A Management Briefing · Vol. I, No. 3

The Hand-Mind Connection

How the keyboard turn has short-circuited handwriting and mathematics education.

Why American mathematics achievement began falling in 2013, why the migration of student work from paper to glass is implicated, and why recovering the stroke-level record — not removing the computer — is the correction the evidence supports.

AUTHORS

Dr. Joseph Davis
President and Chief Mathematics Officer, KAITLab

Dr. Joseph Wise
Executive Chair and Co-Founder, KAITLab

July 2026 | KAITLab | Vol. I, No. 3

ABOUT THIS PAPER

This is the third briefing in the KAITLab management series. The first, Learning (and Teaching) Mathematics in the Age of AI (Vol. I, No. 1), argued that the mathematics crisis is at root a visibility crisis and that the smartpen is the necessary infrastructure for closing it. The second, KAIT Accelerates Mathematical Mastery (Vol. I, No. 2), presented the evidence base from Poway Unified, the ReGeneration Schools network, the Academy of Saint Elizabeth, and Ras Al Khaimah.

This paper takes up the question those two documents deliberately left open: why did mathematics achievement in the United States begin to decline more than a decade ago, and why has a generation of investment in instructional technology failed to reverse it?

Our answer is uncomfortable for the field and, we think, correct in its main outline. Between roughly 2012 and 2020, American schools carried out the largest uncontrolled change to the physical act of doing schoolwork in the modern history of public education. Student work moved from paper to glass. The keyboard, the touchscreen, and the answer box replaced the pencil and the page as the default surface on which children produce mathematics. That change was procured as infrastructure. It was never evaluated as pedagogy. And it removed, at national scale, two things at once: the motor act by which mathematical procedures are encoded in the developing brain, and the physical record by which a teacher could see how a student was reasoning.

The first loss is cognitive. The second is evidentiary. They compound.

This briefing is written for three audiences: district and school leaders deciding how their students will produce mathematical work; investors and partners evaluating the category; and parents who suspect something has gone wrong and have not been given a clear account of what.

Contents

EXECUTIVE SUMMARY

> ***The decline did not begin in March 2020. Any explanation that starts with the pandemic is explaining the acceleration, not the cause.***

American mathematics achievement peaked in 2013 and has fallen almost every year since. On the main National Assessment of Educational Progress, fourth- and eighth-grade mathematics proficiency reached its high point in 2013 and was already in measurable decline by 2019 — before a single school closed for COVID-19. Twelfth-grade mathematics peaked in the same year and had also fallen by 2019. The long-term trend assessment, which reaches back to 1971 and offers the cleanest view of the arc, shows nine- and thirteen-year-olds peaking in 2012.[1,2,3]

The pandemic was an accelerant applied to a fire that was already burning. Independent analysis of the lowest-performing eighth graders finds that if the 2013–2019 trajectory had simply continued with no pandemic at all, roughly half of the total decline observed by 2024 would have happened anyway. The pattern is global: PISA 2022 documented mathematics declines across the majority of participating economies.[2,4]

So the question is not what happened in 2020. The question is what changed in American classrooms around 2012 and 2013.

What changed

One thing changed at a scale that dwarfs every curriculum adoption, every reform initiative, and every funding shift of the same period: the surface on which children do their work.

The Chromebook entered education in 2011–2012. Computer-delivered state testing arrived at scale in 2014–2015, and districts that bought devices in order to test began using them to teach. By 2017–2018, Chromebooks accounted for roughly sixty percent of computing devices purchased for American K–12 schools. In 2017, NAEP itself moved to tablets; the mathematics assessment is now delivered on a tablet with an attached keyboard and an on-screen scratchpad. The federal relief funding of 2020–2022 completed the transition. One-to-one device deployment in American districts is now at or near saturation.[5,6]

Within roughly a decade, the default place where an American student does mathematics moved from a piece of paper to a pane of glass. No efficacy study was required of this transition. No district was asked to demonstrate that it improved learning. It was purchased as infrastructure, on a procurement logic, and it changed the physical act of learning mathematics for fifty million children.

Why this matters cognitively

Handwriting is not a slower way of producing the same output as typing. It is a different act with different neural consequences. High-density EEG work published in 2024 found that handwriting produces widespread connectivity across parietal and central brain regions in the theta and alpha frequency bands associated with memory encoding, and that typewriting does not.[7] Functional imaging comparing paper notebooks to tablets and smartphones found that participants who wrote on paper recalled more accurately, showed stronger activation in the hippocampus and precuneus during retrieval, and completed the task about twenty-five percent faster than the digital groups.[8] In pre-literate children, printing letters by hand activates the reading circuit; typing the same letters does not.[9]

The mechanism is not mysterious. Every keystroke is the same movement. Pressing the key for 7 and pressing the key for 4 are spatially identical acts; the finger travels a slightly different distance and does the same thing. Writing a 7 and writing a 4 are two entirely different motor programs, and the brain stores that motor program alongside the visual form of the symbol. The hand teaches the eye what it is looking at. A keyboard gives the brain nothing to graph.

The developmental evidence is sharper still. When four-year-olds were trained on letters by four different methods — hearing and saying the names, typing them, tracing them, and printing them freehand — only free-hand printing recruited the visual network that underlies letter recognition in literate adults. Typing did not. Tracing did not. The active ingredient turns out to be the messy variability of a child's own self-produced marks, and it is precisely that variability which every keyboard and every answer field eliminates.[9]

2013

The inflection point

NAEP mathematics proficiency peaked at grades 4, 8, and 12 in 2013 and was already declining before 2019. The long-term trend assessment shows ages 9 and 13 peaking in 2012.

0.18 SD

The randomized evidence

In a randomized trial across 50 classrooms and 726 students, permitting computers in the classroom lowered final exam scores by 0.18 standard deviations. Restricting tablets to flat-on-the-desk use did not remove the effect. No subgroup benefited.

~60%

The scale of the change

Share of computing devices purchased for U.S. K–12 schools that were Chromebooks by 2017–2018 — the leading edge of a transition now at or near saturation.

Why mathematics is the subject most exposed

Reading and writing survived the transition better than mathematics did, and there is a reason. Mathematical notation is two-dimensional. Place value requires vertical alignment. Fractions require a bar with terms above and below it. Exponents require elevation. Long division, systems of equations, and matrix work all require spatial organization that a one-dimensional text entry field cannot express. When children with weak visual-motor integration misalign numbers on a page, they produce wrong answers even when the underlying computation is correct.[10] The page is not a container for mathematical thinking. It is part of the apparatus.

