


Please cite the Published Version

Pollard, Eileen  (2024) Back to the future: everything you wish you'd asked Derrida about Chat-GPT when you had the chance! Cultural Studies - Critical Methodologies. ISSN 1532-7086

DOI: <https://doi.org/10.1177/15327086241232722>

Publisher: SAGE Publications

Version: Published Version

Downloaded from: <https://e-space.mmu.ac.uk/634089/>

Usage rights:  [Creative Commons: Attribution-Noncommercial 4.0](https://creativecommons.org/licenses/by-nc/4.0/)

Additional Information: This is an open access article which first appeared in Cultural Studies - Critical Methodologies, published by SAGE Publications

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

Back to the Future: Everything You Wish You'd Asked Derrida About ChatGPT When You Had the Chance!

Cultural Studies ↔ Critical Methodologies
1–8
© 2024 SAGE Publications



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/15327086241232722
journals.sagepub.com/home/csc



Eileen Pollard¹ 

Abstract

This article considers and then reconsiders what ChatGPT produces and how it produces it, using the work of a range of critical theorists and authors. In particular, it imagines what different philosophers, thinkers, and writers would say about this most recent technological leap, if they were somehow brought *back* from the past, into this, our new future. To ventriloquise for them, this article plays fast and loose with the work and styles of Douglas Adams, Virginia Woolf, and Alan Turing, among others, to try to demonstrate what ChatGPT can do, having been potentially “trained” on their work, as well as highlighting the nuances of allusion, subtext, paradox, and contradiction as possibly more human aspects of both writing and reading. Such “play” is followed by a more serious analysis of writing and the suggestion of a “Double Signature Signification” at work in the text produced by ChatGPT, meaning one system of signification for writing the text (schematic) and one for reading it (referent), which overlap perfectly. The article concludes by arguing that it is not the consciousness of ChatGPT that beguiles us, it is the *possibility* of that consciousness, and what gives rise to that sense of possibility is partly the spectral and haunting nature of dialoguing with it. A dialogue with ChatGPT has all the excitement of a séance: it is uncertain, unknown, yet with its traces of the familiar it is also like talking to the dead.

Keywords

generative AI, rethinking critical theory, methods of inquiry, creativity methodologies, double signature signification, discourse, politics, and culture

Epigraphs

We do not know our own souls, let alone the souls of others. Human beings do not go hand in hand the whole stretch of the way. There is a virgin forest in each; a snowfield where even the print of birds' feet is unknown. Here we go alone, and like it better so. Always to have sympathy, always to be accompanied, always to be understood would be intolerable.

Virginia Woolf, *On Being Ill*,¹ 1926.

The text is a tissue of quotations drawn from the innumerable centres of culture. Once the Author is removed, the claim to decipher a text becomes quite futile. To give a text an Author is to impose a limit on that text, to furnish it with a final signified, to close the writing.

Roland Barthes, *The Death of the Author*, 1967.

In the seminar Turing was challenged about how a thinking machine would cope with a contradiction, and said that it would backtrack until it found the error: Jefferson said this wasn't how humans thought, and Turing responded that it was exactly how mathematicians thought. It was at this point that the murmur, “but is a mathematician human?” was heard in the room.

Jonathan Swinton, *Alan Turing's Manchester*, 2022.

To Begin at the Beginning

*“Forty-two,” said Deep Thought, with infinite majesty and calm. (Douglas Adams, *The Hitchhiker's Guide to the Galaxy*, 2020)*

The literary reader will notice a number of allusions in this journal article, like the heading above, which is a reference to the opening of Dylan Thomas' *Under Milk Wood* (1954). Such deliberate pastiche highlights that there can be a *layer* to texts that ChatGPT cannot consciously add, or perceive, a *tissue* of meaning perhaps that is still potentially in the human domain. A literary reader, for instance, is a reader with a particular position based on previous reading experiences, raising the question, what kind of “reader” is

¹Manchester Metropolitan University, UK

Corresponding Author:

Eileen Pollard, Manchester Metropolitan University, All Saints Building, Grosvenor Square, Manchester M15 6BH, UK.
Email: e.pollard@mmu.ac.uk

ChatGPT? And anyone doubting the influence of literature in this technological space need only think of Deep Thought, quoted above from *The Hitchhiker's Guide*, which inspired the name of the first chess playing computer and then the more well-known later one, Deep Blue. It is significant to this article that chess was the stage on which computers were to perform “thought” and that the irony of Adams’ naming his computer *Deep Thought*, a computer that calculates the meaning of life as 42, was clearly lost when life tried poorly to imitate art.

