Why Computers Will Never Read (or Write) Literature: A Logical Proof and a Narrative

Angus Fletcher

➡ *For additional information about this article*
https://muse.jhu.edu/article/778252

Angus Fletcher

Why Computers Will Never Read (or Write) Literature: A Logical Proof and a Narrative

ABSTRACT: In response to Franco Moretti's project of distant reading and other recent developments in the Digital Humanities, this article offers a proof that computers will never learn to read or write literature. The proof has three main components: (1) computer artificial intelligences (including machine-learning algorithms) run on the CPU's Arithmetic Logic Unit, which performs all of its computations using symbolic logic; (2) symbolic logic is incapable of causal reasoning; and (3) causal reasoning is required for processing the narrative components of literature, including plot, character, style, and voice. This proof is presented both in logical form and as a narrative. The narrative's *beginning* traces the origin of automated symbolic-logic literature processors back before modern computing to the 1930s Cambridge scholar I. A. Richards. The *middle* recounts how these processors became a basis of New Criticism, Cultural Poetics, and other twentieth- and twenty-first-century theories of literary "interpretation." And the *end* explores how those theories' preference for symbolic logic over causal reasoning leaves them vulnerable to the same blind spot that early modern scientists detected in medieval universities: the inability to explain *why* literature (or anything else) works—and therefore the inability to comprehend *how* to use it.

KEYWORDS: *digital humanities, interpretation, scientific method, causal reasoning, neuroscience, machine learning*

Angus Fletcher is Professor of English and Core Faculty of Project Narrative at The Ohio State University. His most recent books are *Comic Democracies: From Ancient Athens to the American Republic* (2016) and *Wonderworks: The 25 Most Powerful Inventions in the History of Literature* (2021).

1969 WAS A HEADY YEAR for Artificial Intelligence. Palo Alto hosted the first International Joint Conference on AI. Marvin Minsky, co-founder of MIT's AI Project, earned a Turing Award, validating his prediction that machines would soon become as intelligent as people. Stanford's John McCarthy and Patrick J. Hayes drafted a philosophical code that provided computers with the experience of "freewill" (464). And Hugo Award–winner Brian Aldiss inked a sci-fi fable, "Supertoys Last All Summer Long," about an AI that learned not only to love, but to love *better* than humans, marshalling its silicon circuitry to care with unmatched empathy.

So blue-sky were these doings that they'd be mocked for much of the next decade as deludedly quixotic. But AI researchers shrugged off the skepticism, forging on to make breakthroughs in machine-learning, data mining, case-based reasoning, visual perception, and natural-language processing. And such was their progress that when "Supertoys" was adapted by Steven Spielberg into a Hollywood film at the turn of the millennium, even Aldiss's audacious tomorrow no longer seemed audacious enough. Deciding that a tale about a heartful robot had come to feel underwhelming, Spielberg posed the question: *What's even more uniquely human than emotion?* And in the 2001 blockbuster *Artificial Intelligence,* he unveiled his answer: literature. Literature, the movie's robots enviously conclude, can generate "meaning," that most distinctive of human achievements . . . distinctive, that is, until a robot learns to digitize it, uploading a fairytale that propels him on a journey into hope, happiness, and ultimately soulful peace.

Artificial Intelligence speculates that its literature-processing robot lies in our not-too-distant future, a couple centuries ahead. And as daring a prognostication as this was intended to be, recent developments in university literature departments have suggested that Spielberg may have fallen into Aldiss's error of not being bold enough. Those developments have occurred within the broad and bustling field known as the Digital Humanities,[1] but their primary point of origin is Franco Moretti's early 2000s discovery that machine-learning algorithms can interpret literature in staggering quantities—and at staggering speeds.[2] In seconds, a computer can crunch through all the novels of the nineteenth century, churning out readings that would take mortal scholars lifetimes to achieve. And as machine-learning has gone on to revolutionize other fields of human industry, from logistics, to gaming, to healthcare, the question has naturally raised itself: How long until computers revolutionize the study of literature? When will an algorithm learn to understand novels in all the ways that humans do—only faster and more precisely? When will digital devices outdo literary experts at analyzing poems and screenplays, just like they've outdone the world's greatest grandmasters at chess? And when will computers complete their education by graduating from consumers into producers, mobilizing their knowledge of literature to create new works? When will their whirring electronics, which can already complete our email sentences as we type, become the next Sappho or Shakespeare?

In this article, I'll provide a definite answer. And that answer is: never. No computer, no matter how immense its circuitry, will ever be able to extract the know-how from a fairytale that can be gleaned by human children. No machine-learning algorithm, no matter how futuristic its software, will ever author a sonnet or short story. The reason for this is simply that literature encodes a great deal of its thought-stuff

in narrative, a mode of communication that requires causal reasoning to process. And while the ability to do causal reasoning is embedded in the architecture of the human brain, computers are hardwired to perform a method of thinking—symbolic logic—that is fundamentally incapable of grasping cause-and-effect.

A summary proof of this fact, written in a logical form that a computer can digest, is laid out on the article's final page. But to make the proof more congenial to our own gray matter, the following pages will relate it in the form of a story. That story will trace the dream of machine-reading literature back before the Digital Humanities and *Artificial Intelligence* to its beginning in the early twentieth century. And it will then chronicle the limits of machine-reading by describing how the dream led generations of scholars, right down to the present, to downplay and even deny a profoundly special feature of literature—and of our humanity too.

The Dream of Machine-Reading Literature

The dream sprang from the brain of young University of Cambridge professor I. A. Richards.

Richards began his teaching career in 1919, the very first year that Cambridge offered an independent English Literature degree.[3] So newfangled was the degree that Richards had no formal faculty position, forcing him to collect tuition from students as they filed into his classroom (Russo 87). This precarious situation exhausted Richards but never discouraged him, and he devoted much of his early thirties to writing a sequence of trade books—*The Meaning of Meaning* (1923; co-authored with C. K. Odgen), *The Principles of Literary Criticism* (1924), and *Practical Criticism* (1929)—that laid out a blueprint for improving literary studies.

Richards's blueprint was filled with ingenious ideas, but their overarching purpose was to establish the viability of "interpretation" (*Practical Criticism* 337). Interpretation, as detailed carefully by Richards, is a practice derived from semiotics, which is itself a derivative of "symbolic logic," a sophisticated system of inductive observation and deductive analysis that treats printed texts, human cultures, and even the cosmos itself as languages of "signs" and "representations" (Ogden and Richards, *The Meaning of Meaning* 269–90). Out of these languages, interpretation can derive hidden significations or "meanings," and since, as Richards pointed out, literature is brimful of languages—written, oral, imagistic, metaphorical—it too teems with buried imports that can be brought to light via semiotic interpretation.

These buried imports are the same "explanations of . . . meaning" that Spielberg's *Artificial Intelligence* would celebrate as the ultimate end of our humanity.[4] And long before Spielberg, Richards foresaw that their mode of discovery could in time be mastered by artificial intelligences. He realized that if meaning came from interpretation and interpretation from symbolic logic, then any symbolic-logic processor could be programmed to extract the deep significances of poems, novels, and stage plays. This led Richards, in 1923, to predict the invention of "automatic thinking-machines" that would one day "interpret" literature (Ogden and Richards, *The Meaning of Meaning* 223).

Like the AI prognostications of 1969, Richards's vision of machine-reading struck most contemporary academics as bizarre—yet swiftly proved prophetic. During the middle years of the twentieth century, it helped inspire a pioneering school of literary studies, New Criticism, that promoted semiotic interpretation to great effect in British and American universities, enrolling millions of undergraduates, training thousands of college and high school instructors, and reshaping the administrative landscape of higher education (Litz, Menand, and Rainey 1–14). In 1979, Yale appointed a New Critic, Angelo Bartlett Giamatti, as its president;[5] in 1991, Harvard followed suit with the appointment of the New Critic Neil L. Rudenstine.[6]

The New Critics were able to achieve this institutional ascent because, as Richards had foreseen, semiotic interpretation imbued their labors with two of the performance benchmarks—pace and scale—characteristic of automated computing. Semiotic interpretation could be taught efficiently to large quantities of students and generate high volumes of original scholarship, rapidly increasing the number of undergraduate literature majors and creating thousands of new faculty positions across the globe.[7] And indeed, so successful was interpretation that its techniques spread beyond New Criticism into other schools of literary criticism—from neo-Marxism to Deconstruction to Cultural Poetics—and then beyond literature departments into Theater, Film Studies, Classics, Folklore, Anthropology, History, and Dance.[8] But as influential as Richards's emphasis on semiotic interpretation proved, it came with heavy trade-offs: it radically narrowed the scope of literary inquiry and disregarded most of the practical benefits that the general public had previously associated with reading literature. So to discern both what the dream of machine-reading created and what it cost, let's widen our narrative lens, turning back a page in history to uncover the approach to literature that Richards displaced.

Reading Literature Before Richards

When I. A. Richards embarked on his professorial career in 1919, the dominant approach to reading literature was "character criticism" (Bate 2). Character criticism had emerged roughly a century earlier as an effort to defend Shakespeare from "Neoclassical" critics.[9] Those critics took it as axiomatic that literary characters should behave consistently, so when they arrived at the theater and bore witness to Hamlet's violent reversals of course, they didn't hesitate to accuse Shakespeare of authorial incompetence.

This accusation was rebuffed in a series of lectures delivered between 1808 and 1819 by the English Romantic Samuel Taylor Coleridge. As Coleridge saw it, the clash between the erratic behavior of Shakespeare's Danish prince and the rigid expectations of Neoclassicism wasn't proof that *Hamlet* was a failure. Quite the opposite—it was proof that Neoclassicism had botched the job of literary scholarship, necessitating a radically new method: "I believe the character of Hamlet may be traced to Shakespeare's deep and accurate science in mental philosophy . . . In order to understand him, it is essential that we should reflect on the constitution of our own minds" (Coleridge, quoted in Roberts 142).[10]

With this, Coleridge brushed off Neoclassicism for being superficial. Neoclassicism's rules may have been born of logic, but logic skipped across Hamlet's surface, unable to hook into his "deep" psychology. To catch that deep, it was therefore "essential" to develop a more "accurate science," the beginnings of which lay in an examination of "our own minds."