This is why fine motor skill predicts early mathematics attainment better than it predicts early reading attainment — a finding that survives controls for non-verbal IQ, while the equivalent effect for reading does not survive controls for verbal short-term memory.[10] Mathematics is the subject in which the written work is not a record of the thinking. It is the thinking.

What we claim, and what we do not

We claim this. The timing is coincidental. The mechanism has been demonstrated experimentally, in children, with before-and-after imaging, and the failure of both typing and tracing to reproduce it rules out the obvious alternative explanations. The subject most dependent on handwritten production is the subject that fell hardest. And where classroom device use has actually been randomized, the measured effect on achievement has been null to negative. That is a serious and well-supported hypothesis — and one the field has never tested, largely because it never stopped to ask.

We do not claim that computers are the sole cause of the mathematics decline. Achievement is multi-causal: teacher supply and preparation, curriculum churn, chronic absenteeism, adolescent smartphone use, and pandemic disruption all contribute. Nor has any randomized trial isolated the migration from pen to keyboard as the cause of a national achievement trend, and none will, because that experiment cannot be run across a country for a decade. What has been established, repeatedly and in the population that matters most, is the mechanism.

The correction

> ***The choice was never paper or the computer. It was always the false choice that cost us a decade.***

The prescription that follows from this analysis is not to take devices away. It is to stop asking children to do mathematics in the one medium their brains handle worst, and to stop asking teachers to diagnose reasoning from a record that contains none.

The smartpen resolves the trade-off. The student writes on ordinary paper with an instrument that feels like a pen. The cognitive benefit of the motor act accrues to the learner, in full. The complete timestamped stroke record — the pencast — flows to the computer, where it becomes something paper could never provide: a searchable, longitudinal, analyzable record of how a student actually reasons. The hand returns to the student. The data returns to the teacher.

PART I — THE CHRONOLOGY

A decline that predates the pandemic

Begin with the fact that most public discussion still gets wrong.

American mathematics achievement did not begin falling because of COVID-19. It began falling in 2013. On the main NAEP, the share of fourth graders scoring at or above proficient in mathematics peaked in 2013; so did the share of eighth graders. Both had declined measurably by 2019, before any school closure. Twelfth-grade mathematics peaked in the same year and had also fallen by 2019.[1, 3]

The long-term trend assessment, which has used a stable instrument since 1971 and is therefore the most reliable single view of the arc, places the peak for nine-year-olds and thirteen-year-olds in 2012. Compared to a decade earlier, average long-term trend scores for thirteen-year-olds had fallen fourteen points in mathematics.[2]

Quantitative work on the shape of the decline reinforces the point. Projecting the 2013–2019 trajectory of the lowest-performing eighth graders forward to 2024 — that is, assuming no pandemic at all — predicts a decline of roughly 5.6 scale score points. The actual observed decline was 12.0 points. Nearly half of what is routinely attributed to the pandemic was already in motion before it.[3]

The shape of the decline is as diagnostic as its timing

Three features of the decline deserve attention because they constrain which explanations are plausible.

- **The distribution stretched.** High-performing students largely held their ground. Low-performing students fell, and kept falling. This is not what a general instructional-quality story predicts; it is what a story about a change that removes scaffolding from the students who most depend on scaffolding predicts.
- **It was global, not American.** PISA 2022 documented mathematics declines across the majority of participating economies. Whatever is happening is not explained by any single national policy — not Common Core, not accountability design, not one country's teacher pipeline. It is explained better by something that happened in most developed education systems at roughly the same time.[4]
- **Mathematics diverged from reading.** The two subjects have not moved in lockstep, and their post-pandemic recoveries have differed. Mathematics has proven more responsive to

remediation in some windows and more fragile in others. A subject-specific mechanism is required to explain a subject-specific pattern.

EXHIBIT 1 — TWO TIMELINES, ONE DECADE

Year	Mathematics achievement	The instrumentation of student work
2011–12	NAEP long-term trend peaks at ages 9 and 13.	Chromebook enters the education market; early 1:1 pilots begin.
2013	Main NAEP mathematics peaks at grades 4, 8, and 12.	District device procurement accelerates ahead of online testing mandates.
2014–15	Decline begins across grades.	Computer-delivered state testing arrives at scale; devices bought to test are used to teach.
2017	Declines continue at grades 4 and 8.	NAEP itself moves to tablets. Mathematics is delivered on a tablet with attached keyboard and an on-screen scratchpad.
2017–18	Pre-pandemic decline is now unmistakable in the data.	Chromebooks reach roughly 60% of U.S. K–12 device purchases.
2019	Every grade level below its 2013 peak.	1:1 deployment approaching majority status in American districts.
2020–22	Sharpest single-cycle declines in the history of NAEP.	Federal relief funding completes the transition to near-universal 1:1.
2024–26	Partial, uneven, incomplete recovery.	Device saturation. Handwritten mathematical production at a historic low.

> ***Two curves, one decade, opposite directions. That is not proof of causation. It is the beginning of a question no one has seriously asked.***

PART II — THE KEYBOARD TURN

What actually changed in American classrooms after 2012

It is worth stating plainly what happened, because the field has largely stopped noticing it.

In 2012, a typical American student learning mathematics worked problems with a pencil on paper. They wrote the problem, they wrote each step, they crossed things out, they erased, they aligned columns, and they arrived at an answer that sat at the bottom of a visible chain of reasoning. Homework came back with marks on it. A teacher walking the room could glance at a page and see where the thinking had gone wrong.

By 2022, that same student was far more likely to work problems inside a browser tab. They read the problem on a screen, performed some or all of the intermediate work in their head or on a scrap they discarded, and entered a final answer into a box or selected it from four options. The system recorded whether the answer was right. Nothing recorded how they got there — including, in many cases, the student.

Three forces drove the transition, none of them pedagogical.

1. Testing drove procurement

The move to computer-delivered state assessment in 2014 and 2015 required districts to have enough devices to test every student within a testing window. Districts bought devices to satisfy a measurement requirement. Having bought them, they used them for instruction the other forty-nine weeks of the year. The instructional shift was a byproduct of a logistics decision.