And the title of this journal article too is, of course, a homage to Slavoj Žižek’s *Everything You Always Wanted to Know About Lacan (But Were Afraid to Ask Hitchcock)* published in 1992. But it is important to note that the second part of the title does *not* suggest that this article will examine everything you wished you asked ChatGPT about Derrida: although you are free to undertake such an exercise if you want, at your own risk. Like Žižek’s book, this journal article is experimental, a piece of speculative fiction/philosophy based loosely on the idea of going back in time and bringing critical friends like Jacques Derrida, Michel Foucault, and Jean Baudrillard back to the future, or rather, into our ever fast-changing present. What would they make of ChatGPT? Would they be surprised? And what questions would they ask? This magic realism is also informed by the thinking of famous writers too, especially as ChatGPT has most likely been “trained” on their work, without acknowledgment, or indeed permission. What would Virginia Woolf make of that, for example? Finally, to make a further point about writing, this article is written in an old-fashioned yet academically freestyle, which in itself is a nod to Turing’s seminal, “Computing Machinery and Intelligence.” Turing’s eccentric and creative publication, produced in a naive pre-REF, pre-metrics age, is the basis for the so-called “Turing Test” of whether or not a computer could trick a human into thinking that it (the machine) was human.²

And here I interrupt myself with a question: Is it important to be polite? Bing Chat says, “No need to say please in prompts. It’s a machine . . .” Have they not watched *Battlestar Galactica*?³ So no manners and possibly satire is dead too. I asked ChatGPT to “write a satirical biography of Winston Churchill.” This is what it came up with:

Winston Churchill was a man of many contradictions. He was born in 1874 in Blenheim Palace, a sprawling estate in England, where he was educated by private tutors and nurtured by his aristocratic parents. Despite his privileged upbringing, Churchill was a rebellious child who struggled in school and often found himself in trouble with the authorities.

Now compare with my own satirical biography of myself for an imaginary book cover:

Eileen Pollard does have a PhD but feels mentioning it here is about as relevant as saying she has a driving license (which she

also has). She once tried, and failed, to become a National Teaching Fellow and has since given up on the idea. She is not a Reader, although she did learn to read at a school where not everyone did. She has many passionate beliefs about education, but has resisted reducing them to an odd sort of shopping list for this book cover. Eileen remains very proud of her level two swimming certificate, which she was basically given out of pity upon leaving primary school.

Reader, dear reader, welcome to this article.

What I Think About When I Think About Thinking

In *On Being Ill*, Virginia Woolf (2002) defined the soul as unknowable, inaccessible, yet present in all of us. The workings of ChatGPT is in some ways similar, unknowable yet present, so is it possible that ChatGPT⁴ has a soul? A virgin forest within? Different to ours, naturally, but perhaps equal? And on the other hand of the Hegelian dialectic, is it conscious? Can it *think*? And what about thinking? After all thinking is not “neutral” (Aristotle, Kant, and Hegel) the way we understand what it is “to think” is cultural, informed by a history of ideas and having ideas and framed by what constitutes knowledge in different disciplines and subjects.⁵ Significantly, it was mathematics, rather than any other discipline (physics, for example) that both informed and indeed *drove* the desire to recreate “thinking” and the recreation of “thinking” in a thinking *machine*, specifically. Although this thinking was MORE than calculations, it was still fundamentally a view of thinking informed by mathematics, as well as *underpinned* by mathematics in a more mechanistic way. Mathematics made it all technically possible, but the thinking, cultures, and agendas of mathematicians also directed and informed what was actually DONE. But maths is only one way of *thinking*, so there were inherent limitations—and dangers—from the start. Notably, there is a gap between this way of thinking and other cultures of thought, with a further side effect, perhaps, of presenting this way of thinking as the only *way* of thinking, or at the very least a superior way. For example, the Turing Test presents *trickery* as thinking: and humans do trick one another, but this is not the only way they can be said to think, and is arguably not a particularly sophisticated form of thought. If you disagree, think about the last time you were scammed.