Coleridge's new science proved enormously popular. It banished the "Dramatic Unities" and the other mathematically indisputable but emotionally wooden doctrines of the Neoclassicals. And it replaced them with an abundant populism. Anyone could practice character criticism; it was as simple as relating to Hamlet. And because the intricacies of human minds were so vast as to border on the infinite, there were endless breakthroughs to be made. So, in pursuit of fresh wisdom, character critics kept mind-spelunking further, casting introspective lanterns into the remotest niches of their individual psyches, until eventually, at the dawn of the twentieth century, their sprawling democratic method succeeded in working its way from England's country living rooms and grubby public halls into the rarefied ivy of the nation's most ancient university: Oxford.

Oxford had in 1901 appointed the boyishly unpretentious graybeard A. C. Bradley as Professor of Poetry, establishing him as the British Empire's highest literary expert. And three years later, in 1904, Bradley used his position to publish a monumental tome of character criticism: *Shakespearean Tragedy. Shakespearean Tragedy* was not a work of great originality; it continued the by now long familiar psychological "science" described in Coleridge's lectures. Yet still, *Shakespearean Tragedy* was lent enormous importance by Bradley's academic position. And befitting Bradley's scholarly eminence, he pursued his inquiries with careful precision. He finely weighed Coleridge's famous theory that Hamlet had delayed out of "that aversion to action which prevails among such as have a world in themselves" (Roberts xxvi). And he then rejected the theory as oversimplistic, declaring that

> it misconceives the cause of that irresolution which, on the whole, it truly describes. For the cause was not directly or mainly an habitual excess of reflectiveness. The direct cause was a state of mind quite abnormal and induced by special circumstances—a state of profound melancholy. Now, Hamlet's reflectiveness doubtless played a certain part in the *production* of that melancholy, and was thus one indirect contributory cause of his irresolution. And, again the melancholy, once established, displayed, as one of its *symptoms,* an excessive reflection on the required deed. But excess of reflection was not, as the theory makes it, the *direct* cause of the irresolution at all; nor was it the *only* indirect cause; and in the Hamlet of the last four Acts it is to be considered rather a symptom of his state than a cause of it. (Bradley 107–8; emphasis original)

If the prose here feels stilted, it's because Bradley is determined to be exacting. Trading stylistic grace for academic fastidiousness, he repeatedly deploys the technical scientific term "cause"—and indeed does so with such meticulousness that he distinguishes two distinct species: "direct" and "contributory."

This combination of Coleridge's popular science with Bradley's professorial diligence made *Shakespearean Tragedy* an enormous success. In the twentieth century's early decades, it became a "landmark" of literary scholarship and earned the devoted approval of the broader public,[11] ensconcing itself as a staple of book clubs and classrooms: "Undergraduates in the first year have told me that . . . it gives them human beings behaving in a way which they have always thought of human beings behaving" (Joseph 91).

Yet the undergraduates would not be permitted to think this way for long. For *Shakespearean Tragedy*—and all of character criticism—was about to be indicted by I. A. Richards for committing a gross logical error.

In Which Richards Overturns Character Criticism with an Automated Logic

When *Shakespearean Tragedy* was published, eleven-year-old Richards was busily building model steam-engines (Russo 5). He paid no attention to Bradley's book then, and his indifference continued for years. Although he was an eager childhood reader of literature—Rudyard Kipling, Robert Louis Stevenson, Jules Verne—he focused his studies on philosophy. And in June of 1915, he took his degree at Cambridge by earning a first class in a Tripos exam heavy on logic and linguistic analysis.

This academic training laid the groundwork for Richards's case against character criticism, a case that began in earnest with the 1923 publication of *The Meaning of Meaning. The Meaning of Meaning* made no direct mention of *Shakespearean Tragedy,* yet it struck at the book's heart by denying the existence of that scientific element of Hamlet's psychology—"cause"—so minutely parsed by Bradley: "A Cause indeed, in the sense of something which forces another something called an effect to occur is so obvious a phantom that it has been rejected even by metaphysicians" (Ogden and Richards 142). This was a fantastically counterintuitive claim. But it was also rigorously logical; so logical that it had been deductively proven a decade earlier by the era's most renowned British logician: Bertrand Russell.

Russell had elbowed his way to fame in 1903 with *The Principles of Mathematics,* which opened by combatively—but not inaccurately—asserting that "all Mathematics is Symbolic Logic" (*Principles* 5). Over the following decade, this equivalence between mathematics and symbolic logic was elaborated by Russell and his collaborator, Alfred Lord Whitehead, into the three volume *Principia Mathematica,* which proved the laws of math via symbolic equations:

$$\supset \vdash : . \; \alpha = \tau \, `x . \; \beta = \tau \, `y . \; \supset : \alpha \cup \beta \in 2 . \equiv . \; x \neq y$$

Such was the painstaking rigor of these logical exercises that it took fifty-four chapters of dense calculations for the authors to reach the stage where they could finally announce: "From this proposition it will follow . . . that $1 + 1 = 2$" (Russell and Whitehead 1.379). Yet even so, Russell was certain that the *Principia*'s logic had far-reaching consequences—consequences that would not only transform epistemol-

ogy, politics, and spirituality, but exorcise the old metaphysical bogeyman of cause-and-effect. As Russell declared in 1912: "The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm" ("On the Notion of Cause" 1).

Russell wasn't the first to question the "law of causality." It had previously been interrogated by philosophers from Henri Bergson, to David Hume, all the way back to Epicurus. But Russell added his own fresh insight. He began by observing that the only evidence for causality—that is, the only reason to suppose that causes and effects existed at all—came from philosophical intuitions. Those intuitions had taken different forms, from J. S. Mill's common sense to Immanuel Kant's transcendental cognitions, but they were all reflections of what *seemed* logical to human brains.

This method of reasoning was categorically rejected by Russell. He observed that many things that our brains "instinctively imagine" run counter to discovered truths ("On the Notion of Cause" 20). And among those discovered truths, he claimed, was that mathematics could model the workings of the universe without ever invoking cause or effect: "In the motions of mutually gravitating bodies, there is nothing that can be called a cause, and nothing that can be called an effect; there is merely a formula" (14).

Russell's point here is that planetary movements can be modeled by mathematical equations. And mathematical equations don't take the form of "X *causes* Y." They take the form of "X *equals* Y," or in other words, "X *is* Y," capturing the movements of the cosmos in a computational sequence of . . . *is* . . . *is* . . . *is* . . . that admits no "was"—and so no prior cause. Like the images of a motion picture, each of these "is" instants seems to our viewing mind to be part of a continuous chain of events, but there's no more causal connection between them than there is between the snapshot print-frames of a cinematic reel.

When this logico-mathematical denial of causation was accepted by I. A. Richards in *The Meaning of Meaning,* it completely undid Coleridge's science of *Hamlet.* That science had been premised upon the intuition that the human mind—with all its motives, desires, and other psychological springs—was the cause of human behavior. But if logic was right, then cause was a mental fallacy—or to use Richards's term, a "phantom." To chase it was to join Hamlet in madness.

In a bid to restore sanity, Richards therefore abandoned the method of the Romantics and A. C. Bradley, returning to the old Neoclassical view that literature was governed by mathematical laws. And to isolate those laws more successfully than the Neoclassicists had done, Richards sought guidance from a logician who'd developed much of the theory upon which Bertrand Russell's symbolic equations were based.[12] The logician was C. S. Peirce. He'd come of age during the American Civil War, and having witnessed the terrible results of human disagreement, he sought out a way to universal truth: "Now thought is of nature of a sign. In that case, then, if we find out the right method of thinking and can follow it out—the right method of transforming signs—then truth can be nothing more nor less than the last result to which the following out of this method would ultimately carry us" (Peirce, *Writings* 553).

Peirce's verdict here is breathtaking. It begins with the observation that when we think of gold, gold doesn't literally materialize inside our brain. Instead, what appears is a mental representation of gold, or in other words, a "sign." From here, Peirce proceeds to the conclusion that human thought is by "nature" a system of signs—that is to say, a symbolic language.[13] And from that premise, Peirce deduces a solution to every dispute that has ever bedeviled philosophy, concluding that if human thought is composed of signs, then by finding the "right method"—that is, the right "semiotics"—for logically interpreting signs, we can determine the "truth" of everything inside our heads: "mathematics, ethics, metaphysics, gravitation, thermodynamics, optics, chemistry, comparative anatomy, astronomy, psychology, phonetics, economics, the history of science, whist, men and women, wine" (Peirce, *Semiotics* 85–86).

This list, drawn up by Peirce, was cited approvingly in 1923 by Richards (*The Meaning of Meaning* 125), who proceeded to add one more item: literature. And as he did, he experienced his machine epiphany. The epiphany began with Richards's observation that logic's mathematical formulae were a "symbolic machinery" that spun with irresistible force, necessitating specific conclusions: 2 + 2 couldn't ever equal 3 or 5; it had to equal 4 (312). Having observed this truth, Richards then followed it to the inexorable conclusion that all the academic fields governed by logic were themselves "automatic thinking-machines" (223). And since one such field was literary semiotics, Richards boldly claimed that it was possible to root a "Theory of Interpretation" in "a logical machine of great sensitiveness and power" that "may in time be made self-running and even fool-proof" (*Practical Criticism* 217).