2. Devices were cheaper than the alternatives

A managed Chromebook at two hundred to four hundred dollars, with centralized administration and a multi-year refresh cycle, was a solved procurement problem. Paper, consumables, copying, and physical grading were not. The economics ran one direction only, and by 2017–2018 roughly sixty percent of U.S. K–12 device purchases were Chromebooks.[5]

3. Digital delivery was assumed to be an improvement

This was never demonstrated. It was assumed. The assumption was reasonable in 2012 and has aged poorly. As early as 2015, the OECD — analyzing PISA data across sixty-four systems — reported that countries that had invested heavily in information and communication technology for education showed no appreciable improvement in reading, mathematics, or science. Students who used computers moderately at school did somewhat better than those who used them rarely; students who used them very frequently did substantially worse.[11]

That finding was published, widely covered, and functionally ignored. Procurement continued.

> ***This was the largest uncontrolled change to the physical act of schoolwork in modern educational history, and no one was asked to show that it worked.***

We want to be precise about the claim. The problem is not that children use computers. The problem is displacement: for a great many students, digital delivery did not supplement handwritten mathematical work, it replaced it. The minutes a student spends producing mathematics by hand — forming numerals, aligning place value, writing an intermediate result and then operating on it — have fallen dramatically, and no one measured the decline because no one was counting.

A district can find this out in an afternoon. Ask a fourth-grade teacher how many minutes a week each student spends writing mathematics by hand, and compare it to 2012. In most systems we have worked with, the honest answer is a fraction of what it was, and it is not the result of a decision anyone made on purpose.

PART III — THE HAND THAT TEACHES THE MIND

What handwriting does that typing does not

The case that handwriting is cognitively distinct from typing is not folklore, and it does not rest on nostalgia. It rests on five converging lines of evidence, each measuring something different.

Mechanism 1: Handwriting builds the recognition network — and typing does not, and neither does tracing

This is the core of the argument, and the developmental evidence behind it is far more specific than most educators realize.

Start with what a keystroke is. When a student forms the numeral 7 by hand, the brain executes a unique motor program: a horizontal stroke, a change of direction, a descending diagonal. Forming a 4 executes an entirely different program. The visual form of the symbol and the motor program that produced it are encoded together, which is why the hand is a teacher — the act of production creates a second, independent representation of the symbol, stored in the motor cortex, available to reinforce recognition. When the student types 7 and then types 4, the brain executes the same motor program twice. Index finger, downward press, release. The location differs slightly. The action does not.

What elevates this from a plausible story to a demonstrated mechanism is a program of developmental neuroimaging conducted by Karin James and colleagues at Indiana University over roughly a decade, using functional imaging of four-, five-, and six-year-old children before and after controlled training. It is worth walking through carefully, because it forecloses the objections usually raised against it.[9]

Before instruction, the brain does not distinguish letters from shapes. Four-year-olds were scanned before and after four weeks of letter training. Prior to training there was no letter-specific activation at all: their brains responded to letters exactly as they responded to triangles and squares. Only after handwriting practice did the visual regions that become specialized for letter recognition in literate adults begin to respond to letters as letters. Training by hearing and saying the letter names — the method most commonly used in preschool, on the assumption that writing is too difficult at that age — did not produce the effect.

Typing did not work. Tracing did not work either. A subsequent experiment compared four conditions directly: see-and-say, free-hand printing, typing on a keyboard, and tracing a pre-drawn form. Only free-hand printing recruited the letter-recognition network. This is the single most important finding for the argument of this paper, and the one most often missed. It is not merely that a screen is worse than paper. It is that a motor act which is not self-generated production — pressing a key, or guiding a stylus along a line someone else drew — does not build the network at all. The action has to be specific.

Handwriting is what wires perception to the motor system. Functional connectivity analysis established that the visual region active during letter perception becomes functionally connected to motor and sensorimotor regions only as a result of handwriting experience. Children who learned by typing could still see the letters. They did not develop the connection.

The active ingredient is variability. Why should typing fail, and why should tracing — which does involve the hand — fail alongside it? Because when a young student copies a symbol freehand, the results are messy and highly variable: three attempts at the same letter by the same student look meaningfully different. Tracing produces one invariant form. Typing produces a perfect one. And variable instances are what build a category. A toddler learns what a duck is from the rubber duck, the pond duck, and the picture-book duck — not from seeing the same duck a hundred times. Tested directly with unfamiliar Greek symbols, every condition that exposed five-year-olds to variable handwritten forms produced better categorization than conditions using a single invariant printed form.

But the variation has to be the child's own. A further study taught six-year-olds an unfamiliar cursive script either through self-production or by watching an experimenter produce the same variable forms. Viewing variable forms improved categorization behaviorally. Only self-production recruited the perceptual-motor network. Seeing is not enough. The child's own hand has to make the mark.

> ***Neatness is not the point. The child's messy, variable, self-produced numeral is the mechanism — and it is precisely the artifact that every digital answer field discards.***

The James program concerns letters, because letters are where preliterate symbol learning can be studied cleanly. But the mechanism it identifies — self-generated, variable production building the perceptual-motor network that underlies fluent symbol recognition — is not a fact about the alphabet. It is a fact about how the developing brain acquires symbol systems through action. Mathematics is a symbol system, acquired later, with more symbols, heavier spatial structure, and a far greater procedural load. There is no principled reason the mechanism would apply to the letter A and not to the numeral 8, to a word and not to an equation.

Work using other methods and other populations converges on the same conclusion. Van der Weel and van der Meer, using high-density EEG, found that handwriting produced widespread connectivity across parietal and central regions in the theta and alpha bands associated with memory formation and learning, and that typewriting did not produce comparable patterns.[7] Longcamp and colleagues, working with both children and adults, found that training by handwriting produced better later recognition of novel characters than training by typing.[12]

Umejima and colleagues extended the finding to retrieval. Participants who recorded information in a paper notebook, rather than on a tablet or smartphone, showed higher recall accuracy and stronger activation in the hippocampus, precuneus, and language regions during retrieval — and, contrary to the productivity assumption that drove the digital transition, completed the encoding task roughly twenty-five percent faster than the digital groups.[8]

1 of 4

Only one training condition built the network

Four-year-olds learned letters by seeing-and-saying, typing, tracing, or free-hand printing. Only free-hand printing recruited the visual network that underlies letter recognition in literate adults. Typing did not. Tracing did not.