And what of writing? Cards on the table, ChatGPT’s clichéd sonnet is a master class in exhausted ventriloquism. Did you see what I did there? But seriously, what would Barthes make of the fact that ChatGPT literally chews up tissues of text, then patches them back together again, the Author, lower or upper case “a,” well and truly removed? In particular, this article spotlights some writing techniques and styles that ChatGPT currently struggles to emulate,

which demonstrates, in practice, how text produced by a human still, at the time of writing, differs from that “written” by an algorithm. As such, I offer a short, playful textual analysis interruption to this rapidly evolving *Apocalypse Now* movie, as well as a brief consideration of why, culturally, change often occurs seemingly at a distance from context, impact, or consequences. Indeed, Zygmunt Bauman observed, some years ago, in *Liquid Times* (Bauman, 2007) that “progress” is so accelerated now, that organizations and institutions outgrow their strategies, infrastructures, processes, even buildings, before they are finished: as their narratives are being written, they are outstripped by events, made redundant, in real time. Never was this motion sickness more the case than thirteen years after *Liquid Times* was published, during the pandemic, when policy after policy after policy was scrapped as it was issued. But wait, back up: taking a step back from impact and consequences for a moment, or even context, what about what ChatGPT actually produces? Not functionally, per se, but philosophically. What is the interface between the technology’s text and its apparent parallels with how humans write as well as, crucially, the differences? As language is central, or rather text, this analysis is underpinned by Derrida’s infamous *fait accompli*: “There is nothing outside the text.” And reading ChatGPT text—like all the best products nowadays—is an experience! Its mimicry of “real time” writing, for example, is a particularly seductive, almost-haptic addition. Does it make it *more* like writing, we wonder, appearing one word at a time? More human? AI, pausing for thought? And if so, why do that? And what is the effect of that?

Yet interest in its possible human qualities seems the wrong question, or the wrong one to start with at least. Generative Artificial Intelligence has more in common with the dead and ghosts in a strange way than it does with *living* human beings. In *Ghost Dance* (McMullen, 1983), a very odd film starring Derrida as “himself,” Derrida’s argument (here and elsewhere) is that telecommunication actually *increases* the possibility (rather than reducing it, as science always promises) for ghosts and their ability to haunt us. And what other metaphor than ghostliness is there for this current zeitgeist? ChatGPT and what it writes is literally haunting us, and, at the same time, is *itself* haunted by the ghosts of writers of Christmases past. It cannot “write,” in the way we might say of a person, “oh, they can *really* write”—because there is no satire, no irony, no paradox—but it is still haunting, spectral, unputdownable. We are much more disturbed by technology that can write than technology that can speak, perhaps because this speech is not obviously a human voice (yet), or perhaps because the world, as we are discovering again, is actually discursive. Attendance at ChatGPT workshops is still really high and will remain so. Although surely everyone’s said everything

now! Haven’t they? All that remains is to misquote Marx: there is a specter haunting the globe, the specter of AI. . .

Because everything has been said, or written, or asked of ChatGPT, surely? Except perhaps a textual interruption, a poststructural analysis, although I am sure those are coming soon (listen to The Briefing Room, BBC Radio 4, “The AI Revolution”) and possibly with a “mind-bending” Netflix series to follow, or *Everything Everywhere All at Once*, featuring AI! So, let us begin again, with the simple fact that ChatGPT is text-based. Derrida argued that writing came first, then speech. He saw it all as *writing*, all of it, hence, *il n’y a pas de hors-texte*. If we think about language acquisition, the child speaks then writes, yes. But as we go through the world, writing more and more informs the way we speak, and the way we think. And the language we speak changes and evolves through what is written and the way it is written. Even those who cannot write themselves are affected by these dominant changes. Writing is everywhere, inscribed on buildings and vans and signs, but it is also the filter in our heads, the language we think in—always informed by the dominance, the unstoppable movement of what is written and written, has been written, is being written, especially in our overwhelming, interruptive digital age—we cannot think outside of it, thus there *is* nothing outside of it. In Mantel’s *Wolf Hall* trilogy (2009–2020), Thomas Cromwell envies those who lived in slower times, before the advent of the printing press. Imagine that! Thinking that the Tudor period was fast, liquid with potentially false, spreadable misinformation. Let that sink in.

