This document was translated from English by machine and may contain mistakes. The English original is at https://www.novaracg.com/2026/07/30/order-of-operations-causal-reasoning-and-autistic-learners-what-actually-helps/.

Tell Me Why First: AI Tutors, Order of Operations and Autistic Learners

Heather Grizzle · July 30, 2026

AI x Learning, Vol 5, Issue 1: why autistic students stall on PEMDAS, what AI tutors get wrong, and what to ask before buying one.

1. Inhalt

1. Tell me why first
2. Theo and the homework helper
3. Where the trouble actually lives
4. Loving rules is not the same as obeying them
5. The acronym is the bug
6. Four moves with evidence behind them
7. The tier model, written out
8. Treat a wrong answer like a root cause problem
9. Read the error, not the score
10. A reason the student can hold
11. Practice conditions that hold up
12. When the tutor is a machine
13. What Ms. Alvarez needed
14. The evidence nobody has gathered yet
15. What happens when an autistic kid asks a chatbot why
16. Back at Pine Lake
17. This is a buying question
18. What families can do tonight
19. Questions teachers ask
20. Questions for the rest of this volume
21. The point underneath

AI X LEARNING | VOL 5, ISSUE 1

This piece opens Volume 5 of the AI x series. Volumes 2 through 4 followed AI into IEPs, vocational rehabilitation and Social Security. This one follows it into learning itself: the tutors, practice apps and homework helpers being sold to schools and families, and what happens to disabled learners when the machine doing the teaching was never built with them in mind. It starts with one small piece of math, because how a tool handles a single rule tells you a lot about how it handles the learner.

2. Tell me why first

I'm AuDHD, and my brain has one rule of its own when it meets a new rule: I need to understand the plan before I can act on it. Hand me a procedure with the reason stripped out and I don't just struggle to remember it. I distrust it. I keep poking at it, looking for the part that explains itself, and until I find that part the rule sits in my head like a piece of furniture nobody will tell me the purpose of. I can walk around it. I can even use it, carefully, the way you'd use a borrowed tool you don't trust. But I can't build on it, because I don't know what it's for.

I also learn by doing. Reading about a thing gets me partway there. Working it with my own hands, getting it wrong, seeing where it broke and fixing it, gets me the rest. That combination, needing the why and needing to touch the thing, is common among autistic and ADHD learners, and it's also the combination that ordinary classroom teaching is worst at serving. Most instruction is built for the student who will accept a rule on faith, practice it until it sticks and trust that the reason will turn up later. Some of us were never going to work that way. We aren't being stubborn. We're asking the question that makes the rest of the learning possible.

That's why the order of operations is such a good place to start this volume. PEMDAS looks like a small corner of the math curriculum, a week or two in the middle grades. For a lot of autistic students it's the spot where arithmetic stops being arithmetic and turns into a question about intent: why does one operation outrank another, and who decided that? A kid who can multiply flawlessly can still stall on $3 + 4 \times 2$, not because the multiplication is hard, but because the sequence looks arbitrary. Drilling the acronym harder doesn't fix that. Showing the reasoning underneath it usually does.

And it matters more every year, because the person explaining the rule is less and less likely to be a person. The kid stuck on order of operations at nine o'clock at night isn't waiting until Monday to ask a teacher. They're asking whatever AI tutor the district licensed, or the chatbot on a parent's phone. So this issue does two things. It lays out what actually helps autistic learners with this one rule, drawn from the research. Then it holds the AI tutors up against that standard and asks whether they meet it.

WHAT YOU'LL LEARN

- Why the order of operations creates a specific, predictable failure point for many autistic learners
- How memorizing a rule and understanding it come apart in this topic
- Four teaching moves with research behind them, and a fifteen-minute diagnostic
- What the best field evidence says happens when an AI tutor does the teaching
- The questions a school or family should get answered in writing before an AI tutor reaches a disabled learner

3. Theo and the homework helper

Let me introduce you to Theo. He isn't a real child. Like Nora, Eli and Jordan in Volume 2, he's a composite, built from patterns that show up again and again in research and in the lives of autistic kids and their families, and he goes to school in the same fictional district, Pine Lake.