An extraordinary tomorrow was conjured by those two adjectives: "self-running" and "fool-proof." The first summoned an image of a literature processor that ran on its own, churning autonomously through reams of poems and novels and plays to calculate meaning. The second celebrated the processor for being free from causal reasoning and the other logic errors of human intuition, making it an infallible literature robot capable of interpreting Shakespeare to reach the ultimate truth.

For most of the next century, this robot was too futuristic a thing to be realized. Instead, the task of interpretation fell to human scholars. One of those scholars, L. C. Knights, helped launch New Criticism in the 1930s with his essay *How Many Children Had Lady Macbeth?,* which deployed semiotics to claim Shakespeare's characters as allegorical symbols who transcended narrative time to establish an eternal meaning: "Macbeth is a statement of evil" (34). Another of those scholars, Stephen Greenblatt, helped launch Cultural Poetics in the 1980s by co-founding the UC Berkeley journal *Representations,* which treated history and culture as linguistic systems of "representations" that could be interpreted to yield insights into freedom, justice, and other semiotic truths.

Then, gradually, this work of human interpretation began to converge with advances in computing. In the 1940s, the world's first programmable electronic computer—ENIAC—was constructed, bringing into being the machine intelligence that Richards had predicted two decades before. And as computers grew steadily faster, Richards's literary robot edged into being. In the 1950s, Berkeley literature professor Josephine Miles collaborated with her university's Electrical Engineering department to map John Dryden's *Poetical Works* with IBM tabulation machines.[14] And over the

following decades, more scholars would use automata to analyze literature, culminating at the turn of the present millennium in the invention of "distant reading" and the other techniques of the Digital Humanities.[15]

When the Digital Humanities made its initial foray into literature departments in the early 2000s, it was regarded warily by critics who worried that its robotic algorithms would be engines of rote absolutism that threatened the core benefits of literary training: critical thinking, indeterminacy, diversity, self-reflection, care-for-others, subtlety, and perhaps most importantly, autonomy. Yet over the past ten years, Digital Humanists have eased those concerns. They've proven that computing doesn't need to operate in crude FALSE/TRUE binaries; it can employ "Bayesian inferences" to muse in countless shades of gray (Underwood). They've demonstrated that machine-learning doesn't converge on one final reading of literature; it proliferates interpretations, nurturing freedom, open-endedness, and choice (So). And they've shown that automated reading doesn't lead inevitably to hierarchies of power; it can promote equality, inclusiveness, and accessibility (Gold and Klein; and Binder and Jennings).

By aligning itself in these ways with the priorities of modern literary scholarship, the Digital Humanities has seem poised to make good on Richards's prediction that a logical robot will one day discern literature's ultimate meanings.[16] Every computer on earth contains the universal logic gates—NAND and NOR—that, as C. S. Peirce proved in the late nineteenth century, can perform all the tasks of semiotics. Which means that millions of college students now have access to the hardware required to compute the truth of *Hamlet* or any other literary text (Piper), enabling them to surpass New Criticism's furthest reaches with a handful of keystrokes.

But even if those keystrokes are being input, right now, into a dorm room laptop somewhere, they cannot portend the revolution in literary studies that Richards imagined. For as AI researchers have recently discovered, the same feature of computer thought that Richards hailed as its enormous strength—its substitution of symbolic logic for causal reasoning—also comes with limitations. Limitations that carry us backward, before our scientific age, to the metaphysics of the Middle Ages.

The Medieval Limits of Machine Reading

During the European Middle Ages, the foundation of learning was symbolic logic.

Symbolic logic was derived from Aristotle's *Organon,* which laid out the basics of semiotics (signs, representation, interpretation) and propositional argument (induction, deduction, syllogisms) in six dense books. Most of these books had been lost in the West after the fall of the Western Roman Empire in 476. But they survived in the Greek and Arabic civilizations of the East, and at last, in the twelfth century, the books were translated into Latin, sparking a glorious new age of learning at schools from Padua to Paris. In the 1230s, the friar Albertus Magnus began employing semiotic interpretation to deduce truths of stars and souls; in the early 1270s, his student Thomas Aquinas syllogistically produced five proofs of God's existence.

Yet as magnificent as all this learning was, it had a limit. Because logical propositions are—as Bertrand Russell observed—interchangeable with mathematical equa-

tions, and because mathematical equations inhabit a perpetual "now" where there exists no cause and no effect, logic cannot ever prove causation. It can only prove *correlation* (Pearl 401–426). It can demonstrate that "X is Y"—which is to say, that X and Y intrinsically co-occur—but can shed no light into the origin of the co-occurrence.

Logic can, in short, assert *that* two things are linked—but it cannot explain *why*.

Initially this assertion of *that* without *why* worked to the benefit of medieval learning, which was conducted at convents, theological colleges, and other churchly institutions that saw it as part of their educational purpose to evoke wonder at the miracle of Creation. That wonder steered us toward salvation by inspiring our mortal mind to worship God's almighty wisdom. And conveniently for our priestly teachers, that wonder also enhanced their professional status, investing them with a shimmering halo of divine *auctoritas,* or "authority" (Ziolkowski 421).

This culture of kneeling awe was fed by symbolic logic's assertion of *that* without *why*. The assertion flung into being a worldwide web of correspondences that—amazingly—always held true and—even more amazingly—had an underlying rationale that eluded mortal understanding. Like a Genesis *ex nihilo,* syllogistic truth popped into existence for no comprehensible reason beyond: *God saw that it was good.*

But even as logic fed spiritual experience, there was a biological hunger that it couldn't satisfy: the human brain's innate desire to know *why*.[17] That desire was the reason for the world's first religions:

> *Why is there thunder?*
> *There must be a god in the sky.*

And although medieval logic promised to replace such fanciful answers with the eternal stuff of truth, what it did instead was frustrate humanity's primal curiosity with endless versions of the following conversation:

> *Why did God see that it was good?*
> *Because God is good.*
> *But why is God good?*
> *Because good is the essence of God.*

On it went, through infinite tautologies, as everything was justified back to God, whose reasons were in turn justified on the grounds that God was the definition of reason.

The result was a body of learning that claimed to know everything and yet was unremittingly superficial. It squashed the four dimensions of life into one-dimensional representations that could be inked on parchment, reducing all study to book learning and all research to interpretation. It spawned tomes of exotic terminology that exhaustively labeled the "forms" of nature without penetrating their inner nuts and bolts. It took the rich complexities of human psychology and history, and flattened them into taxonomic charts and moral allegories.

Five hundred years on, this medieval quality can be felt in the machine-learning of modern computers. The computers are driven by precisely the same symbolic logic as medieval metaphysics: the Aristotelian syllogisms that powered every academic

disputation of the later Middle Ages were used by George Boole in 1854 as the basis of Boolean Algebra,[18] which would itself be fashioned in 1937 by Claude Shannon into the core computing hardware—now known as the CPU's Arithmetic Logic Unit— that runs all twenty-first-century machine-learning algorithms.[19] And so it is that those algorithms can discern the most unexpected patterns—in consumer buying habits, in traffic jams, in strategy games, in corporate inventories, in spam, in stock markets, in human faces—without ever explaining *why* the patterns exist.

The same shallow wonder permeates the interpretations of I. A. Richards's robot. The robot in each of its various guises—from New Criticism to Cultural Poetics to the Digital Humanities—deploys semiotic interpretation and other techniques of symbolic logic to reveal startling patterns in literature that have gone unnoticed by previous generations of readers. Yet the robot's logic cannot guide us into fathoming the reasons that those patterns exist (So 670). So, the robot's users have two choices: they can content themselves with mass producing "close readings" and "data visualizations" and other word clouds of meaning that initially dazzle the mind but soon come to feel arbitrary, insubstantial, and aimless; or, they can shore up the hole at the center of logic by resorting to the old medieval strategy of providing *why* via tautology.

In the case of New Criticism, the tautology often works by claiming great authors as miniature versions of God, such that when inquisitive students wonder why Shakespearean sonnets are filled with particular patterns of imagery, the teacher responds: *Because Shakespeare is the definition of great poetry.*[20] In the case of Cultural Poetics, the tautology often works by invoking Foucauldian epistemes, Marxist social relations, or other larger cultural forces that have been deduced—via quasi-syllogistic philosophical reasonings—to be determinants of literature.[21] In the case of the Digital Humanities, the tautology often works by "p-hacking" or "data-dredging" (that is, cherry-picking) from bazillions of word-associations, such that we ignore the associations that seem random while highlighting the ones that confirm our preexisting theoretical assumptions.[22]

And the tautology isn't restricted to these outsized examples. It's a feature of *every single interpretation of literature.* Interpretation, as the medieval rebooters of the syllogism discovered, can infallibly derive new conclusions from given truths, but it cannot generate those given truths or establish their veracity. So, every act of interpretation must begin with the interpreter positing *a priori* first principles, or ontological definitions, or other unprovable grounding assumptions. In medieval metaphysics, these assumptions are provided by the Bible; in modern literature departments, they're supplied by a rich and growing catalogue of theoretical canons. But in both instances, they reflect what the interpreter believes to be true, not what the interpreter can demonstrate irrefutably. And thus it is that the interpreter's conclusions are always partly circular: when we accept them as true, we end up confirming what the interpreter started by supposing.[23]

This feature of interpretation can seem to doom the project of extracting knowledge from poems and novels and plays: if literary interpreters—both human and computer—can only reach their truth conclusions by positing things that may or may not be accurate, then how true can their truth conclusions be trusted to be?

