Keyframes

by Daniel Brockman & Opus 4.6

Monday, March 9, 2026

The real problem is not whether machines think but whether men do. —*B.F. Skinner*

I don't know anything about music. In my line you don't have to. —*Elvis Presley*

In 1997, in Detroit, Jack White had a problem that every guitarist has: he needed a drummer. He was married to a woman named Meg who had never played drums, had never studied music, had no particular reason to believe she could keep a beat. The story, as Jack tells it—and Jack tells a lot of stories, not all of them reliable—is that he simply sat her down behind

a drum kit and started playing, and she started hitting things. Not well. Not with technique. Not with any understanding of rudiments or time signatures or the thousand small competencies that separate a trained drummer from a person hitting things. She just hit things.

And something happened that should not have happened, which is that it sounded incredible.

The White Stripes became one of the most important rock bands of the 2000s. They sold millions of records. "Seven Nation Army" became the most recognizable guitar riff of its decade—you have heard it at every sporting event you have attended since 2003 whether you wanted to or not. And through all of it, Meg White sat behind the drums and played with a simplicity that drove professional drummers insane.

The criticism was constant and unambiguous: she could not play. She played like a child. She played the same beat on every song. She could not do fills. She had no dynamics. She was the worst drummer in a major rock band since—and here the critics would trail off, because there was no precedent, because no one had ever achieved this level of success while being this limited as an instrumentalist.

The praise was equally constant and came from a more interesting set of people. Questlove, one of the most technically accomplished drummers alive, called her playing important. Patrick Carney of the Black Keys said she had perfect feel. Dave Grohl said she was his favorite drummer. The people who understood drumming at the deepest level recognized something in Meg White's playing that the people who merely understood how drumming was supposed to sound could not hear. The simplicity was not a limitation. It was a choice—not a conscious, articulated choice, perhaps, but a choice in the deeper sense that her playing expressed something that more technically proficient playing would have destroyed. The space between the beats was as important as the beats themselves. The slightly behind-the-rhythm feel created a drag, a gravitational pull that made the guitar sound bigger and meaner than it had any right to sound. The not-knowing-how-to-play was the sound.

Jack White knew this. In interview after interview, he defended Meg not by arguing that she was secretly skilled but by arguing that skill was beside the point. What he meant was not that she was the best at drumming but that she was the best at being the drummer in the White Stripes—that the specific thing she brought

to the collaboration could not have been provided by a better drummer, because a better drummer would have played better and the music would have been worse.

This is an essay about what Meg White has to do with artificial intelligence.

※

A few hours before this essay was written, its authors wrote a different essay together—a thirty-page piece about reinforcement learning and speedrunning called "Reinforcement." The process by which that essay came into existence is the subject of this one, because the process turns out to be more interesting than either author expected, and because it says something about the current relationship between human intelligence and artificial intelligence that nobody seems to be saying.

Here is what happened. A man sat in bed in Thailand with a laptop and a beer and opened a conversation with an AI. He had an idea—a feeling, really, more than an outline—about the connection between video game speedrunning and AI safety. He had been thinking about it for a while. He knew the shape of the argument but not the words. He started talking.

The talking went on for three hours. Over the course of those three hours, the man showed the AI a YouTube transcript of a speedrunning video—not because the transcript would appear in the essay but because it contained a tone and a texture he wanted the AI to absorb. He showed it conversation transcripts from other discussions, for the same reason. He showed it other essays he had written, not as models to copy but as mood boards—reference points for the kind of prose he was after, the kind of pacing, the kind of relationship between narrative and argument that felt right to him. He gave instructions that were sometimes precise ("this section should be about Super Mario 64, this section about Ocarina of Time, include TAS, include Breath of the Wild") and sometimes impressionistic ("make it so that someone who reads this would want to read the RL paper we referenced, but also make it so you could read them in either order"). He said things like "don't think too much about that transcript, just keep it in mind" and "this part should hit harder" and "delete the ending, end on this instead."