Theo is twelve and in sixth grade. He's autistic, he reads two grade levels ahead, and he can tell you more about the New York City subway system than most people who ride it. His multiplication facts are fast and clean. His teacher, Ms. Alvarez, likes him and knows him well. She also has twenty-eight other students, and the week Pine Lake's math curriculum reaches the order of operations, she does what the pacing guide says. She writes PEMDAS on the board, teaches the mnemonic about a dear aunt named Sally, and assigns a practice set through the AI homework helper the district started licensing that fall.

Theo gets the first ten problems right. Then he hits $8 \div 2 \times 4$. He's learned that M comes before D, because that's the order of the letters, so he multiplies first, gets $8 \div 8 = 1$, and the home-

work helper marks it wrong. The correct answer is 16. Theo does the thing he always does when something doesn't add up. He asks why.

The helper tells him to remember PEMDAS. Theo points out, correctly, that he did. The helper explains that multiplication and division are done left to right. Theo asks why, then, they're two separate letters in a fixed order. The helper apologizes for the confusion, restates the rule in friendlier words and offers him a hint for the next problem. Theo closes the laptop.

Nothing in that exchange looks like a failure on a dashboard. The district's usage report will show that Theo logged in, completed eleven problems and received tutoring support. What the report won't show is that a kid who loves systems just learned that this system doesn't explain itself, and that the thing the district gave him to ask questions of can't answer the one question he actually has. Theo will be back at this rule next year, and the year after that in algebra, carrying an unexplained exception and a little less trust.

> **ABOUT THEO AND PINE LAKE**
>
> Theo, Ms. Alvarez and Pine Lake are composites created for this series. They are not real people or a real district, and the homework helper is not any specific product. The patterns they illustrate are drawn from the research cited throughout.

4. Where the trouble actually lives

Math ability in autistic students is uneven rather than uniformly low. Chiang and Lin reviewed the research and found that most autistic children and teens score in the average range in mathematics, with a subgroup showing both sharp strengths and sharp gaps.[1] Oswald and colleagues later reported that roughly one in five autistic students of average intelligence met criteria for a mathematics learning disability, and that the profile was driven less by calculation accuracy than by problem-solving and flexible strategy use.[2]

That second finding matters a lot here, because PEMDAS isn't a calculation. It's an agreement about how to read a written string of symbols, which makes it closer to reading comprehension than to arithmetic. A student can be fast and accurate with every single operation and still not know which one the sentence wants first. Theo is that student. His calculation is fine. His model of how the sentence is read is incomplete, and nobody has given him what he needs to complete it.

When we grade a student like Theo wrong and send them back to drill facts they already know, we've diagnosed the wrong problem. Worse, we've told them, without ever saying it, that the gap is in them and not in the explanation. Over enough years that adds up. Autistic students learn to distrust their own correct reasoning because it keeps running into rules nobody will justify.

> **THE CORE MISMATCH**
>
> PEMDAS is usually taught as a rule to obey. Many autistic learners process it best as a system to understand. When the reason behind the order is withheld, the rule becomes one more unexplained social convention to memorize, and memorized conventions fall apart the first time the problem looks unfamiliar.

5. Loving rules is not the same as obeying them

There's a stubborn assumption that autistic students like rules, so a rule-based topic should be easy for them. The research is more specific than that. Baron-Cohen's systemizing account describes a strong drive to work out how input, operation and output connect, meaning the mechanism underneath, not the surface instruction.[3] That matches how my own brain works. Give me the mechanism and I'll run with it. Give me the rule alone and I'll stall, because a rule without its mechanism gives a systemizer nothing to build on.

This is one of the most common misreadings of autistic learners, and it shows up far beyond math. A kid who insists on knowing why the class lines up by the door, or why the essay needs five paragraphs, or why you carry the one, gets labeled rigid or oppositional. Often they're neither. They're trying to load the rule into a model of how things work, and the rule arrived without the parts needed to load it. Give them those parts and the rigidity people complain about turns into the consistency every teacher wishes the rest of the class had.

That gap produces an error you can see coming from a mile off. The student applies PEMDAS strictly in the order of the letters, treating division as always coming after multiplication and subtraction as always coming after addition. On most textbook problems the answer still comes out right, so nobody catches it. Then this shows up:

$$8 \div 2 \times 4$$

The strict-acronym reading gives $8 \div 8 = 1$. The correct reading, left to right within the same tier, gives $4 \times 4 = 16$. The student didn't fail at arithmetic. They faithfully applied an incomplete model, which is what the teaching handed them. If you've ever been marked wrong for following the instructions you were given to the letter, you know what that feels like, and you know how fast it teaches a kid that the instructions can't be trusted.