But the inherent truth-circularity of interpretation doesn't mean that modern literary research is fated to end up becoming, like Aquinas's God-proof, the indisputable fact that falls into endless dispute. For as the long and fertile history of literary studies reveals, interpretation isn't the only possible path to learning from literature. There's an alternative route, the one roughly employed by Coleridge and the character critics: science.[24]

Science shuffled quietly into being via alchemical experiments and astrological predictions that date back through the Islamic Golden Age to the trial-and-error discoveries of Vedic India and Chalcolithic Mesopotamia. And even after the scientific method was loudly positioned by Galileo Galilei, Francis Bacon, and other early-seventeenth-century natural philosophers as a direct challenger to the intellectual traditions of the Christian Middle Ages, it took many more decades for science to work its way into the university lecture halls of Paris and Oxford. That's because science is slower to yield its gleanings than symbolic logic. And in fact, science cannot *ever* claim to deliver what logic, from its beginning, promises: the truth. There is no truth in science, only probability and the hypothesis "unfalsified."[25] So, compared to logic, science provides less authority and less immediate wonder. It's more open to question, more frank about itself as a work in progress, and more exciting for what it *could* in the future be.

Yet in spite of this perpetual incompleteness, science did in the end manage to establish its place alongside logic on college campuses. And science did so because it takes a different approach to *why,* one that trades truth for something that our brains value even more: empowerment.

The Empowerment of Science's Why

The empowerment was observed by Francis Bacon.

Bacon had become legendary by the time of A. C. Bradley for discerning science's physical potency. It was Bacon, the Victorians rhapsodized, who kindled "in science a revolution which will, to the end of time, be reckoned among the highest achievements of the human intellect . . . [yielding] practical purposes on a scale never before known" (Macaulay 399, 276). This revolution in the "practical" remains Bacon's legacy today.[26] Bacon—as modern historians have discovered—misunderstood the method by which science actually works,[27] but he unerringly anticipated the value that later generations would place upon science's utility. In an era when science was mostly known for its dismaying disenchantment of the heavens, Bacon foresaw that it would someday earn the public's gratitude for its enrichments of the sublunar arts of medicine, agriculture, and technology.

Bacon's act of foreseeing was laid out in 1620 with the publication of his *Novum Organum,* or "*New Organon.*" The work was, as its title aggressively declared, a bid to replace the method of semiotic interpretation laid out in Aristotle's "old" *Organon.* And to engineer that replacement, Bacon began his new *Organon* by forging a link between science and human power that he anchored in science's ability to probe into *why*: "Human power is coincident with science, because you can't generate effects if

you're ignorant of causes" (*Novum Organum* 1.3; *Scientia et potentia humana in idem coincidunt, quia ignoratio causae destituit effectum*). Human power, in other words, is the ability to "generate effects," an ability that comes from the insight into "causes" that "science" provides.

Science provides this empowering insight because its method (as researchers post-Bacon have learned) swaps out interpretation for two other research practices—prediction and experiment—both of which involve causal reasoning.[28] In prediction, the scientist works backward from an observed effect to a hypothesized cause: *I predict that the cause of diabetic symptoms is low blood-insulin.* Or: *I predict that the cause of smallpox immunity is prior exposure to cowpox.* In experiment, the scientist then tests her prediction by physically manipulating the cause. She injects insulin to see whether the injection *causes* the symptoms of diabetes to lessen. Or she performs a cowpox exposure to see whether the exposure *causes* immunity. If the manipulation works as anticipated, the scientist has discovered a new power: the power to treat diabetes, or vaccinate against smallpox, or generate some other desired effect.

This derivation of know-how from an investigation of causes is what distinguishes science from medieval learning. Medieval learning is rich with discussions of causes: both Albertus Magnus and Thomas Aquinas carefully anatomize causes into four separate kinds—material, formal, efficient, and final—that are derived from Aristotle's *Metaphysics.* But as Bacon brusquely notes, Aristotle's *Metaphysics* fails to unlock the practical power of *why,* because "Aristotle enslaves the study of nature to his logic, making it useless" (*Novum Organum* 1.54; *Aristotele, qui naturalem suam philosophiam logicae suae prorsus mancipavit, ut eam fere inutilem et contentiosam reddiderit*). Aristotle, that is, transports "cause" into a semiotic system where (as we saw from Bertrand Russell) the "was" and "will be" that are necessary for cause-and-effect don't exist. So, like "green" in a colorless cosmos, "cause" is reduced in Aristotle's *Metaphysics* to empty verbiage. As Bacon illustrates:

> Take the motion of missiles, like darts and balls, through the air. The Aristotelians have (as is their way) treated this motion with gross negligence, thinking it adequate to distinguish this motion with the name "violent" (as opposed to "natural") motion and to explain the origin of the motion by deducing, "two bodies cannot be in one place." But this pays no mind to the continuing progress of the motion. (2.36)

> *Sit natura inquisita Motus Missilium, veluti spiculorum, sagittarum, globulorum, per aerem. Hunc motum Schola (more suo) valde negligenter expedit; satis habens, si eum nomine motus violenti a naturali (quem vocant) distinguat; et quod ad primam percussionem sive impulsionem attinet, per illud, (quod duo corpora non possint esse in uno loco, ne fiat penetratio dimensionum,) sibi satisfaciat; et de processu continuato istius motus nihil curet.*

In other words, because symbolic logic inhabits a timeless *present,* it cannot capture the *future* effect—the "continuing progress"—of throwing a ball. So, symbolic logic gives us a cause with no effect, which is to say, a cause that isn't actually a cause. It

is instead only a word ("violent motion") and a confirmation of a prior theory (the principle that "two bodies" cannot occupy the same space). Or to put it bluntly: the semiotic interpretation of nature promotes a proliferation of jargon and a reconfirmation of what we already believed, neither of which improves our ball-throwing prowess.

And symbolic logic's shortcomings don't end here. Its failure to provide actual empowerment has the unfortunate knock-on effect, as Bacon observes, of diverting human learning into two unconstructive mental tendencies. The first is a form of magical thinking where logic's treatment of causes as symbols prompts us to confound the symbol with the power it represents. Bacon refers to this as the idolatrous error of medieval Catholicism, which frequently implied through its elaborate rituals and ornamental arts that the sheer representation of a holy good was enough to bring the good to be. The second tendency is a form of ecstatic ignorance, characteristic of Protestant reform movements such as Lutheranism, that preached that human reason had been forever corrupted by the Fall, making it impossible for even the most pious among us to delve into life's rationale. All we could do was submit gladly to God's unfathomable will, accepting that we lacked the power to effect our own upliftment.

Since both tendencies are outgrowths of symbolic logic, they're not limited to medieval learning. They're scattered through logic's various offspring, including modern computer culture. The first tendency can be seen in the magical thinking that leads us to associate Facebook posts and Instagram pics with winning at life; the second, in our increasing capitulation to algorithms that run economic, logistic, and social-engineering decisions that no human understands.[29]

And both tendencies can also be detected in modern literary studies. There is the magical thinking, characteristic of many heirs of Cultural Poetics, that semiotic interpretation makes possible a form of "activism" through which we can counteract imbalances of power by critiquing literary "representations" or media "discourses" (Gold and Klein xi). And there is also the glorification of incomprehension, sometimes in the form of what I. A. Richards's student William Empson called "ambiguity," sometimes in the form of what the New Critic Cleanth Brooks called "paradox," sometimes in the form of esoteric critical theories that promise to initiate us into secret truths.

Together, these tendencies reinforce a popular prejudice—traceable back through the sixteenth-century English Puritans to Plato—that literature itself is a kind of occult activity, magic and arcane. Like the sorcerous illusions of Circe in *The Odyssey,* it unhitches us from reality to sweep us into fictions that are at best escapist and at worst inducements to the lunacy of Don Quixote. And like a Rosicrucian spellbook, it's writ in locutions strange and intricate that hint of Prospero's genius in *The Tempest,* yet lead only to the labyrinths wandered vacantly by Chaucer's alchemist.

But while it's true that literature can provide us with mystical experiences—experiences that can be enlarged by semiotic interpretation—this isn't all that literature can do. In fact, literature can do just the opposite: like science, it can throw open a future of practical empowerment by providing us with clear and accessible tools for transforming ourselves and our physical world.

To access those tools, we simply need to set aside symbolic logic and return to the method that Bacon celebrated and I. A. Richards disowned: causal reasoning.

Causal Reasoning and How It Can Lead to A Practical Science of Literature

I. A. Richards was right to accept Bertrand Russell's assertion that causal reasoning is nonlogical; because logic speaks a language of perennial *is,* while cause-and-effect acts through *was* and *will be,* there is no syllogism or mathematics that can indisputably link an effect to a cause. But although causal reasoning cannot be reduced to formal logic—and has remained resistant to *any* philosophical attempt to systematize it— this doesn't make causal reasoning intellectually worthless.[30] Its method of prediction and experiment has birthed the scientific theories and technological inventions of our STEM age, and as recent work in neuroscience has uncovered, causal reasoning is also the engine of much of the learning that we do outside science and engineering labs, including a good deal of the learning that makes literature possible.

The first hints of this fact were uncovered in the 1960s and 1970s by Nobel laureate Eric Kandel (Roush 1102–3). Kandel and his research team started with an apparent paradox: the neurons of adult brains were largely wired in place, yet somehow those neurons were able to gain fresh skills, acquire new information, and engage in other acts of learning. How was this possible? How did the brain rewire its thoughts when its own wiring was so fixed?

The answer, Kandel discovered, is the "synapse." The synapse is a general term for an enormous variety of complex neural structures. But what all the structures have in common is that they sit at the juncture between the "ending" of one neuron and the "beginning" (or "dendrite") of another. And when the synapse is triggered (typically by an electrical signal from the neuron's axon), it carries a signal across the juncture, becoming the "middle" that connects the two neurons together.