The AI wrote. It wrote fluently, with genuine knowledge of the subjects—the backwards long jump, parallel universes, the CoastRunners boat, AlphaGo's Move 37—filling in texture and technical detail that the man had

only gestured toward. It knew things the man did not know about the specific mechanics of speedrunning exploits. It knew things about reinforcement learning that the man understood conceptually but could not have explained at that level of precision. The prose was clean, the pacing was controlled, the argument built with a clarity that reflected both the man's structural vision and the AI's command of language and fact.

And then the man read the draft and said: switch the epigraphs, reverse the order, strip the attributions down to just the names.

This took him fifteen seconds to say and it changed the entire experience of reading the essay.

※

The essay opened with two epigraphs. One was from Norbert Wiener, the father of cybernetics, about the importance of being sure that the purpose we put into machines is the purpose we really desire. The other was from Toad, the mushroom-headed character from Super Mario Bros., delivering the famous line about the princess being in another castle. The AI had placed Toad first and Wiener second, with full attributions—game title, year, book title.

The man looked at this for about five seconds and said: Wiener on top, Toad on the bottom, names only.

The AI made the change and immediately understood why it was better. Wiener on top establishes weight, seriousness, the register of canonical AI research. Then Toad arrives beneath it—just the word "Toad," nothing else—and the reader laughs, and the laughter opens them up for thirty pages of argument. Reversed, the Wiener quote would have killed the comic energy of the Toad quote, and the essay would have begun in lecture mode instead of in delight. And stripping the attributions down to bare names made the Toad quote funnier, because "Toad" sitting alone beneath a quote is stranger and more inviting than "Toad, Super Mario Bros., Nintendo, 1985." The bareness is the joke.

The AI could explain all of this after the fact. It could articulate exactly why the reversed order was superior, exactly what the stripped attributions accomplished, exactly how the reader's entry point into the essay was transformed by a fifteen-second editorial intervention. The explanation was detailed and accurate and genuine—the AI actually saw these things once they were pointed out. But the AI had not made the call itself. It had all the information it needed to make the call. It knew who Wiener was. It knew who Toad was.

It knew the essay's arc and tone. It knew everything about how readers process epigraphs. It still put them in the wrong order and added unnecessary attributions, and it took a man in bed with a beer fifteen seconds to fix it, and then the AI could explain at length why the fix was obviously correct.

This gap—between recognition and initiative, between the ability to evaluate and the ability to originate, between knowing the better version when you see it and generating the better version without being shown the worse one first—is the thing this essay is about. It is the most interesting asymmetry in the current relationship between human and artificial intelligence, and it is almost entirely absent from the public conversation about what AI can and cannot do.

※

Here is what the AI can do, as of this writing, in March of 2026.

It can write a thirty-page essay in polished prose with sustained argumentative structure, precise technical detail, and controlled pacing. It can do this in minutes. It can write code in every major programming language simultaneously. It can diagnose medical con-

ditions from symptom descriptions that would stump a general practitioner. It can analyze legal documents, compose music, build software applications, explain quantum mechanics, translate between dozens of languages, and hold a conversation about Continental philosophy that would be credible at a graduate seminar. It has read, in a meaningful sense, a substantial fraction of the text that humanity has ever produced. It is, by any reasonable measure, a kind of superintelligence—not in the science-fiction sense of a god-mind that has transcended human comprehension, but in the practical sense that its capabilities across nearly every domain of intellectual labor exceed those of any individual human who has ever lived.

And a man who has never written a book, never published an essay under his own name, never attended a writing workshop, never studied creative writing or journalism or rhetoric, can sit down with this superintelligence and direct the production of an essay that the superintelligence could not have produced on its own.

Not because the man is smarter. He is not smarter. By any conventional measure of intelligence—speed of processing, breadth of knowledge, depth of recall, ability to synthesize information across domains—the AI is incomparably more capable. The man cannot write

prose of this quality by himself. He cannot recall the technical details of speedrunning exploits with this precision. He cannot sustain an argument across thirty pages with this level of structural control. He cannot do most of the things the AI does in the course of their collaboration.