And you might not believe me about writing, so try a thought experiment. Think something. Anything. Anything at all.

You thought it in language, right? Or you were confused, but in language. And that language, in your head, was put there and informed by the writing, and the writing and the writing and the writing, behind the language evolution, that long, *long* predates you (apologies, I know this is hard). This discursive swamp swallows up writing that you have actually read *Your Self*, that has informed what and how you speak, and has, from day one—even before it really. OK, so if you are still with me reader (and not on ChatGPT instead) then writing—these signs and symbols we exchange and use, the words we agree on, and the arbitrary relationship they have with what they describe—is very privileged, yes, but it is also *very* easily produced, replicated, mimicked, systematized, and basically, churned out: first by moveable type printers, then typewriters, then keyboards, then computer printers, and now algorithms. But the meaning, the meaning is brought to it by humans, and it escapes in a ghostly fashion. It cannot be fixed, or pinned down, it is open to interpretation, always, and that is what makes it spectral and indeed what makes it possible.

It Really Doesn't Know What a Dog Is: How and Why ChatGPT Messes With Your Head

For some people a language, reduced to its essentials is a nomenclature: a list of terms corresponding to a list of things [. . .] This conception is open to a number of objections. It assumes that ideas already exist independently of words. It does not clarify whether the name is a vocal or psychological entity [. . .] Furthermore it leads one to assume that the link between a name and a thing is something quite unproblematic, which is far from being the case. None the less, this naive view contains an element of truth, which is that linguistic units are dual in nature, comprising two elements.⁶ Ferdinand de Saussure (1986) *“Course in General Linguistics”* (p. 65)

Shall we begin a third time with an analogy? In 1829, Stephenson's Rocket started to move. The steam engine propelling its wheels and, therefore, the locomotive itself, at a whopping thirty miles an hour! Then, and only then, came the realization of the need for brakes. Is this not all a bit ChatGPT? Progress (or profit) before consequences? The technology is moving, moving at speed, as we rush around it panicking, throwing our collective hands in the air, shouting, “Make it stop! Make it stop!” “What about academic integrity?” “What about our discipline?” “What about ethics?” “What about a moratorium?” “What about democracy, economics, power, geopolitics???” The cries and misunderstandings become ever more urgent and increase in scope and impact, literally, with each passing day. It is utterly dizzying trying to keep up.

And yet, I⁷ cannot help thinking that it is because it can *write* that we are panicking so much. Roland Barthes, famously, coined the terms, “first order signification,” meaning verbal signification and “second order signification,” meaning the image.⁸ Generative AI has mastered both. Or has it? The code (referent signification) has been trained and then that training has been massively upscaled and repeated (iterative, iteration) on huge, compendious, and varied corpuses of text, leading to many, many corrections to the code. The gamechanger is of course, again, language-oriented, in that the front end is now no code, making it accessible to anyone.

The key to what I describe below as “Double Signature Signification” is the basic fact that the signification underpinning the way the text is WRITTEN (by AI) is different to but symbiotic with the signification that allows it to be READ (by humans). A note first on “signature” though, Derrida was fascinated by signatures:

[He] focuses on “Signature” as an indicator and measure of the presence of the writer or author. The signature denotes the writer as the source of the text, or the speaker of an utterance, and they hold the form of regulation for the sign or words which are spoken or written.

But there is a *double* signature in AI-produced text, and in more than one sense—the signature of two systems of *signification* and the signature of AI (as an author) and the trace signature of all the human authors of both the source code and the texts that it was trained on.