6. The acronym is the bug

> A memorized rule is only as good as the model it stands in for.
>
> On the acronym problem

Here's a detail that tends to surprise adults: there isn't one order-of-operations acronym. Students in the United States usually learn PEMDAS. Students in the United Kingdom, India and much of the Commonwealth learn BODMAS or BIDMAS, which put Division before Multiplication. Canadian students often learn BEDMAS. If the letters themselves were the rule, students on different continents would be getting different answers to the same problem. They don't, because the letters were never the rule. They're a memory aid for a two-tier convention, and every version of the aid accidentally implies a ranking between operations that are actually equal.

For most students, that accident is harmless. They absorb the left-to-right rule from enough worked examples that the letter order fades into the background. For a student who takes a stated rule seriously and literally, the accident is the whole problem. PEMDAS says M then D. The teacher said to follow PEMDAS. So the student follows it, gets it wrong, and is told the rule doesn't mean what it says. That's not a student who failed to learn. That's a student who learned too faithfully from a tool that was never accurate.

You may have seen the internet fight over problems like $8 \div 2\,(2+2)$, where millions of adults split between 1 and 16. Part of the fight is a real ambiguity in how implied multiplication is written, which is why careful math writing avoids that notation. But a lot of it is grown-ups carrying the same strict-acronym model Theo has, decades later, because nobody ever replaced it. The acronym doesn't wear off. It just gets more confident.

7. Four moves with evidence behind them

King, Lemons and Davidson pulled together the math intervention research for autistic students and found that explicit, systematic instruction with modeling and structured practice

was the most consistently supported approach.[4] Root and colleagues built on that with modified schema-based instruction, which teaches students to recognize the underlying structure of a problem before they pick a procedure.[5] Both findings carry straight over to the order of operations, and none of this is exotic. It's mostly a matter of saying out loud the things we usually leave unsaid.

7.1 Four evidence-aligned moves

1 Teach the convention as an agreement, not a law

Say plainly that mathematicians chose this ordering so a written expression has one meaning and only one. Show a sentence read two ways that gives two different answers. Once the purpose is visible, the arbitrariness disappears.

2 Replace the acronym with a tier model

Teach two tiers of equal rank, multiplication with division and addition with subtraction, instead of six ranked letters. That removes the single most common structural error before it ever forms.

3 Read the structure before computing

Have the student mark the tiers on the expression before calculating anything. Structure first, arithmetic second, the same order modified schema-based instruction uses.

4 Cut the extra load while the rule is forming

Worked examples with faded steps, the same visual layout every time and no decorative context. Cognitive load research supports stripping out everything that isn't the target relationship while the student's mental model is still being built.[6]

The first move does more work than it looks like. Telling a student that order of operations is an agreement, something people decided so that everyone would read the same expression the same way, turns the rule from an arbitrary command into a solution to a problem. Autistic learners are often very good with solutions to problems. Show Theo that $3 + 4 \times 2$ could mean 14 or 11 depending on who's reading it, and that a world where it means both is a world where nobody can share math, and the convention stops being a decree. It becomes a fix he can respect.

The second move is the one that would have saved Theo's homework night. Two tiers, equal rank inside each, left to right. There's no letter order to misread because there are no letters. The third and fourth moves are about protecting that new model while it forms: look at the shape of the expression before touching the numbers, and keep everything else on the page quiet so working memory goes to the one relationship that matters.

> **WHAT TO AVOID**
>
> Word problems layered on top of a brand-new PEMDAS rule at the same time. Language demands and a new procedure compete for the same working memory, and the result usually gets read as a math failure when it's really a load failure.

8. The tier model, written out

Here's the whole convention restructured the way I'd want it handed to me: four rows, one direction rule per row, and nothing hiding in an acronym.