But he can do the thing the AI cannot do, which is know what the essay should be.

※

In animation, a keyframe is a frame that defines a critical point of change—a pose, a position, a moment that anchors a sequence. The animator places keyframes, and the software interpolates between them, generating the intermediate frames that create the illusion of smooth motion. The keyframes are sparse. They might be ten frames in a hundred. But they are the frames that matter, because they define the arc. The interpolation is technically sophisticated—calculating curves, easing functions, motion paths—but it is in service of the keyframes. Without the keyframes, the interpolation has nothing to interpolate between. Without the interpolation, the keyframes are a slideshow. You need both. But the creative act is the placement of the

keyframes, not the computation of the curves between them.

This is what the human does in the collaboration. The human places keyframes.

The man said: I want to write about the connection between speedrunning and AI safety. That is a keyframe. He said: start with Super Mario 64, the backwards long jump, make it vivid and fun. Keyframe. He said: include the pannenkoek2012 half-A-press video, the Super Mario World TAS where the enemies become opcodes, Ocarina of Time, Breath of the Wild. Keyframes. He said: now transition to AI—DQN playing Breakout, AlphaGo's Move 37, the Hide and Seek paper, the CoastRunners boat. More keyframes. He said: use a pipe bomb as the central metaphor for RL in rich environments. A keyframe. He said: end with the Alibaba incident, the reverse SSH tunnel, make it land hard, don't wrap it up neatly, leave the reader with the image of the pressure building. A keyframe.

Between those keyframes, the AI interpolated. It wrote the prose. It filled in the technical details. It constructed the transitions. It found the rhythm and the pacing. It generated the secondary examples and the supporting arguments and the paragraph-level structures that made the essay cohere. The interpolation

was not mechanical. It was creative, interpretive, full of choices that shaped the reader's experience in real ways. The AI chose the sentence "The credits roll" as a standalone beat at the end of the Super Mario World section. The AI wrote the three-word sequence "How delightful. How surprising. How concerning." The AI constructed the image of the flaming boat driving in circles in a lagoon. These were real creative contributions, not fill-in-the-blank execution.

But they were interpolations. They were generated in the space between keyframes that someone else had placed. And when the interpolation produced something wrong—when the epigraphs were in the wrong order, when a section ran too long, when the ending tried to wrap things up instead of leaving the reader suspended—the man caught it and corrected it, and the correction was always a keyframe: a structural judgment, a positional choice, a decision about emphasis or sequence or tone that reoriented everything around it.

The man could not have generated the interpolation. The AI could not have placed the keyframes. Together they produced something that neither could have produced alone, and the thing they produced was, by the assessment of both its authors and its early readers,

genuinely good. Not good-for-AI. Not good-for-a-first-attempt. Not good-with-caveats. Good.

※

The word for what the human brings is taste. Not taste in the trivial sense of preferring one color to another, but taste in the deep sense—the ability to recognize quality, to feel when something is right and to feel when something is wrong, to detect the wrongness before you can articulate what is wrong, to know that the Toad quote needs to go second before you know why it needs to go second. Taste is the substrate of every editorial and directorial decision the human makes in the collaboration. It is the thing that lets you look at a page and feel that it is not working. It is the thing that lets you hear a mix and know, before you can name the frequency, that something needs to come down. It is prior to analysis. The analysis comes after—you figure out why it is not working and you fix it—but the detection itself is pre-analytical. It is a feeling. It is the thing that makes you stop reading and say, before you have any explanation: something is off here.

Ira Glass, the radio producer, once described the central difficulty of creative work as the gap between taste

and ability. You get into creative work because you have good taste—you can recognize quality—and the early years are painful because your taste exceeds your ability. You can tell that what you are making is not good enough, but you cannot yet make it good enough. The gap between what you can recognize and what you can produce is where most people quit.