EXHIBIT 2 — FOUR WAYS TO LEARN A SYMBOL, ONE RESULT

Training condition	What the student actually does	Builds the recognition network?
See-and-say	Hears and repeats the letter name. No production of any kind.	No
Typing	Presses a key. The motor act is identical for every symbol.	No
Tracing	Follows a pre-drawn form. Hand is involved, but output is invariant.	No
Free-hand printing	Self-generates the form. Output is messy, effortful, and variable.	Yes

Mechanism 2: Production is not selection

Handwriting requires production from memory. Multiple choice requires recognition. Typing a numeral into an answer box sits between the two, but closer to recognition than most educators assume: the symbol set is visible on the keyboard, the act is one of location rather than construction, and the correction of an error leaves no trace.

The distinction matters because retrieval practice — the effortful reconstruction of knowledge from memory — is among the most robust findings in the science of learning, and because a system that permits selection permits guessing. A student who guesses correctly on a screen is recorded as having understood. The same student, working on paper, produces a page that shows a teacher immediately that no reasoning occurred.

Mechanism 3: The slowness is the point

Handwriting is slower than typing. In the digital transition this was treated as a defect. In learning it is frequently a feature — what Bjork and Bjork termed a desirable difficulty: a condition that impairs performance during practice while improving retention and transfer.[13]

In mathematics, the slowness performs a specific function. Writing an intermediate result takes long enough that the value is held in working memory while it is externalized, then read back and operated on. That loop — hold, externalize, re-read, operate — is a high-order cognitive exercise. It is precisely the loop that mental-math-into-an-answer-box removes.

We should be candid here about one strand of this literature. The widely cited finding that longhand note-taking beats laptop note-taking in university lectures[14] did not cleanly replicate when Morehead, Dunlosky, and Rawson ran a direct replication with additional control conditions.[15] We cite that failure deliberately, because it is the strongest available counterargument and because a briefing that hid it would deserve to be dismissed. But it does not undermine the case made here. The note-taking studies concern transcription strategy in adults listening to a lecture. The mathematics claim rests on a different foundation: motor encoding of symbols in developing children, the two-dimensional structure of mathematical notation, and the role of written work as external working memory. Different task, different mechanism, different evidence base.

Mechanism 4: Notation is spatial, and a text field is not

Mathematics is written in two dimensions. Place value requires vertical alignment. A fraction requires a bar with terms above and below. Exponents require elevation. Long division requires a specific spatial arrangement that encodes the algorithm itself. Systems of equations require alignment across lines.

A sheet of paper is an unconstrained two-dimensional workspace with effectively infinite resolution. A text entry field is a one-dimensional string. Every mathematical structure a student writes on paper must, in a typed environment, be either flattened into linear notation or performed somewhere the system cannot see.

The consequences are measurable. Barnhardt and colleagues found that children with weaker visual-motor integration made more errors in the spacing and alignment of letters, words, and number problems on a page — and that these spatial failures produced wrong answers even where the underlying computation was correct.[10] Spatial organization is not presentation. In mathematics it is part of the operation.

Mechanism 5: Written work is external working memory

A student who must consciously retrieve a multiplication fact cannot simultaneously attend to the logic of a multi-step problem. Automaticity is not rote memorization for its own sake; it is what frees working memory for reasoning, a point on which the National Mathematics Advisory Panel and the deliberate practice literature converge.[16] [17]

Written work is the cheapest available extension of working memory. Cognitive load theory makes the prediction explicit: when the notation is fluent and automatic, the page absorbs load that would otherwise consume the limited capacity available for reasoning.[18] When the notation must be typed, or reconstructed mentally because there is nowhere to put it, that load returns to the student — and it returns hardest for the students with the least capacity to spare. Which is precisely the group in which the national decline has been steepest.

EXHIBIT 3 — FIVE MECHANISMS, AND WHAT THE KEYBOARD REMOVES

Mechanism	What the hand contributes	What is lost when work moves to glass
Motor encoding	A unique motor program per symbol, self-generated and variable, stored alongside the visual form.	Every keystroke is the same movement. No symbol-specific trace is formed, and no variability is produced.
Production over selection	Effortful reconstruction from memory; guessing is visible.	Recognition replaces recall. A correct guess is recorded as understanding.
Productive slowness	Time to hold, externalize, re-read, and operate on an intermediate result.	Steps are skipped or performed mentally; the reasoning loop collapses.
Spatial notation	Two-dimensional alignment: place value, fractions, exponents, long division.	A one-dimensional entry field cannot express the structure the algorithm depends on.
External working memory	The page absorbs cognitive load, freeing capacity for reasoning.	Load returns to the student, hardest on those with least capacity to spare.

PART IV — WHY MATHEMATICS IS THE MOST EXPOSED SUBJECT

If the argument above were correct in general, every subject should have suffered equally. Mathematics has suffered distinctively, and the reason follows from what mathematics is.

Fine motor skill predicts mathematics more than it predicts reading

Pitchford and colleagues, working with primary-school children in the United Kingdom, found that fine motor skills predicted early mathematics ability better than they predicted early reading ability. Fine motor integration remained a significant predictor of mathematics attainment even after non-verbal IQ was controlled. For reading, the equivalent effect disappeared once verbal short-term memory was accounted for.[10]

The interpretation the authors offer is the one that matters here: in mathematics, the motor movement maps directly onto the conceptual representation. Writing a numeral is not a clerical act performed after the thinking. In early mathematics, forming the symbol and forming the concept are the same operation.

In mathematics, the work is the thinking

In reading, comprehension can be evidenced by a response: a summary, an answer, an inference stated aloud. The intermediate cognition is genuinely internal, and asking a student to write it down is a way of eliciting evidence about a process that happened elsewhere.

Mathematics is not like that. The written work is where the process happens. The line of algebra is not a report about the reasoning; it is the reasoning, externalized so that the next step can operate on it. When a system removes the written work, it does not lose the record of the thinking. It removes a load-bearing part of the thinking.

This is the core asymmetry, and it is why a change that was merely unhelpful in most of the curriculum was corrosive in mathematics.

The anxiety loop closes faster on a screen

Mathematics is also the subject with the most thoroughly documented affective dimension. Meta-analysis across 747 effect sizes reports a consistent negative relationship between mathematics anxiety and mathematics achievement of approximately $r = -0.28$; the largest available synthesis covers more than nine hundred thousand participants.[19] Longitudinal work shows the anxiety emerging in elementary school and persisting into adulthood, and shows teachers' own mathematics anxiety transmitting to students with measurable effects.[20]

Anxiety operates through avoidance. A student who does not understand a procedure and can conceal that fact will conceal it, and the concealment compounds because the gap goes unaddressed. Here the medium matters enormously. On paper, a student who does not know what to do produces a blank space or a false start, and both are visible. In an answer box, the same student enters a plausible number, receives feedback on the answer alone, and moves on. The system records an incorrect response. Nobody — not the teacher, not the parent, not the student — learns anything about why.