Order of operations as ranked tiers

Tier	Operations	Direction within tier
1	Grouping symbols (parentheses brackets braces)	Innermost outward
2	Exponents and roots	Left to right
3	Multiplication and division	Left to right (equal rank)
4	Addition and subtraction	Left to right (equal rank)

Source: Standard convention, restructured for instruction

Keeping Table 1 in front of the student during practice gives them an outside structure to check against, which takes the memory load off while the understanding settles in. It also gives the teacher and the student a shared language. Instead of "you did it wrong," the conversation becomes "which tier is that, and which direction does that tier run?" That's a question Theo can answer, and answering it is how the model gets built.

9. Treat a wrong answer like a root cause problem

Before I worked on sign language AI, I worked in quality assurance at the American Red Cross, where my job was problem management. That meant digging into the issues that kept coming back and finding out why, instead of patching the same symptom over and over. The discipline has a name, root cause analysis, and once you learn it you can't stop seeing it everywhere, including in a sixth grader's math homework.

A student who keeps getting PEMDAS wrong is a recurring issue. Re-teaching the acronym louder is patching the symptom. So is sending the student back to an app for more practice problems of the same kind, which is what most homework tools do by default when a student misses one. The useful question is which link in the chain is breaking, and you can usually find out in one sitting of about fifteen minutes.

1. **Test pure calculation**

   Give five single-operation problems. If these fail, the issue is arithmetic fluency, not PEMDAS, and the order of operations work should pause.

2. **Test the two-tier boundary**

   Give $3+4\times2$ and $10-6\div2$. Errors here mean the tier order itself hasn't formed.

3. **Test direction within a tier**

   Give $8\div2\times4$ and $12-5+3$. This isolates the strict-acronym error described above.

4. **Test nesting**

   Give $2\times(3+(4-1))$. Errors here point to working memory during nested evaluation rather than to the rule.

5. **Ask for the reason**

   Ask the student why multiplication comes before addition. An answer of "because PEMDAS says so" means the rule is stored without a model behind it, and move one of the framework applies.

Run this with Theo and the result is clear in about five minutes. Step one is clean. Step two is clean. Step three fails both problems in the same direction, and when Ms. Alvarez asks him in step five why multiplication comes before addition, he says "because it's earlier in PEMDAS," which is the most honest possible answer and the exact root of every mistake he made. Nothing about Theo needs fixing. One explanation does.

* * *

10. Read the error, not the score

Two students can write down the same wrong answer for completely different reasons, and the reason is what decides what you do next. A score tells you something went wrong. The written working tells you where.

Wrong answer

Start from the student's written working, not the final number.

→ **Single operation correct?**

Single operation correct?

Yes → Tiers applied? No → Build arithmetic fluency first

Tiers applied?

Yes → Direction within tier correct?

No → Teach the two-tier model with worked examples

Direction within tier correct?

Yes → Check nesting and working memory load

No → Correct the strict-acronym misconception explicitly

Check nesting and working memory load

Provide a written tier checklist and allow step-by-step recording rather than mental evaluation.

→ **End**

This is also the first place AI tools tend to fall short, and it's worth naming now because we'll come back to it. Most practice software sees the final number. It knows 1 is wrong and 16 is right. Some newer AI tutors can read a student's typed working, but whether they diagnose

the type of error, or just notice that an error happened and offer a hint, depends entirely on how they were built and instructed. A tool that can't tell the strict-acronym error from an arithmetic slip will keep serving the same remedy to two students who need opposite things.

11. A reason the student can hold

The strongest single explanation I know is that operations differ in how much they pack inside one symbol. Multiplication is repeated addition, so one times sign carries a whole bundle of additions. Unpacking the bigger bundle first keeps the meaning of the sentence intact.

$$3+4\times 2=3+(2+2+2+2)=11 \tag{1}$$

Written out like Equation 1, the order stops being a decree. The student can now work out the rule for themselves instead of recalling it, and a rule you can rebuild from scratch survives an unfamiliar problem in a way a memorized one doesn't. For a learner like me, that's the whole difference between a rule I'll use and a rule I'll quietly route around.

The same idea extends upward. An exponent is repeated multiplication, a bigger bundle still, so it gets unpacked before multiplication does. Parentheses are the writer's way of saying "treat this as one thing," so they come first of all. Laid out that way, the tiers aren't a list to memorize. They're a ladder from the most packed symbols to the least packed, and a student who understands the ladder can climb it in any problem, including ones they've never seen.

> **Causal transparency**
>
> How clearly a learner can see why a procedure produces the result it produces. High causal transparency supports transfer to unfamiliar problems. Low transparency supports only recognition of formats the student has already practiced.