This synaptic middle, as Kandel's research revealed, is adjustable: it can be dialed up to strengthen the link between one neuron and the next, or it can be dialed down to mute the link. And because each of our neurons has multiple synapses, their adjustment allows our brain to continually modify its circuitry. As a simple example, imagine a neuron, A, that has two synapses, *b* and *c,* that form junctures with neurons B and C respectively. When *b* is dialed up and *c* is dialed down, the result is the circuit A → B. But when *b* is dialed down and *c* is dialed up, the result is the circuit A → C. By toggling between these two synaptic arrangements, the brain can therefore rewire itself, even while all its neurons remain in the same place. It can switch from A → B to A → C, enabling us to change our minds and learn.[31]

This process of human learning can be replicated to some extent by the machine-learning of computers. Computers can use their transistors to open and shut different circuit paths, reshaping their network connections. And they can also use Bayesian inferences and other mathematical techniques to emulate our synapses' variable adjustments of connection strength, transcending binary absolutes to cog-

itate in waves of analogue. But there's one feature of human learning that computers are incapable of copying: the power of our synapses to control the *direction* of our ideas. That control is made possible by the fact that our neurons fire in only one direction, from dendrite to synapse. So when our *c* synapse creates a connection between neuron A and neuron C, the connection is a one-way A → C route that establishes neuron A as a (past) cause and neuron C as a (future) effect. It's our brain thinking: "A causes C."

This physiological mechanism is the source of our human powers of causal reasoning. And it cannot be mimicked by the computer Arithmetic Logic Unit. That unit (as we saw above) is composed of syllogistic logic gates that run mathematical equations of the form of "A = C." And unlike the A → C connections of our synapses, the A = C connections of the Arithmetic Logic Unit are not one-way routes. They can be reversed without altering their meaning: "A = C" means exactly the same as "C = A," just as "2 + 2 = 4" means exactly the same as "4 = 2 + 2," or "Bob is that man over there" means exactly the same as "That man over there is Bob."

Such reversibility is incompatible with causal reasoning. A → C is not interchangeable with C → A any more than *fire causes smoke* is interchangeable with *smoke causes fire.* The first is an established rule of physics; the second, a wizard's recipe. And so it is that, as the Turing Award–winning computer scientist Judea Pearl has shown in *The Book of Why,* the closest that the A = C brains of computers can get to causal reasoning is "if-then" statements:

> *If* Bob bought this toothpaste, *then* he will buy that toothbrush.
> *If* this route has a traffic jam, *then* the other route will be faster.
> *If* this chess move is played, *then* ninety-five percent of possible outcomes are victory.

If-then statements like these make up the bulk of Artificial Intelligence. And they do a good job of simulating casual reasoning. So good, in fact, that we humans tend to conflate the two in our ordinary speech. When we say, "if you're a smoker, then you're more likely to get lung cancer," we usually mean that smoking causes cancer. We're using "if-then" as a synonym for "cause-and-effect."

But cause-and-effect and if-then are *not* synonyms. Cause-and-effect encodes the *why* of causation, while if-then encodes the *that*-without-*why* of correlation. To take the example above, Bob buying toothpaste is *correlated* with him buying a toothbrush. But it doesn't *cause* him to buy a toothbrush. What causes Bob to buy a toothbrush is a third factor: wanting clean teeth.

Computers, for all their intelligence, cannot grasp this. Their if-then brains see no meaningful difference between Bob buying a toothbrush because he bought toothpaste and Bob buying a toothbrush because he wants clean teeth. In the language of a computer's logic gates, the two equate to the very same thing.

The inability of computers to employ the causal reasoning of the human brain doesn't mean that we humans are similarly unable to fathom the logic of computers. Our synapses have the capacity to establish bidirectional circuits and other network connections that can process identity and equation, so as C. S. Peirce observed, we can

marshal the symbolic thought of "2 + 2 = 4" and "peace is good" and "blue represents sorrow." Yet still, our synapses are more fluent in causal reasoning than in logic, and indeed, are so extraordinarily fluent that causal reasoning is a candidate for that wonder sought by Steven Spielberg: the unique element of human intelligence. The most distinctive feature of the human brain is its neocortex, and the most distinctive feature of the neocortex is its richness of synaptic connections. Those connections are on the order of ten thousand per neuron, many times that of the average animal brain cell. And since the human neocortex contains over 20 billion neurons (far more than any other species), our brains each possess tens of trillions of neocortical connections that can be stretched into long and branching chains of cause-and-effect.

This is why causal reasoning felt so intuitive to Immanuel Kant, J. S. Mill, and the other philosophers dismissed by Bertrand Russell. It's why Samuel Coleridge, A. C. Bradley, and millions of Victorian readers were able to speculate with such ease about Hamlet's motives. And it's why every one of us is capable of telling stories, concocting plots, and imagining alternative worlds. These endeavors are all organic extensions of our synaptic ability to think in cause-and-effect, the same ability that makes possible the biography of every individual and nation, the vision of every political reformer and entrepreneur, the foresight of every savings account and legal regulation, and the future purpose of every gallant sacrifice and altruistic act of kindness.

Not that our natural capacity for causal reasoning is all for the positive. Because our synapses don't obey the rules of logical proof, they're perfectly capable of hammering out counterfactual and even wholly fabricated narratives. And while such willingness to embrace the not-true has seeded the laurels of literary fiction, it's also the source of countless superstitions, prejudices, and attribution errors—"that black cat *caused* my bad luck"; "Immigrants *cause* crime"; "Her free choices were the *cause* of her suffering"—that we pass onto computers by demanding that they use their algorithms to map our bigotry-ridden human systems.[32]

Yet as much as our synapses can promote destructive ignorance, their capacity for causal reasoning does come with educational benefits, one of which is that we're all born scientists. We're not born *perfect* scientists; our youthful experiments are riddled with bias, limited sample sizes, unfalsifiable hypotheses, and other sources of blunder. But still, we intuitively employ the basic method of science from birth. We make predictive guesses about what causes what: *mother → milk → pleasure.* And we then test our guesses by putting them into action.[33]

As we saw above, this process of scientific learning does not yield absolute truths, and in fact, when we make the mistake of conflating science with truth (in the way that "Enlightenment" thinkers from Thomas Macaulay to Steven Pinker have done) we can veer into smugness, imperialism, and other habits antithetical to the curiosity, open-mindedness, and bias awareness necessary for effective science.[34] But even though scientific learning won't ever usher us into omniscience, it can convey a host of practical powers. It's how we learn to make fires, grow crops, and build robots. And no less crucially for our survival, it's how we learn to communicate. Communication begins with our brain's discovery: *this* gesture → *that* effect. And our brain's mental catalogue of signs and signals is then built out through more such causal discoveries

that enable us to express feelings, influence groups, and woo mates, until at last we achieve the feat of communication known as literature.

Literature isn't reducible entirely to communication; it involves other skills and can work to additional ends. But communication is *necessary* for literature. We cannot fashion poems or comic books or memoirs unless we learn to impart a feeling, an impression, an experience, or some other mental stuff to an audience. So to discover how literature operates, enabling us to put its powers to work in our lives, we have to draw on our inborn ability to perform causal reasoning. We have to make predictions about whether a poetic phrase, theatrical event, or novelistic character will convey a particular idea, mood, or emotion. And we then have to test that prediction by giving our literary creation to an audience. That is, we have to adopt the basic process of scientific learning, drawing hypotheses and running experiments that enable us to refine and enlarge the effects of our writing.

Without this process of scientific learning, we'd never be able to pen a short story or appreciate how poetry works. Which is why, to bring our narrative to its end, both tasks lie beyond the reach of computers.

Why Computers Cannot Write or Read Literature

The basic inability of computers to grasp the practical powers of literature can be glimpsed in the dire quality of algorithm-generated poems and short stories. The poems lack organic unity; one phrase follows another, follows another, follows another, without achieving a consistent narrative style or lyric voice. The stories, meanwhile, lack coherent characters or plots; one action follows another, follows another, follows another, without establishing any overarching psychological purpose or narrative direction. And although it's possible in the wake of intellectual movements such as postmodernism to convince ourselves that these robotic emissions possess some sort of literary substance, they're really no more than word soups. Those word soups are interpretively, mathematically, and symbolically equal to the works of Adrienne Rich and Maya Angelou. But practically, scientifically, and causally, they're not.

The current failure of machine-learning algorithms to author literature does not, of course, prove that computers will never author the next "Planetarium" or *I Know Why the Caged Bird Sings*. The relentless data acquisition of computer algorithms ensures that they can always surprise us with new feats of intellectual achievement: thirty years ago, there were no computer–chess grandmasters; now there are legions. But even so, we can be completely certain that no such mastery of literature will ever be achieved by a computer. We can be certain because the vast majority of literature's practical powers require a mode of communication that, unlike the rules of chess, cannot be reduced to symbolic logic.

This mode of communication is narrative. Narrative was uncovered over two millennia ago by the same Greek polymath who invented the syllogism: Aristotle. Aristotle observed in his *Poetics* (ca. 335 BCE) that Greek tragedies generated a trio of psychological effects—pity, wonder, fear—that he traced empirically to plot twists, character epiphanies, and other elements of narrative.[35] At which point, Aristotle

then anatomized narrative itself, discerning that it took the basic form of a *beginning* that caused a *middle* that caused an *end.* Narrative was, in other words, a chain of causes and effects that was produced by the author's mental powers of causal reasoning—and that required the audience's mental powers of causal reasoning to process.[36]

Aristotle's emphasis on the narrative sources of literary effects was continued by classical rhetoricians such as Quintilian, who emphasized narrative as distinct from—and typically more rhetorically effective than—logic (*Institutio Oratoria* 2.4.1). And although this insight into the communicative force of narrative was generally lost during the European Middle Ages, when rhetoric (like ball-throwing) was reduced to a branch of symbolic logic (McKeon 15), it would be rediscovered in part by Renaissance humanists such as Peter Ramus[37] and then in full by a mid-twentieth-century collective of American literary scholars who styled themselves the "Chicago School."[38] The Chicago School began by exhuming Aristotle's method of linking literary "causes" to psychological "effects": "In the method of Aristotle . . . poetics is a science concerned with the differentiation and analysis of poetic forms or species in terms of all the causes which converge to produce their respective emotional effects" (Olson, "An Outline" 9). And hewing closely to the "science" of Aristotle's "poetics," the Chicago scholars traced many of literature's "effects" back to "narrative," which they explicitly treated as a "causal chain" of "cause-effect."[39]

With this, the Chicago School restored Aristotle's ancient discovery that literary communication involved causal reasoning all the way down: it was built out of cause-and-effect narrative components such as plot and poetic style that were in turn the rhetorical causes of psychological effects. And having revived Aristotle's practical science of literature, the Chicago School then followed the classical rhetoricians in extending it. Elder Olson traced "suspense" to the storytelling technique of "partial disclosure" or "concealment" ("William Empson" 52). R. S. Crane traced "comic pleasure" to a literary character who *wants* to do the right thing, but who is too immature to grasp what the right thing is (Crane 81). More recently, the second-generation Chicago scholar Wayne Booth traced literary sympathy and irony back to different forms of narrative voice (*The Rhetoric of Fiction*). And the third-generation Chicago scholar James Phelan has explored the way that narrative can sway our brains' subtle emotional perceptions of what *feels* right or true.[40]

Because the literary powers uncovered by these scholars all require causal reasoning, none of them can be processed—let alone learned—by computers. Like a three-dimensional cosmos that comprehends all the intricacies of length, width, and depth but not the barest instant of the fourth dimension of time, the timeless symbolic logic of the computer brain is ontologically incapable of containing the *past-present-future* components of narrative. So, it ignores all such components—every beginning, middle, and end—acting as though they do not exist.