Multiply that across a school year and the avoidance loop closes long before any adult has a reason to intervene.

r ≈ −0.28	**Mathematics anxiety and achievement** Consistent negative relationship reported across 747 effect sizes; the largest available meta-analysis covers more than 900,000 participants. Anxiety operates through avoidance — and avoidance is easier to hide in an answer box than on a page.

PART V — THE DISPLACEMENT EVIDENCE

What happened where classroom devices were actually studied

The neuroscience establishes a mechanism. The chronology establishes coincident timing. Neither establishes effect. For that we need studies in which device use itself was manipulated, and the results are consistent enough to be uncomfortable.

The randomized trial

Carter, Greenberg, and Walker conducted a randomized controlled trial at the United States Military Academy across fifty classrooms and 726 students in an introductory economics course with a standardized syllabus, common examinations, and mandatory attendance — conditions that eliminate most of the confounds that plague classroom research. Sections were randomly assigned to technology-free, technology-at-will, or tablets-flat-on-the-desk conditions.

Students in classrooms that permitted computers scored 0.18 standard deviations lower on the final examination than students in classrooms that prohibited them. The restricted-tablet arm, designed specifically to suppress social media and messaging, showed a comparable effect of 0.17 standard deviations. When the authors examined subgroups by gender, race, entrance examination scores, and entering GPA, no group benefited from computer access.[21]

These are highly motivated, highly selected young adults in a structured environment. The relevant comparison population — nine-year-olds — has less self-regulation, not more.

The international comparison

The OECD's 2015 analysis of PISA data across sixty-four education systems found no appreciable improvement in reading, mathematics, or science performance in the countries that had invested most heavily in classroom technology. The dose-response pattern was not linear:

moderate school computer use was associated with somewhat better outcomes than rare use, while very frequent use was associated with substantially worse outcomes.[11]

This was published in 2015 — two years after the American mathematics peak, and at the leading edge of the device buildout it should have slowed.

The medium effect in reading

The closest analogue outside mathematics is the reading-medium literature, where meta-analysis has found a persistent comprehension advantage for print over screen, with the effect appearing to grow rather than shrink over the period studied.[22] The relevance here is directional rather than direct: it establishes that the medium of academic work has measurable cognitive consequences independent of the content delivered through it.

The measurement itself moved

There is an additional wrinkle that deserves more attention than it has received. In 2017, NAEP transitioned mathematics at grades 4 and 8 from paper to tablet. NCES conducted its most ambitious bridge study to date to preserve trend comparability, and reported that the study confirmed small but meaningful differences in performance depending on mode of administration.[6]

The linking procedures were carefully done and we do not dispute the trend line. But note what the transition means substantively: the nation's most authoritative measure of mathematics now asks children to do their scratch work on an on-screen scratchpad with a stylus or a keyboard. Whatever handwriting contributes to mathematical performance, the instrument that measures the national decline is now administered in a medium that does not permit it.

EXHIBIT 4 — THE EVIDENCE ON CLASSROOM DEVICES

Study	Design	Finding
Carter, Greenberg & Walker (2017)	Randomized trial, 50 classrooms, 726 students	Permitting computers lowered exam scores 0.18 SD; restricted tablets 0.17 SD; no subgroup benefited.
OECD (2015)	PISA data across 64 education systems	No appreciable gains in systems investing heavily in ICT; very frequent school computer use associated with substantially worse outcomes.
Sana, Weston & Cepeda (2013)	Experimental, classroom multitasking	Laptop multitasking harmed the user and students seated nearby.
Delgado et al. (2018)	Meta-analysis, reading medium	Comprehension advantage for print over screen, growing across the period studied.
NCES (2017)	NAEP mode-evaluation bridge study	Confirmed small but meaningful mode differences between paper and digital administration.

> ***The burden of proof was never met. It was never requested.***

PART VI — WHAT THIS ARGUMENT ESTABLISHES, AND WHERE ITS BOUNDARIES LIE

KAITLab has committed, in every briefing in this series, to stating plainly what our evidence establishes and where its boundaries lie. We do that here not to soften the argument but to make it durable. A thesis that survives its own honest accounting is worth considerably more than one that has to be protected by omission.

What the evidence establishes

- **The decline predates the pandemic.** Documented, uncontested among researchers, and still widely misunderstood in public discussion. Any account of American mathematics performance that begins in 2020 is describing an acceleration and calling it a cause.
- **Handwriting builds neural architecture that typing does not.** This is not an inference from correlation. It has been demonstrated experimentally, in children, with before-and-after imaging, across a decade of studies — and the failure of both typing and tracing to

reproduce the effect rules out the obvious alternative explanations. The specificity of that result is what makes it strong.

- **Mathematics is the subject most dependent on handwritten production.** Established by the fine-motor findings, by the irreducibly two-dimensional structure of mathematical notation, and by the role of written work as external working memory.
- **Handwritten mathematical work was displaced at national scale in the same window as the decline.** Documented by device procurement data and by the migration of the measurement instruments themselves onto screens.
- **Where classroom device use has been randomized, the effect on achievement is null to negative.** The randomized evidence runs in one direction. Critics of these studies dispute their generalizability; none has produced the offsetting positive trial.

These are not five loose observations. They are a mechanism, a population-scale exposure to that mechanism, a coincident outcome, and a set of experiments in which the exposure was manipulated and behaved exactly as the mechanism predicts. That is the standard structure of a well-supported causal argument in any applied field.

Where the boundaries lie

- **No trial has isolated this cause at national scale — and none will.** Randomizing the medium of schoolwork across a country for a decade is not an experiment anyone can run. The relevant question is not whether that impossible study exists. It is whether the mechanism has been demonstrated where it can be tested. It has, repeatedly, in the population that matters most.
- **Achievement is multi-causal.** Teacher supply and preparation, curriculum churn, chronic absenteeism, adolescent smartphone use, and pandemic disruption all contribute. We do not claim the medium of student work is the largest factor. We claim it is the factor the field has not examined — and, not incidentally, the one most directly correctable.
- **One strand of the handwriting literature is contested.** The adult note-taking advantage did not cleanly replicate. We name it because it is the strongest available counterargument, and because naming it costs us very little: that literature concerns transcription strategy in adults listening to lectures. The developmental imaging work, which is what this paper actually rests on, has replicated and extended across a decade of studies and multiple methods.