12. Practice conditions that hold up

- ☐ Same visual layout for every expression while the rule is being learned
- ☐ Tier reference chart visible and allowed at all times

- ☐ Working shown line by line, one operation removed per line
- ☐ Errors talked through by tier, never just marked wrong
- ☐ Word problems introduced only after the procedure is solid in bare numbers
- ☐ The student's own interests used for practice sets wherever possible

> **ONE SMALL CHANGE WITH A BIG PAYOFF**
>
> Ask the student to circle the tier-three operations before computing anything. It's a ten-second habit, and it removes most direction-within-tier errors.

That last item on the checklist deserves more than a line. Interest-based practice isn't a reward or a bribe. For many autistic students, a deep interest is where their attention and working memory are most available, so practice built on it costs less. Theo will grind through thirty expressions about subway transfers that he'd abandon after five about apples. A good human teacher knows that about Theo. Whether an AI tutor can know it, and whether anyone has told it, is one of the questions this volume will keep asking.

13. When the tutor is a machine

Everything above is about human teachers, and it matters more now, because more and more of this teaching is being handed to software. Theo's homework helper isn't hypothetical in kind, only in name. Districts are licensing AI tutors and practice tools at a pace that has outrun the evidence about what they do to learning, and families are filling the gaps at home with general-purpose chatbots that were never designed to teach anybody anything.

The best evidence so far on what happens next comes from a field experiment with nearly a thousand high school math students in Turkey, published in PNAS in 2025.[7] Researchers gave some students a standard ChatGPT-style tutor and others a version with guardrails built to protect learning, one that coached instead of handing over answers. During practice, both groups did better than students without AI: 48 percent better with the standard tool and 127 percent better with the guarded one. Then the AI was taken away for the exam. Students who had used the standard tool did 17 percent worse than students who never had AI at all. Students who had used the guarded tutor did about the same as the control group.

The researchers' read was that without guardrails, students used the AI as a crutch. They got through the practice and learned less. Sit with what that means for a school that judges a tool by its dashboard. The tool that harmed learning produced better practice scores than no tool at all. Every report a district administrator would normally look at said it was working. The harm only showed up when the support was gone and the student had to stand on their own, which is the one moment no usage report measures.

It's worth being careful about what the study does and doesn't show. It was one subject, one school, one country and a handful of sessions, with typical students rather than students with disabilities. It doesn't prove that AI tutors harm learning in general, and the guarded version didn't. What it shows is that design decides the outcome, and that the design choice that makes a tool feel most helpful in the moment, just answering the question, is the one most likely to hollow out the learning underneath.

14. What Ms. Alvarez needed

It would be easy to read Theo's first homework night as a story about a bad teacher. It isn't. Ms. Alvarez did what her pacing guide told her to do, with the materials her district bought, in the week the calendar gave her, for twenty-nine kids at once. She knows Theo asks why. She just didn't have the time, the training or the tools to answer him the way he needed, and the AI homework helper was supposed to be the thing that filled that gap.

That's the promise every AI tutor is sold on: the patient, individual attention a teacher with twenty-nine students can't give. It's a real need, and it's not a crazy promise. But a tool can only fill a gap it was designed to see. If nobody told the tool that some students need the reason before the rule, need less language rather than more and will push back hard when something doesn't add up, it fills the gap with the same explanation that left the gap in the first place, only faster.

What Ms. Alvarez needed was not a better chatbot in the abstract. She needed three things the district could have given her: an explanation of order of operations built on the tier model instead of the acronym, a tool configured to teach that same explanation, and fifteen minutes and a diagnostic to find out which kids were carrying the strict-acronym error before it hardened. None of those cost much. All of them depend on somebody upstream deciding that students like Theo were part of the plan from the start.

15. The evidence nobody has gathered yet

Here's the uncomfortable part. Almost everything we know about AI tutors and learning, including the study above, comes from typical students. The research on what helps autistic learners with math was done with human teachers and paper worksheets. The two bodies of evidence have barely touched each other. So a district buying an AI tutor for a classroom with autistic students in it is, in most cases, buying something nobody has tested on the students most likely to be failed by it.