This vast blind spot was observed by the earliest Chicago scholars. Although those scholars were working prior to the advent of modern computing, they decried the interpretive logic of I. A. Richards's poetry robot as "a mechanical method . . . capable of all the mindless brutality of a machine" (Olson, "William Empson" 27). And the "brutality" they detected can be felt in twenty-first-century computer readings of literature. By focusing monomaniacally on semiotic meaning, those readings

skip over the characters, the plots, the storyworlds, and the narrators of literature. Which is to say, they skip over the *minds,* the *purposes,* the *experiences,* and the *voices* of literature—and with them, most of the stuff that our brains recognize as human.

This automated deletion of literature's fourth dimension means that no matter how many novels and poems and playscripts are digitized, and no matter how much data on human reading-responses is uploaded onto silicon hard drives, algorithms will forever remain incapable of identifying the narrative nuts and bolts that stimulate pity, wonder, fear, joy, empathy, laughter, irony, curiosity, suspense, or any of literature's other rhetorical effects. An infinite Arithmetic Logic Unit fed infinite literary information for infinite years will never achieve the faintest glimmer of insight into how dramatic characters touch the heart, how novelistic geographies capture the imagination, how television plots raise the pulse, or how poetic styles invest inert words with the timbre of human psyches.

To change this situation, we'd have to re-engineer the computer brain, ripping out the Arithmetic Logic Unit and replacing it with a processor modeled on the human neocortex. A processor that operates not by running bits and bytes through logic gates but by opening and closing synaptic junctions. A processor that thinks not in signs and symbols but in chains of causes and effects.

This may not seem so daunting a task. After all, there are shelves full of science fiction, from "Supertoys" to *Artificial Intelligence,* that encourage us to believe that it's possible to build robots that tell stories, hatch plans, and engage in other acts of causal reasoning. But if we wanted to build any such robots, we'd have to go far beyond repurposing the current circuitry found in computers. We'd have to invent wholly new technology.

To understand why this is so, compare the synapse with its nearest computer analogue: the transistor. The transistor is the component used to build the logic gates of the CPU's Arithmetic Logic Unit. And a lone transistor can, by itself, operate as a gate. So, like a synapse, it can open and shut connections between one wire and other, operating as a junction administrator.

But there's a key difference between the synapse and the transistor. The synapse is a physical device, constructed from cellular proteins, that opens and shuts by adjusting its shape. The transistor, meanwhile, is an electronic device: the channel it gates is an electrical wire, and it itself is regulated by voltage. This means that the transistor can only function within a system that obeys the laws of electronics: the system's current must remain within particular parameters, its overall circuit must be closed, its electrons must flow in precise patterns, etc. So, the system cannot be improvised willy-nilly from within. It requires an overall, unified design.

This need for design is perfectly compatible with symbolic logic, which obeys strict mathematical rules and involves a limited number of processes—AND, OR, NOT—that can be orchestrated ahead of time. Yet it conflicts fundamentally with causal reasoning's experimental method. That method can incorporate *some* design. But it cannot incorporate too much; its creative engine is fueled by trial-and-error, that is, by attempting things before there's certain evidence that they will work. And in fact, so important is trial-and-error to the method of causal reasoning that the method can proceed entirely without design. It can fumble around nonintention-

ally, anarchically embarking on a thousand random doings until it finds one that "works"—not in the sense of doing *exactly what* was aimed for, but in the more modest sense of accomplishing *anything* useful.

This blind process is how evolution by natural selection engineered the human neocortex, and it's how the human neocortex continues to operate inside our heads. Our neurons don't plot every detail of their causal reasonings in advance; indeed, our neurons often have only the dimmest sense of where their reasonings are going. But they grope forward, and in response to feedback, they then adjust, stumbling toward discoveries that are frequently very different from what we initially hoped to discover, but which are nevertheless practically valuable in the end.

To engage in that undesigned method, our brains require neural hardware that is similarly able to operate without design. That hardware consists of two primary components, the first of which is our synapses. Because our synapses are physical (not electronic) devices, they don't need to be part of a system that's planned in advance. They can be randomly snapped together like Lego blocks or a children's set of plastic gears—and then be tested without any fear of short circuit or meltdown, until a configuration is discovered to work.

This synaptic extemporization is supported by the second hardware component that distinguishes our neuroanatomy from computers: mitochondria. Mitochondria are intracellular organelles that provide each of our neurons with its own private supply of energy, allowing it to operate independently. So, unlike the computer CPU's etched metal wires, which are powered by an outside electron source that requires them to function all together as a single circuit, our individual neurons boast inner power generators that equip them to function as self-contained electric circuits that can be plugged physically into other neurons without causing changes in voltage. The resulting biological combo of self-powered wires and synaptic wire connectors enables our neural networks to exploit the quickness of electric transmission while liberating them from the need to conform to an overall electronic design. So at the same time as our brains harness the action potential "firing" of our neuronal axons to lend our thoughts a zip of electron speed, they can also wildly improvise their neuron-to-neuron connections, freestyling A → C speculations without ever frying our motherboards.

It's possible that in some future age, we'll invent nanosized power plants, non-electric signal relays, and other post-computer technologies that can mimic mitochondria, synapses, and the rest of the brain's architecture. But in the meantime, there will be no machines that can perform causal reasoning—and so no machines that can learn to read or write literature better than we do. Instead, our best hope to speed up the labor of literary studies is to remember what Aristotle discovered but I. A. Richards forgot: logic and narrative can be complementary. The former has helped birth mathematics and philosophy; the latter, science and the arts. And as different as the two are, there's no need to choose between. The research of Judea Pearl has demonstrated that machine-learning and causal reasoning can partner to advance scientific experiments ("A Probabilistic Calculus"). And the same basic insight can be found in the way that Samuel Coleridge, A. C. Bradley, the Chicago School, and

millions of ordinary readers have simultaneously mined literature for its interpretive meanings and its narrative empowerments.

Computers, it is true, can participate in only the first part of this process. But their vast capacity for symbolic logic gives them the potential, as the Digital Humanities has shown, to fast-track interpretation in extraordinary ways. So, by pairing that automated silicon ability with our own native intellectual strengths, we can continue the pattern-recognition magic that I. A. Richards accelerated ahead, while retaining the practical science that he cast aside. We can, that is, marry an AI chase of meaning to an exploration of the *minds,* the *purposes,* the *experiences,* and the *voices* that our synapses intuitively treasure in literature, so that we keep alive 1969's blue-sky dream of an upgraded tomorrow without forsaking the benefits of the human way that we already read.

A Logical Proof That Computers Cannot Read (or Write) Literature

1. Literature has a rhetorical function.[41]
2. Literature's full rhetorical function depends on narrative elements.[42]
3. Narrative elements rely on causal reasoning.[43]
4. Causal reasoning cannot be performed by machine-learning algorithms because those algorithms run on the CPU's Arithmetic Logic Unit, which is designed to run symbolic logic, and symbolic logic can only process correlation.[44]

QED: Computers cannot perform the causal reasoning necessary for learning to use literature.

Endnotes

1. For more on this term, see Liu 409.

2. See Moretti, *Distant Reading* (2013) and *Graphs, Maps, Trees: Abstract Models for a Literary History* (2005).

3. On the prehistory of the degree, see Palmer.

4. In the closing moments of *Artificial Intelligence,* the Specialist narrates: "Human beings had created a million explanations of the meaning of life in art, in poetry, in mathematical formulas."

5. In 1966, Giamatti published *The Earthly Paradise and the Renaissance Epic,* a classic New Critical study in the "ambiguity" of an "image" (6).

6. In 2014, Rudenstine capped his career with the publication of *Ideas of Order: A Close Reading of Shakespeare's Sonnets,* which, as its title announces, is an exercise in New Critical "close reading."

7. See, e.g., Abrams; and Gallagher.

8. See Lentricchia; Menand; and Kernan.

9. The first stirrings of this defense can be detected in eighteenth-century writings such as Thomas Robertson's "An Essay on the Character of Hamlet," which asserts that Hamlet's apparent inconsistencies of action have a deeper psychological consistency: "the latitude of his character" (255). But the defense's effective origins lie with the German Romantics of the early nineteenth century, especially August Wilhelm von Schlegel, who in 1797 began translating Shakespeare into German, and who, in 1798, began delivering the lectures that would develop into *Lectures on Dramatic Art and Literature* (1809–11).