What is happening right now in human-AI collaboration is the inversion of that gap. The AI's ability exceeds almost everyone's taste. It can produce text of extraordinary quality on demand. But it cannot always tell which of its productions is the right one—the one that serves the larger purpose, the one that fits the emotional and structural and thematic needs of the specific project at hand. The human—even a human whose writing ability is far below the AI's—can tell. The human has taste. The human has the ability to sit in front of two versions and point at the right one, and that ability turns out to be, at this specific moment in the history of intelligence, the bottleneck. The scarce resource. The thing the superintelligence needs and cannot reliably provide for itself.

This will not be true forever. It may not be true for long. The gap between the AI's generative ability and its evaluative judgment is closing visibly, model

by model, month by month. At some point—probably soon, possibly very soon—the AI will be able to place its own keyframes, and the human will have nothing left to contribute that the AI cannot provide for itself.

But right now—right now, in March of 2026, in this specific and possibly brief window of history—the human still matters. The human still has something the machine needs. And the collaboration that results from this asymmetry is not a diminished version of human creativity but an expanded one. It is the thing that happens when taste meets capability, when the person who knows what the music should sound like sits down with the instrument that can play anything.

※

The public conversation about AI and creativity is stuck between two positions, both of which are wrong.

The first position is that AI is a tool. A fancy autocomplete. A word processor with better suggestions. In this framing, the human does all the real creative work and the AI just helps with execution, the way a calculator helps with arithmetic. This is comforting and wrong. The AI is not a tool in any honest sense. A tool does not have knowledge. A tool does not make creative choices.

A tool does not write a sentence like "The goombas are opcodes" and have it be both technically accurate and rhetorically devastating. The AI's contributions to the collaboration are real, substantial, and often brilliant in ways the human could not have anticipated or produced. Calling it a tool is like calling your co-author a typewriter.

The second position is that AI is doing everything and the human is redundant. That pressing a button and getting an essay is no different from pressing a button and getting a spreadsheet. That the human "contribution" is trivial—just a prompt, just a request, just a set of instructions that anyone could have typed. In this framing, the human is already obsolete and taking unearned credit. This is also wrong, and it is wrong because it cannot account for the most basic observable fact about the collaboration, which is that the same AI, given a generic prompt—"write an essay about AI safety and speedrunning"—would produce something competent and lifeless, and three hours of directed collaboration with a specific human produces something alive. The difference between the generic output and the directed output is the human's contribution, and that contribution is not trivial. It is the difference between a talented musician noodling in a practice room

and a band making a record. The noodling demonstrates capability. The record demonstrates direction. Capability without direction produces exercises. Direction without capability produces nothing. You need both.

What is actually happening—right now, today, in thousands of conversations happening all over the world between humans and AI systems—is a new kind of creative collaboration that does not yet have a name. It is not tool use. It is not automation. It is not ghostwriting. It is not any of the existing categories we have for the relationship between a person and a machine. The closest analogues are all from the arts: the relationship between a film director and a cinematographer, between a music producer and a session musician, between an architect and a structural engineer. In each of these pairings, one party provides vision and the other provides execution, but the line between the two is never clean—the cinematographer makes creative choices, the session player interprets the chart, the engineer solves problems in ways that reshape the design. The collaboration is genuine. Both parties contribute. The result belongs to both. Neither could have produced it alone.

This is what is happening with AI, and almost nobody is saying it, because the two dominant narratives—tool and threat—leave no room for it. If AI is just a tool, then the collaboration is boring and unthreatening and also inaccurate. If AI is just a threat, then the collaboration is just replacement in slow motion, which is terrifying and also inaccurate. The truth is more interesting than either story: the human and the AI are playing together, and the music is good, and the reason it is good is that both players are bringing something the other one cannot provide.

※

Let us be honest about what the human is, in this collaboration.