In philosophy, the basic epistemological questions are “What is thinking?” “What is intelligence?” “What is creativity?” “What is writing?” and “What is reading?” Consider again the origin of AI in mathematics: How would a mathematician answer these questions compared with a social scientist? How would a mathematician define “writing”? Generative AI can produce text, but can it appreciate multivalent meaning? Or narrative? Or how humans construct narratives, telling stories to make sense of themselves? Basically, AI ascribes a quantitative VALUE to a word, whereas a human ascribes a qualitative MEANING.⁹ Hence, on the basis of value (calculated through compendious amounts of training), AI predicts that the next word in the phrase, “it's raining cats and. . .” is “dogs,” but does it know what “dogs” actually *means*? Can it point to a dog, for example? The word, or sign, “dog,” can have a VALUE and a MEANING because its relationship to both is arbitrary: There is nothing dog-like about the word dog, it does not look like a dog, the letters d-o-g do not sound like a dog. And likewise, the VALUE d-o-g has within ChatGPT (following its billionth training on use of “dog”) has no absolute relationship with a real dog either. Thus, symbiosis between the signification underpinning the way the text is written, and the signification determining the way the text is read, as the VALUE is calculated on the many, many instances in which MEANING has been expressed, reexpressed, reproduced, and iterated.

This arbitrariness, this *hollowness* to words, is the crux of the matter and what allows Double Signature Signification to operate. For AI, USE of a word determines VALUE. For humans, USE of a word determines MEANING. If AI calculates that the next word in the phrase “raining cats and. . .” is “dogs,” on the basis of USE of that word across billions of texts, it ascribes the letters d-o-g with a PROBABILITY that could be seen as a SIGNIFIED. But does AI know that the letters d-o-g, in that order, SIGNIFIES a furry animal with four legs and a tail? Sure, it could draw together a description of a dog (based on predictions of next-word-probability, as aforementioned) but could it be said to KNOW that d-o-g SIGNIFIES a dog? Not a picture, also based on the word or description recurring with those images, but an actual *dog*. It could not just point to a dog. It could only bring up representations of dogs in descriptions or images based on probability-based calculations about use. AI therefore has a different way of using language that dovetails with ours (after all, it was trained billions and billions of times on *our* texts written in so-called “natural” language). However, crucially, AI-produced text is NOT

based on our understanding of signification—that is, learn what a word means by someone pointing to its SIGNIFIED—in AI a word’s signified is a number¹⁰ based on previous use of that word in human-authored text and its next-word-prediction (hilariously called hallucinating!) is decided on the probability that that word should be used there (rather than because of what the word points to, or signifies). To us, it is a chain of signification. To AI, it is a series of best guess predictions calculated on use-probability.

Human: “dog,” signified = furry animal with four legs and a tail.

AI: “dog,” signified = use-probability (it has been used three billion times here, so the probability it should be used here is x, and that value is “closed” within the system—it is not outward, it does not advertise itself).

It is the same and not the same. It is parallel but different. We, as humans, read the text in the same way as human-authored text, because the words are the same (Searle’s Chinese Room). These words are little vessels for meaning, but the MEANINGS are actually different—for the Author (AI) and the Reader (human). This operation of different but simultaneous and symbiotic meaning has simply not been the case before. And the method that chooses the next word is different too. Ultimately, the feeling (ontologically) that these words SIGNIFY the same thing for both AI and humans is illusory. A very convincing illusion, but an illusion nonetheless.

All this is possible due to the arbitrary relationship between a SIGN and its SIGNIFIED. It is not fixed, therefore a SIGN can SIGNIFY different meanings simultaneously. AI language is a closed system. Highly, highly trained, and fast. But closed. And AI would NOT be able to do this “writing” without texts written by us. An obvious, but important point. AI only deals in representation, not “reality” per se because it is entirely image and text based (it is ALL mimesis). In code, x in *code* language represents y in *natural* language (this is a straightforward substitution cipher, like a Caesar cipher). In ChatGPT, x in ChatGPT language represents x in natural language but its SIGNIFIED in ChatGPT is understood only mathematically in its closed AI system (there is literally nothing outside of the text), a signified calculated on the basis of vast corpuses and the characteristics and context of x in those corpuses. But for humans, the SIGNIFIED of x is understood qualitatively, through consensus, social/actual reference points. But because of the CONTEXT for where x appears in the AI produced text (and the context is shared) it does not matter that the SIGNIFIED is not shared as well. Which is why when the context is wrong (the prediction that that was the next best word was incorrect) ChatGPT text does not make sense to a human reader.