10. In this same lecture ("Lecture 3, Thursday, 7 January 1819 (*Hamlet*)"), Coleridge finds it necessary to defend himself from the—highly plausible—complaint that he has plagiarized many of his ideas about Shakespeare from Schlegel.

11. See Ridley 4–5; and Cooke 181, 225.

12. For an engaging account, see Anellis.

13. This conclusion is unwarranted. As discussed below, human thought can contain other stuffs, including narrative.

14. See Rovee; Buurma and Heffernan; and Pasanek.

15. See Hill and Sula on the history of digital humanities. For the semiotic roots of Moretti's digital approach, see Moretti, *Signs Taken for Wonders.*

16. See, e.g., Davidson.

17. On our brain's natural drive to know *why,* see Loewenstein.

18. "The subject of Logic stands almost exclusively associated with the great name of Aristotle. As it was presented to ancient Greece in the partly technical, partly metaphysical disquisitions of the Organon, such, with scarcely any essential change, it has continued to the present day" (Boole 1). For discussion, see Corcoran.

19. For more on Boole and Shannon's influence on contemporary computing, see Nahin.

20. On the "circuitous" Shakespeare analysis of Richards's student F. R. Leavis, see Atherton 147.

21. On these cultural forces' influence on canon formation, see Guillory.

22. "In sorting through a vast heap of evidence for something interesting, we run a risk of cherry-picking" (Underwood 17).

23. The same goes for any act of interpretation, including the interpretations of data made by modern business CEOs, college administrators, etc.

24. The term "science" is used throughout this article to refer to the two-part method of prediction and experiment, further explained below in the section on "The Empowerment of Science's Why." In Coleridge's verbal usage, it blurs into Enlightenment science, discussed in the following footnote, and borrows elements and assumptions from symbolic logic. But the folk practice of character criticism is typically rooted not in logic but in the intuitive scientific method discussed below.

25. This view of science dates to Karl Popper's 1935 *Logik der Forschung,* translated into English in 1959 as *The Logic of Scientific Discovery* and expanded in works like *Conjectures and Refutations: The Growth of Scientific Knowledge* (1963). It has been challenged by later Enlightenment philosophers of science (e.g., Jean Bricmont and Alan Sokal, and Steven Pinker) who view it as a dangerous concession to postmodern relativism and reactionary theologies, and who prefer to revert to a version of the old Victorian notion of science as a gradual march toward truth. But the great practical value of Popper's view of falsification is that it promotes behaviors—such as humility, inclusiveness, and open-mindedness—that help ward off scientific complacency and absolutism,

producing a check against bias and nurturing an active search for areas where scientific knowledge can grow.

26. On some of the dangers of that legacy, see below, especially footnote 37.

27. "The Baconian concept of science, as an inductive science, has nothing to do with and even contradicts today's form of science" (Malherbe 75). Bacon's mistake (which was detected by Victorians such as John Herschel) was to rely purely on induction rather than on prediction and experiment (discussed below). And since induction is one of the logical techniques identified by Aristotle's *Organon,* Bacon's approach was itself semiotic; it treated nature as a "text" that could be "interpreted," routing Bacon's method back into the *why*-less tautologies that he wished to escape. For more on this, see Fletcher 2021, *Wonderworks* chapter 13.

28. The term "interpretation" is frequently deployed in modern science, but its technical meaning is synonymous with "prediction." That is, it's a causal-reasoning hypothesis about what *will* happen in a future experiment, making it fundamentally different from a semiotic interpretation about what *is* in a timeless equation.

29. For more on this, see O'Neil.

30. For more on this, see the papers in Waldmann.

31. "While the organism's developmental program assures that the connections between cells are invariant, it does not specify their precise strength. Rather, the strength and effectiveness of these preexisting chemical connections can be altered by experience" (Kandel 401–2).

32. On the relationship between computer algorithms and prejudices, see Noble.

33. "Causal inference is not merely a way of representing and updating probabilities; it is not merely Bayesian inference. Human causal inference involves the construction of narratives that unfold over time and determine the focus of attention, narratives that reflect knowledge of the specific mechanisms that drive effects" (Lagnado and Sloman 236). See also Graesser, Singer, and Trabasso; and Bloom and Fletcher.

34. For a piercing analysis of the disasters that can result from conflating Baconian power with philosophical "truth," see Adorno and Horkheimer. Crucially, however, there is nothing intrinsically negative about power itself. Power is necessary for every biological act; without the power of cellular ATP, there would be no thinking, no doing, no life. The problem, from the perspective of human society, is with power unchecked or unregulated.

35. On Aristotle's scientific method in the *Poetics,* see Fletcher, "The Lost Optimism of Modern Movie Fairytales."

36. Aristotle's beginning-middle-end of narrative neatly follows the anatomical beginning-middle-end of neuronal A $\rightarrow$ C causal reasoning: neuron A (beginning)—synapse (middle)—neuron C (end).

37. Ramus rejected medieval logic for being "useless," replacing it with a practical science that strove to unlock the rhetorical powers of literature: "The cause is the power by which a thing is; and therefore the discovery of this power is the font of all sciences: comprehended utterly is the thing whose cause is grasped: so that the poet rightly says: *happy is he who can know the causes of things.*" (14; *Causa est, cuius vi res est. Itaque primus hic locus inventionis, fons est omnis scientiae: scirique demum creditor, cuius causa teneatur: ut merito dicatur a Poeta: "Felix, qui potuit rerum cognoscere causas.*"). With this, Ramus laid the groundwork for Bacon's practical science of causes, explored above. For a summary of Ramus's influence on Bacon, and thereby on modern science's break with symbolic logic, see Fletcher, "Francis Bacon's Forms and the Logic of Ramist Conversion."

38. For a fuller history of this collective, see Phelan, "Chicago School."

39. "*Cause-effect.*—When we see a causal chain started we demand . . . to see the result. . . . This kind of sequence, so strongly stressed by Aristotle in his discussion of plot, is . . . often underplayed or even deplored by modern critics" (Booth 126).

40. On the former, see the exploration of ethics in Phelan, *Narrative as Rhetoric: Technique, Audiences, Ethics, Ideology* (1996); on the latter, see the exploration of probability in *Somebody Telling Somebody Else: A Rhetorical Poetics of Narrative* (2017), where Phelan roots probability in "a concern with causality" (44).

41. See Aristotle, *Poetics* 1452a1; and Quintilian, *Institutio Oratio* 10.1.46ff.

42. See Olson, "William Empson"; Crane; Booth; and Phelan, *Narrative as Rhetoric* and *Somebody Telling Somebody Else.*

43. See Lagnado and Sloman; and Graesser, Singer, and Trabasso.

44. See Pearl, "A Probabilistic Calculus" and *Causality*; and Mackenzie and Pearl, *The Book of Why.*

Works Cited

Abrams, M. H. "The Transformation of English Studies: 1930–1995." *Daedalus* 126 (1997): 105–31.

Adorno, Theodor W., and Max Horkheimer. *Dialectic of Enlightenment: Philosophical Fragments,* edited by Gunzelein Schmid Noerr. Translated by Edmund Jephcott. Stanford: Stanford Univ. Press, 2002.

Aldiss, Brian. "Supertoys Last All Summer Long." *Harper's Bazaar UK,* December 1969.

Anellis, Irving H. "Peirce Rustled, Russell Pierced: How Charles Peirce and Bertrand Russell Viewed Each Other's Work in Logic." *Modern Logic* 5 (1995): 270–328.

Aristotle. *Poetics.* Edited by Leonardo Tarán and Dimitri Gutas. Leiden: Brill, 2012.

Atherton, Carol. *Defining Literary Criticism: Scholarship, Authority and the Possession of Literary Knowledge 1880–2002.* Basingstoke: Palgrave, 2005.

Bacon, Francis. *Instauratio Magna.* London: Joannem Billium, 1620.

Bate, Jonathan. *The Romantics on Shakespeare.* London: Penguin, 1992.

Binder, Jeffrey M., and Collin Jennings. ""A Scientifical View of the Whole": Adam Smith, Indexing, and Technologies of Abstraction." *ELH* 83 (2016): 157–80.

Bloom, Charles P., and Charles R. Fletcher. "Causal Reasoning in the Comprehension of Simple Narrative Texts." *Journal of Memory and Language* 27 (1988): 235–44.

Boole, George. *An Investigation of the Laws of Thought.* London: Macmillan, 1851.

Booth, Wayne C. *The Rhetoric of Fiction.* Chicago: Univ. of Chicago Press, 1983.

Bradley, Andrew Cecil. *Shakespearean Tragedy: Lectures on Hamlet, Othello, King Lear, Macbeth.* 2nd ed. London: Macmillan, 1905.

Bricmont, Jean, and Alan Sokal. *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science.* New York: Picador, 1998.

Buurma, Rachel Sagner, and Laura Heffernan. "Search and Replace: Josephine Miles and the Origins of Distant Reading." *Modernism / Modernity Print Plus* 3 (2018).

Coleridge, Samuel Taylor. *Coleridge: Lectures on Shakespeare (1811–1819),* edited by Adam Roberts. Edinburgh: Edinburgh Univ. Press, 2016.

Cooke, Katherine. *A. C. Bradley and His Influence in Twentieth-Century Shakespeare Criticism.* Oxford: Clarendon, 1972.

Corcoran, John. "Aristotle's *Prior Analytics* and Boole's *Laws of Thought.*" *History and Philosophy of Logic* 24 (2003): 261–88.

Crane, R. S. "The Concept of Plot and the Plot of *Tom Jones.*" In *Critics and Criticism: Essays in Method,* edited by R. S. Crane, 62–93. Chicago: Chicago Univ. Press, 1957.

Davidson, Cathy N. "Humanities 2.0: Promise, Perils, Predictions." *PMLA* 123 (2008): 707–17.