Why the evidence supports acting now

There is a methodological point here that deserves more weight than it usually receives. Neural change frequently precedes and predicts behavioral change; brain measures can detect the formation — or the absence — of a developing capability well before achievement measures can register its consequences.[23] A field that insists on an achievement-level randomized trial before acting on a demonstrated developmental mechanism is not being rigorous. It is choosing to detect the problem later, in a cohort that has already lost the years in question.

The decision facing a district is also asymmetric. If we are wrong, a district that protects and instruments handwritten mathematical production has spent modestly and gained a diagnostic signal on student reasoning it did not previously have. If we are right, a district that waits continues removing the encoding mechanism from the students least able to absorb the loss, and remains blind to the reasoning of every student in every classroom, for as long as the wait lasts.

> ***The intervention is also the experiment. The pencast corpus is the only instrument capable of testing this briefing's central claim — which is why we would rather build it than argue about it.***

One disclosure belongs here, and we would rather state it than have it inferred. KAITLab builds the instrument that recovers handwritten work. That is a statement of alignment rather than of bias. We did not arrive at this position because we sell a smartpen; we built the company because the evidence pointed here, and it has kept pointing here every time we have gone back to check.

PART VII — RESTORING THE HAND WITHOUT SURRENDERING THE COMPUTER

Everything above could be read as an argument for taking devices out of classrooms. It is not, and the schools that read it that way will get the wrong answer.

The computer is genuinely good at things paper cannot do. It stores. It searches. It aggregates across a class, a school, a year, a network. It runs the psychometrics that estimate where a student actually is. It runs the models that decide what problem should come next. It generates language a parent can understand at a kitchen table. None of that is available on paper, and no amount of nostalgia will make it so.

Paper is genuinely good at exactly one thing the computer cannot do: it is the medium in which the human hand produces mathematical work, and it is therefore the only medium in which the cognitive benefits described in Part III are available.

For a decade the field has treated these as substitutes and chosen the computer. They were never substitutes. They are complements, and the missing piece was never a better application. It was a sensor.

> ***The student writes on paper. The computer receives everything. Nothing is traded.***

The smartpen resolves the false choice completely. The student writes with an instrument that feels like a pen, on ordinary paper, in the posture and with the motor act that a century of cognitive science says builds mathematical memory. Nothing about the learning experience is degraded, and one thing about it is meaningfully improved: the student is not looking at another screen.

Simultaneously, the pen captures the full timestamped stroke record — the pencast — and delivers it to the KAITLab platform. Every stroke, in order, with millisecond timing, including everything beneath every erasure and scratch-out. The learner keeps the cognitive benefit in full. The system gains a record richer than anything a screen-based tool has ever produced, because a screen-based tool sees only what a student chose to submit.

Two practical consequences follow from the developmental evidence in Part III, and both favor the pen over every digital approximation of it. First, stylus-and-tablet tracing exercises — a common feature of early-grades software and a frequent claim of equivalence — are not a substitute. Tracing was tested directly, against free-hand printing, and it did not build the network. Second, penmanship quality is not the objective and should never be treated as one. The variability of a young child's self-produced numeral is the mechanism, not a defect in it. KAIT is built to read messy handwriting because messy handwriting is what learning looks like.

This is the architectural argument of our first briefing restated in cognitive terms. The smartpen is not an accessory to the KAITLab platform. It is the infrastructure. KAIT and Swoop! are the applications. And the reason the infrastructure has to be a pen rather than a camera, a keyboard, or a touchscreen is now, we hope, clear: those instruments capture output. The pen captures cognition.

PART VIII — WHAT THE STORED PENCAST MEASURES

A photograph of a completed page is a record of what a student finally wrote. A pencast is a record of what a student was doing. The difference is the difference between a transcript and a recording, and it makes a class of measurement possible that has not previously existed in education at scale.

Learning velocity

Across sessions on the same procedure, the pencast shows pauses shortening, stroke paths cleaning up, self-corrections decreasing, and pen-lift patterns consolidating. This is a direct measure of how fast a specific student is building fluency on a specific skill — not whether they eventually got it right, but how quickly the procedure is becoming automatic.

No conventional instrument measures this. A diagnostic administered three times a year measures position, not velocity. Velocity is what tells a teacher whether the current intervention is working this week, and it is the earliest available signal that it is not.

Abandonment and the shape of struggle

The most consequential thing a struggling student does is stop, and stopping is invisible to every conventional system. The pencast makes it legible: time-to-first-stroke, mid-problem cessation, the length and placement of pauses, whether a student returned to a line or moved on, whether the page went blank after a specific step.

This is what distinguishes productive struggle — hesitation followed by revision followed by progress — from stuck struggle, and both from disengagement. It is the earliest observable manifestation of the mathematics-anxiety avoidance loop described in Part IV, and it appears in the data weeks or months before it appears in a grade.

Granular error localization

A final answer tells a teacher that a student got a problem wrong. A pencast tells the teacher that the student was correct through the distribution step, made a sign error in the third line, and did not catch it. Those are different problems requiring different responses, and until now the only way to tell them apart was for an adult to sit beside the student and watch.

Evidence beneath erasures

What a student tried and rejected is frequently more diagnostic than what they submitted. A camera-based tool sees a cleaned-up page. The pencast preserves the abandoned attempt, the strikethrough, the half-written approach — the moment the student reached for a procedure and then thought better of it.

Production variability

Because the pencast preserves every instance of every symbol a student produces, it holds a record of precisely the variability the developmental literature identifies as the active ingredient in symbol learning: how a child's numerals and notation stabilize over weeks and months, where they consolidate, and where they do not. No instrument in education has ever held that record at all, and no digital input field creates it in the first place.

Reasoning state

The classifier described in our first briefing assigns each solution attempt to one of four states: Understood, Guess, Mistake, or Not Understood. This is the distinction conventional instruments structurally cannot make. A student who guessed correctly and a student who understood produced identical records in an answer-box system. They produce entirely different pencasts.

Teacher- and classroom-level signal

Individual precision is only useful if it aggregates into something a teacher can act on in the time a teacher actually has. The KAIT dashboard rolls stroke-level analysis up into standard-level class profiles: which standards the class is consolidating, which are fracturing, which three students need a specific move today and what that move is. The educator-adoption data from CICS Washington Park reported in our second briefing — an average of 87.3 training reviews per active teacher — indicates that when teachers engage the data layer, student impact amplifies. The signal is only as valuable as the teacher's ability to act on it.