The way in which AI produces text, and the language (or discourse) which AI specialists use to talk about that production is really revealing, as they do not actually talk about “words,” they talk about “characters,” “tokens,” “sequences,” and “patterns.” This discourse is revealing because it is not actually seeing a unit of meaning as a WORD, and this assists the operation of two intersecting systems of signification. The AI system, which I am going to call a “schematic signification system,” is closed and based on patterns and is about completing a sequence, and it uses units of text that are not necessarily “words,” as we would recognize them in natural language. This schematic system of signification is entirely dependent on a “source” that is text-based—and that is not the case for what I will call “referent signification,” which is the kind of signification we use in natural language, which is built on layers of referring (references can be drawn from any source, text-based or otherwise). In this system, the signifier (or word) “refers to” an object, a thing, a feeling, and so forth—and the unit is a WORD. And the smallest unit is a LETTER. But the schematic system does not deal in words. What is really interesting though, is that because the schematic system begins with a source that is text-based, that has been produced by natural language speakers and writers, the outputs of the two completely overlap: you can place one on top of the other, they intersect, they are in fact producing the same product. Much like John Searle’s (1980) *Chinese Room* thought experiment: although this is *all* languages, not one chosen as mystifying, not orientaling. But interestingly, although not all source code is written in English, most of it is and nearly all contain some English words. What effect does this have, I wonder?

In our human sense of signification, AI appears to be producing sentences; in AI’s idea of signification, it is completing a sequence: the schematic system produces something that is readable to humans, although it is arrived at in an entirely different way. Words in natural language have a meaning (or signified) that DETERMINES how the word is used. Characters/tokens in AI have a meaning (or signified) that is ONLY how the word has previously been used (as well as its relationship to previous words, or tokens, in the sequence) and the likelihood of the appropriateness¹¹ of where it is *about* to be used. In AI, the token signifies a mind-blowing number of calculations, coding, and testing; hence the use of the word “schematic”—each token represents a compendious SCHEMA of calculations. This “referring” is a very different kind of referent, with its origin in codes and ciphers: it only *refers* to relationships of previous use/in relation to preceding and immediately preceding words. It does not refer to anything *external* to that. It is a closed/isolated system. It really does not know what a dog is. USE is one layer of what determines meaning for humans in natural language, but it is only one layer. Each individual use of a token in AI signifies a prediction that is calculated

through a vast schema of backwards-oriented chunking, batching, integering, sampling, averaging, vectoring, teething, matricing, weighting, aggregating, affinities, summing, and self-attention and so forth of bags and bags of words. Plus, there is a meta-level of writing to the “writing,” which is of course the source code. And it can only concentrate¹² on the next token. It cannot “think” to the end of the sentence and humans do not think or write like that. The units of meaning for humans are words, then sentences. Try writing a sentence only thinking about the next word. . . Impossible right? You are always thinking about the end of the sentence as well.

Is generative AI source code re-creating versions of our cognitive processes (perception, attention, conditioning, and memory)? If it is re-creating our cognitive processes, this re-creation is achieved through writing (source code). Are we anthropomorphising ChatGPT though? For example, we describe it predicting the next word as “hallucinating.” And the technology is only as good as the society that uses it: Humans have been consistently frightened of what computers can achieve and their potential for agency, but are they simply a mirror of what frightens us in ourselves?

