pilot that expires, a shared commitment to getting math right for every child.

That is the partnership we are building, with schools that are serious, families that are ready, and educators who believe their students deserve better. We would be honored to build it with you.

ABOUT THE AUTHORS

Dr. Joseph Davis holds a doctorate from Harvard University. A former classroom math teacher, principal, and school superintendent, Joe serves as Chief Mathematics Officer at KAITLab with a belief that great teaching and intelligent technology are not in competition; they are partners. He leads KAITLab's instructional design, educator development, and product vision.

Beau Scott serves as the Executive Director of Customer Success for KAITLab Beau's career has had a deep focus on elementary mathematics and STEM learning. He holds a B.A. in Elementary Education and an M.S. in Educational Leadership from Purdue University. His experience includes serving as the Elementary Math and Science Specialist for the Indiana Department of Education, where he authored the State's Math and Science Frameworks.

SELECTED REFERENCES

The following works are cited and/or inform the evidence presented in this briefing. Additional references available upon request.

Bransford, J. (Ed.). (2000). How People Learn: Brain, Mind, Experience, and School: Expanded Edition. National Academies Press.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practices in verbal recall tasks: A review and quantitative synthesis. Psychological Bulletin, 132(3), 354.

Csikszentmihalyi, M. (1988). The flow experience and its significance for human psychology. Optimal Experience: Psychological Studies of Flow in Consciousness, 15-35.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry, 11(4), 227-268.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in acquisition of expert performance. Psychological rReview, 100(3), 363.

James, K. H., & Englehardt, L. (2012). The effects of handwriting experience on functional brain development in pre-literate children. Trends in Neuroscience and Education, 1(1), 32-42.

Linden, W. J. v. d. (1997). Handbook of Modern Item Response Theory (W. J. v. d. Linden & R. K. Hambleton, Eds.). Springer.

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop notetaking. Psychological Science, 25(6), 1159-1168.

Nakamura, F. K., & Ito, K. (2023). Analysis of STACK answer data using pen-stroke data from a calculation notebook and item response theory. *International Journal of Emerging Technologies in Learning (Online)*, *18*(1), 279.

National Center for Educational Statistics. (2024). NAEP 2024 Mathematics Assessment Results. U.S. Department of Education.

National Mathematics Advisory Panel. (2008). Foundations for success. U.S. Department of Education.

Okayama, L. Y., Takii, K., & Ogata, H. (2024). Identifying key indicators of proficiency in junior high math: Roles of daily handwriting learning logs. *International Conference on Computers in Education*.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. Educational Studies in Mathematics, 22(1), 1-36.

Tonosaki, T. S., Hamada, S., Horikoshi, I., & Ogata, H. (2024). Toward contextualized handwriting process analysis: Comparison between problem types in math. *International Conference on Computers in Education*.

Vygotsky, L. S. (1978). Mind in Society: Development of Higher Psychological Processes (M. Cole, Ed.; M. Cole, Trans.). Harvard University Press.

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. Frontiers in Psychology, 10.