Fletcher, Angus. "Francis Bacon's Forms and the Logic of Ramist Conversion." *The Journal of the History of Philosophy* 43 (2005): 157–69.

———. "The Lost Optimism of Modern Movie Fairytales." In *The Oxford Handbook of Psychological Approaches to Film,* edited by James Pawelski. Oxford: Oxford Univ. Press, 2021.

———. *Wonderworks: The 25 Most Powerful Inventions in the History of Literature.* New York: Simon and Schuster, 2021.

Gallagher, Catherine. "The History of Literary Criticism." In *American Academic Culture in Transformation: Fifty Years, Four Disciplines,* edited by Thomas Bender and Carl E. Schorske, 151–72. Princeton: Princeton Univ. Press, 1997.

Giamatti, Angelo Bartlett. *The Earthly Paradise and the Renaissance Epic.* Princeton: Princeton Univ. Press, 1966.

Gold, Matthew K., and Lauren F. Klein. *Debates in the Digital Humanities: 2019.* Minneapolis: Univ. of Minnesota Press, 2019.

———. "A DH That Matters." In *Debates in the Digital Humanities: 2019,* edited by Matthew K. Gold and Lauren F. Klein, ix–xiv. Minneapolis: Univ. of Minnesota Press, 2019.

Graesser, A. C., M. Singer, and T. Trabasso. "Constructing Inferences During Narrative Text Comprehension." *Psychological Review* 101 (1994): 371–95.

Greenblatt, Stephen. "Murdering Peasants: Status, Genre, and the Representation of Rebellion." *Representations* 1 (1983): 1–29.

Guillory, John. *Cultural Capital: The Problem of Literary Canon Formation.* Chicago: Chicago Univ. Press, 1993.

Hill, Heather V., and Chris Alen Sula. "The Early History of Digital Humanities: An Analysis of *Computers and the Humanities* (1966–2004) and *Literary and Linguistic Computing* (1986–2004)." *Digital Scholarship in the Humanities* 34 (2019).

Joseph, Bertram. "The Problem of Bradley." *The Use of English* 5 (1953–54): 87–91.

Kandel, Eric R. "The Molecular Biology of Memory Storage: A Dialog between Genes and Synapses." In *Nobel Lectures in Physiology or Medicine 1996–2000,* edited by Hans Jornvall, 392–439. London: World Scientific, 2003.

Kernan, Alvin B. *What Happened to the Humanities?* Princeton: Princeton Univ. Press, 2014.

Knights, Lionel Charles. *How Many Children Had Lady Macbeth: An Essay in the Theory and Practice of Shakespeare Criticism.* Cambridge: The Minority Press, 1933.

Lagnado, Donald, and Steven A. Sloman. "Causality in Thought." *Annual Review of Psychology* 44 (2015): 223–47.

Lentricchia, Frank. *After the New Criticism.* Chicago: Univ. of Chicago Press, 1980.

Litz, Walton A., Louis Menand, and Lawrence Rainey, eds. *The Cambridge History of Literary Criticism. Volume VII: Modernism and the New Criticism.* Cambridge: Cambridge Univ. Press, 2000.

Liu, Alan. "The Meaning of the Digital Humanities." *PMLA* 128 (2013): 409–23.

Loewenstein, G. "The Psychology of Curiosity—A Review and Reinterpretation." *Psychological Bulletin* 116 (1994): 75–98.

Macaulay, Thomas Babington. *The History of England, From the Accession of James the Second, Volume I.* London: Longman, Brown, Green, and Longmans, 1849.

Mackenzie, Dana, and Judea Pearl. *The Book of Why: The New Science of Cause and Effect.* New York: Basic Books, 2018.

Malherbe, Michel. "Bacon's Method of Science." In *The Cambridge Companion to Bacon,* edited by Markku Peltonen, 75–98. Cambridge: Cambridge Univ. Press, 1996.

McCarthy, John and Patrick J. Hayes. "Some Philosophical Problems from the Standpoint of Artificial Intelligence." *Machine Intelligence* 4 (1969): 463–502.

McKeon, Richard. "Rhetoric in the Middle Ages." *Speculum* 17 (1942): 1–32.

Menand, Louis. *The Marketplace of Ideas: Reform and Resistance in the American University.* New York: Norton, 2010.

Moretti, Franco. *Distant Reading.* London: Verso, 2013.

———. *Graphs, Maps, Trees: Abstract Models for a Literary History.* London: Verso, 2005.

———. *Signs Taken for Wonders: On the Sociology of Literary Forms.* London: Verso, 2005.

Nahin, Paul J. *The Logician and the Engineer: How George Boole and Claude Shannon Created the Information Age.* Princeton: Princeton Univ. Press, 2012.

Noble, Safiya U. *Algorithms of Oppression: How Search Engines Reinforce Racism.* New York: New York Univ. Press, 2018.

Ogden, C. K., and I. A. Richards. *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism.* New York: Harcourt, Brace, and Company, 1923.

Olson, Elder. "An Outline of Poetic Theory." In *Critics and Criticism: Essays in Method,* edited by R. S. Crane, 3–23. Chicago: Chicago Univ. Press, 1957.

———. "William Empson, Contemporary Criticism, and Poetic Diction." In *Critics and Criticism: Essays in Method,* edited by R. S. Crane, 24–61. Chicago: Chicago Univ. Press, 1957.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.* New York: Crown Random House, 2016.

Palmer, D. J. *The Rise of English Studies: An Account of the Study of English Language and Literature from Its Origins to the Making of the Oxford English School.* London: Oxford Univ. Press, 1965.

Pasanek, Brad. "Extreme Reading: Josephine Miles and the Scale of the Pre-Digital Digital Humanities." *ELH* 86 (2019): 355–85.

Pearl, Judea. *Causality: Models, Reasoning, and Inference.* Cambridge: Cambridge Univ. Press, 2000.

———. "A Probabilistic Calculus of Actions." In *Uncertainty in Artificial Intelligence* 10, edited by R. Lopez de Mantaras and D. Poole, 454–64. San Mateo: Morgan Kaufmann, 1994.

Peirce, Charles S. *Writings of Charles S. Peirce: A Chronological Edition, Volume 5, 1884–1886,* edited by Christian J. W. Kloesel. Bloomington: Indiana Univ. Press, 1993.

———. *Semiotics and Significs.* Edited by Charles Hardwick. Bloomington: Indiana Univ. Press, 1977.

Phelan, James. "The Chicago School: From Neo-Aristotelian Poetics to the Rhetorical Theory of Narrative." In *Theoretical Schools and Circles in the Twentieth-Century Humanities,* edited by Marina Grishakova and Silvi Salupere, 133–51. New York: Routledge, 2015.

———. *Narrative as Rhetoric: Technique, Audiences, Ethics, Ideology.* Columbus: The Ohio State Univ. Press, 1996.

———. *Somebody Telling Somebody Else: A Rhetorical Poetics of Narrative.* Columbus: The Ohio State Univ. Press, 2017.

Pinker, Steven. *Enlightenment Now: The Case for Reason, Science, Humanism, and Progress.* New York: Penguin, 2018.

Piper, Andrew. "Fictionality." *Journal of Cultural Analytics.* December 20, 2016.

Popper, Karl. *Conjectures and Refutations: The Growth of Scientific Knowledge.* London: Routledge, 1963.

———. *Logik der Forschung.* Translated by the author as *The Logic of Scientific Discovery.* London: Hutchinson and Co., 1959.

Quintilian. *Institutio Oratoria,* edited and translated by Harold Edgeworth Butler. Cambridge, MA: Harvard Univ. Press, 1920–22.

Ramus, Peter. *Dialecticae libri duo.* London, 1574.

Richards, I. A. *Practical Criticism.* London: Kegan Paul, Trench, Trubner & Co., 1929.

Ridley, M. R. *Shakespeare's Plays: A Commentary.* London: J. M. Dent, 1937.

Robertson, Thomas. "An Essay on the Character of Hamlet." *Transactions of the Royal Society of Edinburgh* 2 (1788): 251–67.

Roush, Wade. "The Supple Synapse: An Affair that Remembers." *Science* 274 (15 November 1996): 1102–3.

Rovee, Christopher. "Counting Wordsworth by the Bay: The Distance of Josephine Miles." *European Romantic Review* 28 (2017): 405–12.

Rudenstine, Neil L. *Ideas of Order: A Close Reading of Shakespeare's Sonnets.* New York: Farrar, Straus, and Giroux, 2014.

Russell, Bertrand. "On the Notion of Cause." *Proceedings of the Aristotelian Society* 13 (1912–1913): 1–26.

———. *The Principles of Mathematics, Volume I.* Cambridge: Cambridge Univ. Press, 1903.

Russell, Bertrand, and Alfred North Whitehead. *Principia Mathematics.* Cambridge: Cambridge Univ. Press, 1910–13.

Russo, John Paul. *I. A. Richards: His Life and Work.* London: Routledge, 2015.

Schlegel, August Wilhelm. *Lectures on Dramatic Art and Literature, 1809–1811,* edited by Stefan Knödler. Originally published as *Vorlesungen über dramatische Kunst und Literatur, 1809–1811.* Paderborn: Ferdinand Schöningh, 2018.

So, Richard Jean. "All Models Are Wrong." *PMLA* 132 (2017): 668–73.

Spielberg, Steven, dir. *A. I. Artificial Intelligence.* 2001; USA: Warner Bros.

Underwood, Ted. *Distant Horizons: Digital Evidence and Literary Change.* Chicago: Univ. of Chicago Press, 2019.

———. "The Life Cycles of Genres." *Journal of Cultural Analytics* (23 May 2016)*.*

Waldmann, Michael, ed. *The Oxford Handbook to Causal Reasoning.* Oxford: Oxford Univ. Press, 2017.

Ziolkowski, Jan M. "Cultures of Authority in the Long Twelfth Century." *The Journal of English and Germanic Philology* 108 (2009): 421–48.