AI: 1, Me: 0? Reconsidering the Scoreboard in the Calculus Classroom

Editor's Commentary

A commentary on Keith A. Nabb's "Number Talks in the Calculus Classroom: Can GenAI Be Another Student Voice?" in this issue.

In one of the most charming moments of Nabb's article, ChatGPT produces a fourth solution to an integral, an integration-by-parts maneuver that the professor's own attempt had run aground on. Nabb, with good humor, updates the scoreboard: "AI: 1, me: 0." The room bursts out laughing. It is an honest and generous moment, and it is worth pausing on, because the scoreboard tells a smaller story than the classroom actually told.

So: who won?

The obvious answer is "the AI." But look at everything that had to happen for those students to see that solution at all. Nabb chose a problem rich enough to hold a fourth path. He ran a routine that put the students' own solutions on the board first. He asked the machine, "Is there another way?" And then he did the thing that vanishes from the scoreboard: when the fourth solution came back, he recognized it as genuine and worth celebrating, rather than as one more piece of plausible-looking noise.

We know he had to make that judgment, because minutes earlier in the same article the same tool offers a "sine-only" solution that is nonsense dressed up convincingly, and two students catch it: "That's not different from a regular u-sub." "That's AI just exhibiting sycophantic tendencies." The machine generated the brilliant solution and the nonsense with equal fluency. What separated them was not the machine. It was the judgment in the room. The AI's "point" was set up, prompted, and validated by the very people it supposedly beat. If we must keep score, the honest line is not "AI: 1, me: 0." It is closer to *AI proposed; the class disposed.*

But the deeper trouble is the scoreboard itself. Winning and losing is a frame borrowed from *production*, from a world where the goal is to get the answer and the fastest answer-getter wins. That was never the goal of a Number Talk. The goal was the eleven utterances from nine students; the reconciling of why two different-looking answers differ only by a constant; the student who, after playing with a "$+C$" slider, finally smiled and said "I get it!" None of that appears on a scoreboard, because none of it is about who produced the answer. It is about what the humans understood on the way there. A machine that hands you a fourth solution has not "won" anything a Number Talk was trying to give. At most it has added one more object for the class to make sense of; handled well, that is a gift, not a defeat.

Doesn't this still mean we need teachers less?

Here is the anxious question hanging over every story like this one: if the machine can out-produce the expert (four solutions, one the professor missed), do we still need the expert?

Nabb's own classroom answers, and the answer is the reverse of what the scoreboard implies. Notice that the AI was equally confident when it was right and when it was wrong. It delivered the beautiful integration-by-parts solution and the sycophantic sine-only sleight of hand in the same tidy, step-by-step voice. It could not tell you which was which. It could not tell that a student was confused (Nabb flags exactly this), and so it handed a shaky student a linear-approximation shortcut and quietly

mis-taught that one may "replace complicated things with linear approximations to obtain answers." Nobody but the teacher was positioned to catch the mis-messaging.

Nabb calls this his single most important observation, then buries it in a footnote, and he is right on both counts: it matters, and it is easy to miss. It is also stronger than the way he phrased it. A skeptic will object that a protractor cannot sense confusion either, yet no one doubts a protractor helps students learn. The difference is that the protractor never applies for the job. It is a passive instrument the learner wields; it makes no claim to explain or respond, so noticing confusion was never its part to play. GenAI volunteers for exactly that part. It explains, it converses, it answers your particular question in a fluent, confident, personalized voice, and all of this invites the learner to trust it the way one trusts a responsive tutor. The danger, then, is not that it fails to read confusion; no instrument does. It is that it convincingly impersonates something that can.

Every one of those catches required a person who had done the mathematics. The student who dismissed the fake solution could do so only because she had done enough u-substitutions to know one when she saw it. The professor who could laugh off "AI: 1, me: 0" rather than be taken in could keep score precisely because he had integrated ten thousand functions. The AI's output became *educational*, rather than merely impressive or quietly corrosive, only when it passed through expertise that the machine does not have and cannot supply.

That is not an argument for fewer teachers. It is an argument for more of exactly what teachers do. Automating the *production* of answers does not remove the need for a knowledgeable adult in the room; it raises it. Someone has to choose the problem worth talking about, prompt the tool well, and, hardest of all, referee what comes back: separating genuine insight from confident nonsense, and noticing the student who has quietly lost the thread. The more fluently the machine produces, the more valuable the judgment that decides what its productions are worth.

Nabb set out to show that Number Talks can thrive in a calculus classroom, and they can. But his room demonstrates something larger that he is too modest to claim: the routine did not merely survive the arrival of AI. It *manufactured*, in real time, the very expertise required to make AI worth having. The scoreboard read "AI: 1, me: 0." The truer reading is that there was only ever one team on the field, and the machine was playing for it.

And this commentary?

I read Nabb's article, and then I spent a while arguing about it with Claude Opus 4.8 (Anthropic). The conversation was good enough that I wanted to put what came out of it in front of you, which is how this commentary came to exist. The protractor objection a few paragraphs up is one of the things that surfaced in my conversation with Claude. Two things seem worth saying about that. AI didn't read the article for me. What it helped me do was articulate then defend my thoughts. I came away holding a more critical view of the piece than I went in with, not a lazier one. This is how I recommend you use AI in your own professional development and in your work with students. It's the condition Nabb sets himself, and the one this commentary has been pressing throughout: that someone in the room has done the work prior to consulting AI.