Some of My Best Friends Are Books: Dialoguing With the Dead

Thus the Pad provides not only a receptive surface that can be used over and over again, like a slate, but also permanent traces of what has been written, like an ordinary paper pad: it solves the problem of combining the two functions by dividing them between two separate but interrelated component parts or systems.¹³ But this is precisely the way in which, according to the hypothesis which I mentioned just now, our mental apparatus performs its perceptual function. The layer which receives the stimuli [. . .] forms no permanent traces; the foundations of memory come about in other, adjoining, systems. Sigmund Freud, 1924, *A Note Upon the “Mystic Writing-Pad”* (p. 230)

In a video called “The Unborn Author” that I made ten years ago, when I was trying (unsuccessfully) to become a New Generation Thinker, there is the seed of what is happening now when I note that it is not so much that the “author is *dead*” as that the author was never born in the first place, now computers can write novels. *Liquid Times* states that everything is outdated before it is finished, that was published in 2007, the year of the iPhone, and everything has got exponentially faster since then. Yet there is something of Freud’s *Mystic Writing-Pad* too, “a receptive surface that can be used over and over again,” apparently blank and ever-ready, and at the same time full of traces “of what has been written.” And its seeming mimicry of how our human “mental apparatus” performs, stimuli and memory.

But does AI actually have more in common with ghosts, than with people and their ontologies and feelings and souls? Their messy, unaccountable being? Is not being unborn like being dead? Or is it better to scrap this question, and leave the dead in this article as those writers and philosophers who were genuinely once alive and whose work has most likely been cannibalized by this “generative” technology? So, focusing on the *writing* instead, is the palimpsest perhaps a better metaphor? It comes from the idea that, when writing was on parchment, because it was so costly and valuable, it would be scrapped clean and used again. But when it was written on again, traces of what was written on it before would *resurface* through the new writing. Freud expresses a similar idea above and elsewhere, applying it to the unconscious and memory—and the uncanny.¹⁴ With ChatGPT, whatever is written, always has untraceable traces of what has been written before—as with all writing—but in a different more uncomfortable way. Bags and bags of words chewed up, learnt from, regurgitated, without acknowledgment, without consent.

Moreover, despite the fact ChatGPT is presented as question and answer, can it be said to be dialectic? Really? In the philosophical sense, a dialectic is two opposing views in dialogue, but in the Hegelian application it is very specifically, thesis, antithesis, and synthesis. But can ChatGPT genuinely *oppose* a human’s view with its “own” thereby offering an antithesis? I am not sure because the “dialogue” is so dull. It is the genuinely *dead* with whom I would actually like to dialogue about seductive yet faux dialoguing with AI. And these “best friends” are Derrida’s *Specters of Marx* and “Structure, Sign and Play.” Baudrillard’s *Simulacra and Simulation*, and Foucault’s *The Archeology of Knowledge*. Yet my methodology is deliberately as disorganized and “human,” as ChatGPT’s is organized and quantitative. I am drawing too, as the astute reader will have noticed, on Virginia Woolf’s (2004) *A Room of One’s Own* and alluding to many, many well-known contemporary novels, and with all these texts, my allusion to and analysis of them will not necessarily be *coherent* but it is here acknowledged.

Later in his career, Derrida became preoccupied with specters and, in particular, the opening line of *The Communist Manifesto*: “*A spectre is haunting Europe—the spectre of communism*” (1848). So, as speculative fiction, I have exchanged Marx and Marxism, for ChatGPT and AI in this paragraph from Derrida’s (1994) *Specters of Marx*:

The specters of ChatGPT. Why this plural? Would there be more than one of them? Then, if the specter is always animated by a spirit, one wonders who would dare speak of a spirit of ChatGPT, or more serious still, a spirit of AI. Not only in order to predict a future for them today, but to appeal even to their multiplicity, or more serious still, to their heterogeneity. (pp. 1–2)

Does not this multiplicity, this heterogeneity, this appeal to AI, capture some of the disorientation and noise and fragmentation of this forever moment (zeitgeist, spirit of the age)? This ongoing, elongated, *hallucinated* moment, of AI produced text and all its ramifications? The fact of what it produces as text literally freaks us out, frightens us, fascinates us, perhaps replaces us or displaces us. Makes us obsolete. Put simply, it haunts us. I am not simply being pretentious or obtuse with my French philosophy, I really do think it brings something new to light in our already crowded and noisy public sphere (of Habermas).