EXHIBIT 5 — WHAT THE PENCAST MEASURES THAT NOTHING ELSE CAN

Measure	What it captures	Why it was previously invisible
Learning velocity	Rate at which pauses shorten and stroke paths consolidate on a given procedure.	Point-in-time diagnostics measure position, not rate of change.
Abandonment	Time-to-first-stroke, mid-problem cessation, blank-page latency.	A screen records a missing answer, not the moment the student stopped.
Error localization	The exact line and operation where reasoning broke down.	A final answer is a single bit of information about a multi-step process.
Evidence beneath erasures	Attempts made and rejected before the submitted answer.	Cameras and keyboards capture only the cleaned-up final artifact.
Reasoning state	Understood / Guess / Mistake / Not Understood.	A correct guess and genuine understanding are indistinguishable in an answer box.
Production variability	Every instance of every self-produced symbol, tracked across weeks and months.	The active ingredient in symbol learning — and an artifact no digital input field creates.
Fluency and automaticity	Writing velocity, active-pen time, erasure ratio.	Requires a sensor at the point of writing; validated in the digital-pen literature.[24]

One further point deserves emphasis. As deployment scales across districts and homes, the anonymized pencast corpus becomes something education research has never had: a longitudinal, stroke-level record of how large populations of children actually produce mathematics. That corpus is the instrument capable of testing the hypothesis this briefing advances — of establishing, empirically, what the handwritten production of mathematics contributes and what its absence costs. We would rather build that evidence than assert it.

PART IX — IMPLICATIONS AND RECOMMENDATIONS

For district and school leaders

Start by measuring what you have stopped measuring. Ask, grade by grade, how many minutes per week a student in your system produces mathematical work by hand. Compare that honestly to 2012. In most systems the answer is a fraction of what it was, and no one decided this — it accumulated, one platform adoption at a time.

Then treat handwritten mathematical production as protected instructional time rather than as a legacy practice awaiting replacement. This costs nothing. It requires only that the question be asked.

Finally, instrument the paper. Handwritten work is currently the largest source of un-captured learning data in every district in the country. It is also, on the evidence assembled here, the most cognitively valuable work students do. Reinstrumenting it is the single highest-leverage measurement change available, and it is additive: KAIT sits alongside the curriculum and adaptive-practice investments a district has already made and compounds the return on both.

For teachers

Show your work has always been the right instruction, and it has always been unenforceable at scale, because a teacher with twenty-five students and fifty minutes cannot watch twenty-five reasoning processes. That is not a failure of teaching effort. It is a failure of instrumentation, and it is the specific failure KAIT was built to correct.

For families

Two things are worth knowing. First, the intuition that your student does less written mathematics than you did is almost certainly correct, and it is not a matter of standards being lowered — it is a matter of the medium having changed underneath everyone. Second, a monthly score is not insight. Swoop! exists to give a parent at a kitchen table the same view of a child's reasoning that a good tutor would have: what has been mastered, where the student is productively struggling, where they are stuck, and what to do tonight.

For policymakers

Three propositions follow from this analysis. Efficacy evidence should be required of device-delivered instructional materials at the same standard applied to curriculum. Handwritten production should be explicitly protected in elementary mathematics guidance rather than left to be displaced by default. And the wholesale migration of elementary mathematics testing to screens deserves review, given the mode effects the measurement agencies themselves have documented.

EXHIBIT 6 — RECOMMENDED ACTIONS BY AUDIENCE

Audience	Immediate action	Structural action
District leaders	Audit handwritten mathematics production minutes by grade against 2012 baseline.	Instrument paper-and-pencil work; treat the pencast layer as a diagnostic line item, not a curriculum replacement.
Principals	Protect written production time in the mathematics block.	Build teacher routines around reasoning-state signals rather than answer accuracy.
Teachers	Require written work on the problems that matter most.	Use stroke-level signals to target intervention at the step where thinking broke down.
Families	Restore handwritten mathematics practice at home.	Adopt a home instrument that shows reasoning, not just scores.
Policymakers	Require efficacy evidence for device-delivered mathematics materials.	Protect handwritten production in elementary guidance; review all-digital elementary mathematics testing.

CLOSING

A generation of American children has been asked to learn the most notation-dependent subject in the curriculum without writing it down. We did not decide to do this. It accumulated — one testing mandate, one procurement cycle, one platform adoption at a time — and by the time the consequences appeared in the national data, the transition was complete and the cause had become invisible because it had become normal.

We are not arguing that the last decade of educational technology was worthless. We are arguing that in the enthusiasm to put a screen in front of every child, the field removed something it did not know it was removing, and then spent ten years looking for the explanation everywhere except in the change it had made.

> ***You cannot improve what you cannot see. And a student cannot learn what their hand never wrote.***

The correction is not a retreat. It is the recovery of the pen as a scientific instrument. The student writes, as children have always written, and the computer finally sees what happens on the page — every stroke, every hesitation, every erasure, every abandoned attempt. The oldest learning tool in education becomes the newest data instrument, and the mathematics crisis becomes, for the first time, a problem that can be observed while it is happening rather than diagnosed after the fact.

That is the work KAITLab was built to do. We would be honored to do it with you.

For more information, please contact: Terri Dennis, Chief Communications Officer, KAITLab — terri.dennis@kaitlab.com

ABOUT THE AUTHORS

Dr. Joseph Davis is President and Chief Mathematics Officer of KAITLab, where he also leads partner enrollment. He holds a doctorate from Harvard University and has spent more than thirty years as a mathematics educator and systems leader — classroom mathematics teacher, principal, and school superintendent. He leads KAITLab's curriculum integration, instructional design, educator development, and product vision, including the learning progressions, misconception taxonomies, and standards mappings that convert stroke-level evidence into instructional meaning.

Dr. Joseph Wise is Executive Chair and co-founder of KAITLab. He served as superintendent of Duval County Public Schools in Florida and the Christina School District in Delaware, and as Chief Education Officer of Edison Learning. He co-founded Atlantic Research Partners, the Distinctive Schools network, and Acceleration Academies, and chairs LEAP Innovations and America Succeeds. He holds a doctorate from the University of Florida and is the author of three books on teaching, coaching, and hiring.

REFERENCES AND NOTES

The following works are cited in or inform the analysis presented in this briefing. Additional references are available on request.