That gap isn't anyone's secret. It's just the ordinary way products get built. A company tests on the users it can recruit easily, measures what it can measure cheaply, usually engagement and practice scores, and ships. Disabled learners are a smaller market, harder to recruit for studies and more expensive to test with properly, so they get folded into the average. The average then gets reported as if it described everyone.

I've spent this year doing the same kind of work on sign language AI, and the pattern is identical. A tool gets demonstrated to people who aren't its hardest users, the demo goes well, the purchase happens, and the people the tool was supposed to serve find the failures in daily life, one by one, with no channel to report them. The AI x series exists because that pattern keeps repeating across every public system AI is entering. Learning is just the place where the people finding the failures are children.

What would good evidence look like here? Not complicated. Take a tool, give it to autistic students and a comparison group, and measure what they can do without it a few weeks later, not just how they did with it. Ask their teachers what the tool explained and whether the explanation was right. Publish the results, including the bad ones. Any vendor willing to do that would stand out immediately, which tells you something about how rarely it happens.

16. What happens when an autistic kid asks a chatbot why

Now put Theo into that picture, because the general finding gets sharper for him in at least three ways.

The first is the answer machine. An AI that simply tells Theo $8 \div 2 \times 4 = 16$ gives him a number and no model. He can copy it, finish the set and learn nothing about why his own reasoning failed. For a student whose whole problem is a missing explanation, a tool that skips explanations is the classroom problem made faster and available all night.

The second is the acronym at scale. A lot of AI tutoring sounds like the average of a million worksheets, because in a real sense it is. If the most common explanation of order of operations on the internet is "remember PEMDAS," then an AI tutor's most likely explanation is "remember PEMDAS," delivered warmly and confidently to every student who asks, every night, in every district that licensed it. A human teacher who teaches the acronym badly reaches one classroom. A tool that does it reaches all of them, and nobody is watching the explanation it gives, only whether the student logged in.

The third is how language models handle disagreement. Researchers have documented a tendency in these systems toward sycophancy: shifting their answers to agree with the person they're talking to, even when the person is wrong.[8] Picture Theo pushing back the way he does, insisting that M comes before D because the rule said so. A tutor tuned to be agreeable may apologize, hedge or, worst case, concede a point it shouldn't. For a student who is testing whether the system is trustworthy, a tutor that folds under pressure is about the most damaging thing it could do. It teaches him that the explanation was never solid in the first place.

There's a quieter fourth problem, too. When AI tutors try to be helpful, they tend to reach for analogies, stories and lots of words. Pizza slices, a line at a coffee shop, a team of workers taking turns. For many students that's warm and useful. For an autistic student in the middle of building a new model, it can be pure extra load: more language to process, a metaphor to decode literally, a story with details that don't map onto the math. The research on cognitive load says to strip everything that isn't the target relationship while the schema forms. An eager chatbot does the opposite, and it does it because it was built to sound helpful to the average user.

None of those failures show up in a demo. A sales demo uses a cooperative user asking a clean question. They show up on the exam, weeks later, in a kid who got every practice problem right.

17. Back at Pine Lake

Let's run Theo's homework night again, differently.

This time Pine Lake's tutor was set up before anyone bought it. The district asked how it handles a student asking why, tested it with a few of its own autistic students and their teachers, and configured it to coach rather than answer. When Theo gets $8 \div 2 \times 4$ wrong, the tutor doesn't give him 16. It asks him to show his working. It sees that he multiplied first, and



instead of restating the rule, it tells him the reason: multiplication and division are the same kind of step, so neither outranks the other, and when two steps have the same rank you read them the way you read words, left to right. Then it gives him $12 \div 3 \times 2$ to try the idea out.

Theo pushes back. He says PEMDAS puts M first. The tutor doesn't fold. It tells him the letters are a memory trick, that the British version of the trick puts D first, and that both versions mean the same thing: those two steps share a rank. Theo goes quiet for a minute. Then he asks whether addition and subtraction work the same way, and the tutor says yes and lets him prove it to himself on two more problems.

Nothing magical happened there. The tool did what the research says good teaching does: explained the mechanism, held its ground, kept the load low and let the student do the work. The difference between the two homework nights wasn't the technology. It was whether the adults who bought the technology asked the right questions first.