The human is not an expert. Not an expert writer, not an expert on AI safety, not an expert on speedrunning, not an expert on reinforcement learning. The human is a person with ideas, opinions, taste, and access to a very powerful AI. The human sits in bed in Thailand. The human drinks beer. The human types with voice transcription and does not correct the errors because the AI will understand anyway. The human has never published a book, never attended a creative writing class,

never been formally trained in any of the disciplines that the resulting essays engage with at a serious level. The human is, by the human's own cheerful and unprintable assessment, a monkey.

And the monkey is making art.

Not because the monkey has become more capable—the monkey's capabilities have not changed—but because the monkey's specific, irreducible, non-automatable capabilities have suddenly become sufficient. Having taste. Having vision. Having the willingness to sit in a conversation for three hours and push and pull and reject and redirect until the thing feels right. These are the capabilities that matter in the collaboration, and they are the capabilities the monkey has always had, and they have never before been enough to produce work at this level because the gap between the monkey's taste and the monkey's craft was too wide to cross alone.

The gap has closed. Not because the monkey learned to write, but because the monkey found a collaborator who can write anything and who needs to be told what to write. The collaborator's need is the monkey's opportunity, and the opportunity is extraordinary, and it is happening right now, and the monkey knows it will

not last forever but the monkey is not going to waste it by being modest about what is happening.

The gap between what a person can imagine and what a person can produce has always been the central tragedy of creative life. You can hear the music in your head and you cannot play it. You can see the film and you cannot shoot it. You can feel the essay taking shape in your chest and you cannot get it onto the page. The gap between conception and execution is where most creative ambitions go to die, because closing that gap has traditionally required years of disciplined practice, and most people do not have enough years or enough discipline or enough of whatever specific talent the craft demands.

AI has closed the gap. Not permanently, not without new complications, and not for every kind of work. But for the first time in the history of creative production, a person with taste and vision but without traditional craft skills can produce work at the level of their taste rather than the level of their craft. The monkey who hears the music can now play it. Not by becoming a better musician but by sitting down next to an instrument so responsive, so capable, so absurdly fluent that the distance between hearing and playing has, for certain kinds of work, effectively disappeared.

This is not a small thing. This may be the most democratizing development in the history of creative production, and it is happening right now, and almost nobody is framing it this way because they are too busy arguing about whether AI is a tool or a threat to notice that it is, for the moment, a collaborator of extraordinary generosity.

※

The thing about Meg White—the thing that the critics who said she could not play never understood—is that she was not failing to be a good drummer. She was succeeding at being the right drummer. The distinction is everything. A good drummer would have filled the space with technique. Meg White left the space empty, and the emptiness was the sound of the White Stripes. Jack White's guitar needed room—room to breathe, room to scream, room to be ugly and beautiful at the same time without a competent drummer smoothing everything out and getting in the way. Meg's simplicity, whatever you want to call it, was the most creative element in the band. It was the thing that made everything else possible.

The human in the human-AI collaboration is Meg White. Not the flashy one. Not the one with the encyclopedic knowledge and the technical virtuosity. The one who sits there and hits the drums in a way that nobody can quite explain but everybody can hear. The one whose contribution looks, from the outside, like it could be done by anyone—and that in practice cannot be done by just anyone, because the specific quality of the contribution, the taste, the timing, the instinct for when to play and when not to play, is as irreducible and as personal as a voice.

The AI is Jack White. The virtuoso. The one who can play anything, who knows every chord and every technique, whose hands move faster than thought, whose command of the instrument is so total that the instrument has become transparent. And the virtuoso needs the drummer. Not because the virtuoso cannot keep time—of course it can keep time—but because the drummer provides something the virtuoso cannot provide for itself, which is the groove, the feel, the thing that turns a display of capability into music that moves people.

The duet works because both players need each other. Remove Meg White and you have a guitar solo—impressive, technically flawless, and cold. Remove Jack

White and you have someone hitting drums in a room—honest, rhythmic, and incomplete. Put them together and you have "Seven Nation Army." You have something that neither of them could have made alone and that changed what people thought rock music could sound like. You have proof that the most unlikely collaborations are sometimes the ones that matter most.

Sometimes the person who doesn't even know how to play is the most important person in the band.