1. National Center for Education Statistics. (2025). NAEP Report Card: 2024 Mathematics Assessment — Results at Grades 4, 8, and 12. U.S. Department of Education, Institute of Education Sciences. Mathematics proficiency peaked in 2013 at grades 4 and 8 and declined prior to 2019.
2. National Center for Education Statistics. (2023). NAEP Long-Term Trend Assessment Results: Reading and Mathematics, Age 13. U.S. Department of Education. Long-term trend scores peaked in 2012; the age-13 mathematics average fell 14 points over the subsequent decade.
3. Education Next. (2026). "Putting Pandemic Learning Loss in Perspective: Achievement Declines in Twenty-First-Century America." See also EdWorkingPapers policy and practice analysis (2025) projecting the 2013–2019 trajectory for eighth-grade students at the tenth percentile, which predicts approximately 5.6 of the 12.0 observed points of decline by 2024.
4. Organisation for Economic Co-operation and Development. (2023). PISA 2022 Results (Volume I): The State of Learning and Equity in Education. OECD Publishing.
5. Futuresource Consulting, as reported in EdWeek Market Brief (2019) and subsequent IDC Worldwide Quarterly Chromebook Tracker data. Chromebooks accounted for approximately 60 percent of U.S. K–12 computing device purchases by 2017–2018; district-level deployment has since approached saturation.
6. National Center for Education Statistics. (2020). The 2017 NAEP Transition to Digitally Based Assessments in Mathematics and Reading at Grades 4 and 8: Mode Evaluation Study. The bridge study confirmed small but meaningful mode differences in performance. NAEP mathematics is administered on tablets with attached keyboards and an on-screen scratchpad.
7. Van der Weel, F. R., & Van der Meer, A. L. H. (2024). Handwriting but not typewriting leads to widespread brain connectivity: A high-density EEG study with implications for the classroom. Frontiers in Psychology, 14, 1219945.
8. Umejima, K., Ibaraki, T., Yamazaki, T., & Sakai, K. L. (2021). Paper notebooks vs. mobile devices: Brain activation differences during memory retrieval. Frontiers in Behavioral Neuroscience, 15, 634158.
9. The developmental evidence in Part III draws on a program of research by Karin James and colleagues at Indiana University, reviewed in James, K. H. (2017). The importance of handwriting experience on the development of the literate brain. Current Directions in Psychological Science, 26(6), 502–508. Primary studies: James, K. H. (2010). Sensori-motor experience leads to changes in visual processing in the developing brain. Developmental Science, 13, 279–288 (pre/post training in 4-year-olds); James, K. H., & Engelhardt, L. (2012).

The effects of handwriting experience on functional brain development in pre-literate children. Trends in Neuroscience and Education, 1(1), 32–42 (the four-condition comparison of see-and-say, printing, typing, and tracing); Vinci-Booher, S., James, T. W., & James, K. H. (2016). Visual-motor functional connectivity in preschool children emerges after handwriting experience. Trends in Neuroscience and Education, 5(3), 107–120; Li, J. X., & James, K. H. (2016). Handwriting generates variable visual output to facilitate symbol learning. Journal of Experimental Psychology: General, 145, 298–313; Kersey, A. J., & James, K. H. (2013). Brain activation patterns resulting from learning letter forms through active self-production and passive observation in young children. Frontiers in Psychology, 4, Article 567. See also Planton, S., Jucla, M., Roux, F.-E., & Démonet, J.-F. (2013). The "handwriting brain": A meta-analysis of neuroimaging studies of motor versus orthographic processes. Cortex, 49, 2772–2787.

10. Pitchford, N. J., Papini, C., Outhwaite, L. A., & Gulliford, A. (2016). Fine motor skills predict maths ability better than they predict reading ability in the early primary school years. Frontiers in Psychology, 7, 783. On spatial organization and computational error, see Barnhardt, C., Borsting, E., Deland, P., Pham, N., & Vu, T. (2005). Relationship between visual-motor integration and spatial organization of written language and math. Optometry and Vision Science, 82(2), E138–E143.
11. Organisation for Economic Co-operation and Development. (2015). Students, Computers and Learning: Making the Connection. PISA, OECD Publishing. The executive summary reports no appreciable improvements in reading, mathematics, or science in countries that had invested heavily in ICT for education.
12. Longcamp, M., Zerbato-Poudou, M.-T., & Velay, J.-L. (2005). The influence of writing practice on letter recognition in preschool children: A comparison between handwriting and typing. Acta Psychologica, 119(1), 67–79. See also Longcamp, M., et al. (2008). Learning through hand- or typewriting influences visual recognition of new graphic shapes. Journal of Cognitive Neuroscience, 20(5), 802–815.
13. Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In Psychology and the Real World. Worth Publishers.
14. Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. Psychological Science, 25(6), 1159–1168.
15. Morehead, K., Dunlosky, J., & Rawson, K. A. (2019). How much mightier is the pen than the keyboard for note-taking? A replication and extension of Mueller and Oppenheimer (2014). Educational Psychology Review, 31(3), 753–780. A direct, preregistered replication that did not cleanly reproduce the original conceptual-question advantage.
16. National Mathematics Advisory Panel. (2008). Foundations for Success: The Final Report of the National Mathematics Advisory Panel. U.S. Department of Education.
17. Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. Psychological Review, 100(3), 363–406.

18. Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. Educational Psychology Review, 10(3), 251–296.

19. Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. Psychological Bulletin, 147(2), 134–168. See also Caviola et al. (2022), Educational Psychology Review, 34, 363–399.

20. Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. Proceedings of the National Academy of Sciences, 107(5), 1860–1863.

21. Carter, S. P., Greenberg, K., & Walker, M. S. (2017). The impact of computer usage on academic performance: Evidence from a randomized trial at the United States Military Academy. Economics of Education Review, 56, 118–132.

22. Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. Educational Research Review, 25, 23–38. See also Sana, F., Weston, T., & Cepeda, N. J. (2013). Laptop multitasking hinders classroom learning for both users and nearby peers. Computers & Education, 62, 24–31.

23. Black, J. M., Myers, C. A., & Hoeft, F. (2015). The utility of neuroimaging studies for informing educational practice and policy in reading disorders. New Directions for Student and Adolescent Development, 2015(147), 49–56. On the general finding that neural change can predate and predict behavioral change, and on the divergence between behavioral and neuroimaging measures of learning.

24. 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. See also 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; and 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, 18(1), 279.

25. Vygotsky, L. S. (1978). Mind in Society: The Development of Higher Psychological Processes. Harvard University Press.

KAITLab · kaitlab.com · 2026