18. This is a buying question

That makes AI in learning a procurement question as much as a teaching one, and that's the lens this whole volume will use. When a district licenses a tutor, it's buying a teacher it can't interview, can't observe and can't easily correct. When a family downloads one, they're doing the same with even less information. Either way, the product decisions that determine whether a disabled learner is helped or quietly harmed were made by somebody else, usually with typical students in mind, and almost never published.

If a public school district provides the tool, it's the district's program, and the district's civil rights obligations come with it. Title II of the Americans with Disabilities Act and Section 504 of the Rehabilitation Act don't stop applying because the instruction came from a vendor's server. A tool that a district requires students to use has to work for the students who have to use it, and "the vendor didn't test with autistic kids" is not a defense anyone should want to make.

Before a school or a family puts an AI tutor in front of a disabled learner, someone should be able to answer a few plain things in writing.

☐ Does the tool give the answer, or coach toward it, and can a teacher set which?

☐ When a student asks why a rule exists, does it explain the reason or restate the rule?

☐ Which explanation of order of operations does it teach, the acronym or the tiers?

☐ Does it hold a correct answer when a student pushes back, or does it fold?

☐ Can it read a student's working and name the type of error, or only mark the final number?

☐ Can a teacher turn down the analogies and extra language for a student who needs less load?

☐ Has it been tested with autistic and other disabled learners, and where are the results?

☐ Who checks its math explanations for errors, and how does a teacher report one?

Most of those questions can be answered in an afternoon with a test account and a list of hard prompts, the way Ms. Alvarez might test a new student teacher. Ask it why multiplication comes first. Tell it it's wrong when it's right. Make the error Theo made and see whether it notices what kind of error it was. If the vendor can't or won't let a district do that before signing, that tells you what you need to know.

19. What families can do tonight

Not every family has a district that asked the right questions. A lot of AI tutoring happens at a kitchen table, on a parent's phone, with whatever chatbot is closest. Here's what I'd want a parent of a kid like Theo to know.

1. **Ask it the hard question yourself first**

   Before your child uses an AI tool for math, ask it why multiplication comes before addition, and why multiplication and division are done left to right. If it just says "because of PEMDAS," it will say that to your kid too.

2. **Tell it how to teach**

   Most general chatbots will follow instructions. Tell it not to give final answers, to explain the reason behind every rule, to keep explanations short and plain with no stories or analogies, and to hold its answer if your child disagrees and it's right. Save those instructions so you don't have to retype them.

3. **Check the learning without the tool**

   Every so often, give your child a few problems with no AI at all. If they can do them, the tool is helping. If they can only do them with the tool open, it's being used as a crutch.

4. **Write down what goes wrong**

   When the tool explains something wrong or confuses your child, take a screenshot. If the tool came from the school, those screenshots are evidence your child's teacher and IEP team can use.

None of that should be the family's job. A child's right to an education that works for them doesn't depend on whether their parents know how to prompt a chatbot. But until the people buying these tools do the testing, families are the ones closest to the failures, and a parent who knows what to look for is the fastest early warning system a kid has.

20. Questions teachers ask

+ Should I stop teaching the acronym entirely?

Not necessarily, but demote it. Use it as a memory aid after the tier model is understood, never as the main explanation. The acronym encodes a false ranking between multiplication and division, and that misinformation is behind a large share of the errors that stick around.

+ My student computes correctly but can't explain the rule. Is that a problem?

It predicts trouble later. Getting the procedure right without understanding why tends to hold until the expressions get unfamiliar, then falls apart fast. Deal with it while the arithmetic is still going well.

+ Is this an autism-specific approach?

The approaches come from research with autistic students, but explicit instruction, reading structure first and cutting extra load help mixed classrooms too. Nothing here has to be delivered separately.

+ How long should the tier chart stay available?

Until the student stops looking at it. Pulling a support on a schedule instead of on evidence that the student is independent brings back the memory load the support was there to remove.

+ My district already licensed an AI tutor. What can I do as one teacher?

Test it yourself with the hard prompts above before you assign it to a student who needs the why. Write down what it said. If it fails, you have something concrete to bring to your department or your district's technology office, and you have a reason to keep that student on human-explained practice in the meantime.

+ A parent says their autistic child does better with the AI than with me. Should I worry?

Not necessarily. Some autistic students genuinely prefer a patient, private, text-based tutor, and that preference is real and worth respecting. The question is what the student can do when the tool is gone. Check that the learning holds without it, the way the Turkish study did, before you treat practice scores as proof.

+ Can an AI tutor be written into an IEP?

A specific tool can show up as an accommodation or supplementary aid if the team decides it serves the student, but the team should know what it's choosing. A tool named in an IEP is a tool the district has committed to, so the questions in the buying checklist matter even more there, and the team should be able to say how it will know the tool is helping.

---

21. Questions for the rest of this volume

Like the earlier volumes, this one is meant to be used, not just read. Each issue will end with questions for the people who buy, assign and live with these tools. These are the ones this opening issue leaves on the table.

☐ If a tool raises practice scores while lowering what students can do on their own, who in a district would ever find out?

☐ What does it mean for a disabled learner to consent to an AI tutor, and who gives that consent for a twelve-year-old?

☐ When an AI tutor explains a rule wrong to every student in a district, whose error is it: the vendor's, the district's or the teacher's who assigned it?

☐ Should AI tutors be required to disclose what explanations they give for core skills, the way a textbook can be read before it's adopted?

☐ Who should test these tools with disabled learners before purchase, and who pays for that testing?

☐ What would it take for a family to find out, after the fact, what an AI tutor actually told their child?

The rest of Volume 5 takes those questions into other corners of learning: AI reading tutors and the children whose speech they mishear, captioned and signed lessons generated by machines, chatbots used as study partners by students with ADHD, and the data these tools collect about how disabled kids learn. Each issue starts the same way this one did, with one small, specific thing a tool does, and follows it out to what it means for the learner.

22. The point underneath

The trouble autistic students have with PEMDAS is rarely about arithmetic and rarely about capacity. It's about a teaching habit of handing over a convention without its justification, which asks the learner to trust a rule they aren't allowed to inspect. Some of us were never going to do that. We weren't being difficult. We were asking the one question that makes math make sense, which is why.

Answer it, and the topic opens up, sometimes fast. Refuse to answer it, or answer it badly at scale through a machine nobody tested, and you teach a kid something far bigger than a wrong rule. You teach them that the systems around them don't explain themselves, and that asking is pointless. Theo deserves better than that, and so does every kid who closes the laptop instead of asking again. The rest of this volume asks the same question of the machines: do they answer why, or do they just answer?

> **TIPS**
>
> Know something about a sign language AI system that the public should know? Send it to tips@novaracg.com. Every tip is read by me and kept confidential. If you work in the UK, writing to us is not by itself a protected disclosure under the Public Interest Disclosure Act 1998, so before you share employer information, check your options with Protect (protect-advice.org.uk), the free whistleblowing advice line.

23. Notizen

1. Chiang, H. M., & Lin, Y. H. (2007). Mathematical ability of students with Asperger syndrome and high-functioning autism. Autism, 11(6), 547 to 556.
2. Oswald, T. M., Beck, J. S., Iosif, A. M., McCauley, J. B., Gilhooly, L. J., Matter, J. C., & Solomon, M. (2016). Clinical and cognitive characteristics associated with mathematics problem solving in adolescents with autism spectrum disorder. Autism Research, 9(4), 480 to 490.
3. Baron-Cohen, S. (2009). Autism: the empathizing-systemizing (E-S) theory. Annals of the New York Academy of Sciences, 1156, 68 to 80.
4. King, S. A., Lemons, C. J., & Davidson, K. A. (2016). Math interventions for students with autism spectrum disorder: A best-evidence synthesis. Exceptional Children, 82(4), 443 to 462.
5. Root, J. R., Browder, D. M., Saunders, A. F., & Lo, Y. (2017). Schema-based instruction with concrete and virtual manipulatives to teach problem solving to students with autism. Remedial and Special Education, 38(1), 42 to 52.
6. Sweller, J., van Merrienboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. Educational Psychology Review, 31, 261 to 292.
7. Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. Proceedings of the National Academy of Sciences, 122(26), e2422633122.
8. Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askell, A., Bowman, S. R., et al. (2023). Towards understanding sycophancy in language models. arXiv:2310.13548.

23.1 Zitierweise

Grizzle, H. M. (2026, July 30). Tell Me Why First: AI Tutors, Order of Operations and Autistic Learners. Novara Consulting Group. https://www.novaracg.com/de/2026/07/30/order-of-operations-causal-reasoning-and-autistic-learners-what-actually-helps/