Try to imagine what Baudrillard would have said as well, he who was so disturbed by the simulacrum of Disneyland and the potential for falsification and distance in the television reportage of the Gulf War: think “fake news,” X, TikTok, and now ChatGPT. It is almost as if the world is too fast now for any type of *slow* analysis, and perhaps deliberately so. And yet, I always felt uneasy of Baudrillard’s four stages of the image: one that reflects a basic reality, one that masks or perverts a basic reality, one that masks the absence of a basic reality, and one that bears no relation to any reality. I used to teach these “stages” using *Toy Story*. So, Woody and Buzz as reflecting “go West” and “space race” versions of the American dream. These dreams are masked or perverted because they are represented by toys, that can speak—and in cartoon form. Therefore, the “basic reality” of the American dream is absent—in the film at least. And then, Buzz’s so-called Spanish mode “bears no relation to any reality” and neither do the animated “bloopers.” But are the last two stages really possible? Isn’t some “basic” reality always implied, however fragile?

To conclude by returning to the question, although it takes the *form* of a dialogue, can it be truly dialectic, or even dialogic, this strange “conversation” with specters and ghosts of previous writing selves? Or does dialogue remain the preserve of humans? Ultimately, it is not Virtual Reality, or robots, or anything “real” that has changed our world. It is computers not thinking, but writing. And where has the Turing Test got us to? Trickery, AI and scams. This discussion has been about writing, that is all that can be analyzed and understood. We want to believe there is more: aliens, conscious machines, and dark forces. But the threat is so much more *meta* than literal, it is not robots, aliens, and monsters, but our written systems of representation that we have developed in such a way they can potentially be used against us. As usual, it is Frankenstein, not the monster, that is the problem; however, this Frankenstein has made a creature that can write, which is not only a proxy for “thinking” (which scares us) but also writing can control (which scares us even more). It is what is within us that we fear, and should fear, and that is part of what is so frightening about AI, we are looking at ourselves, askance, and often seeing what is there as if it were elsewhere. It is *our* writing, no one

else’s, coming back differently, uncannily familiar yet unfamiliar, and we are desperate to understand this as consciousness, that is our need, and our burden. But overall, if it is consciousness, Jim, it’s not as we know it.

Acknowledgments

The author acknowledges that, like any philosophical work, the ideas in this article have been arrived at through dialogue, discussion, and disagreement. In particular, the author wishes to acknowledge the intellectual nutrition of conversations with a wide range of people, which made the author’s position clearer to her. All errors however remain the author’s own.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Eileen Pollard  <https://orcid.org/0000-0003-3634-4393>

Notes

1. Note this quotation about the unknowability of the human soul is from Virginia Woolf’s introspective look at what it means to be ill, and being ill is so *human*: is there perhaps a connection to be made here?
2. However, in an unexpected turn of events, it turns out that this “test” is also, or perhaps, primarily, about gender: “The object of the game for the interrogator is to determine which of the other two is the man and which is the woman,” and whether or not a computer would be good at this game? Obviously, there is not space in this piece to go down such a rabbit hole, other than to say it puts a different slant on this seminal article and perhaps goes some way to demonstrating my point about the “cultures” surrounding mathematics and mathematicians (Turing, 1950).
3. In *Battlestar Galactica*, robots that look exactly like humans try to take over. They were originally more robotic (cruelly referred to as “toasters”) and humans treated them like slaves, so they rose up against them, went to war, went into exile, and then secretly returned having re-created themselves in our image (called cylons). It is a slippery slope, perhaps not saying “please” was the start.
4. This journal article concerns ChatGPT specifically and although many points apply to other Generative AI platforms, such as Google Bard or Bing Chat, for example, the focus here is ChatGPT.
5. Ashwin (2009).
6. In the case of AI, a calculation and a specific arrangement of characters in a token in a specific instance. The calculations for exactly the same token elsewhere would be different. The relationship between a calculation and a token is specific

to that instance, it is not transferrable, meaning the signifier of that token could have an infinite number of possible calculation-signifieds.

7. And here, I use “I” as Virginia Woolf (2004) does in *A Room of One’s Own*, as having no real “being,” and I use the term advisedly.
8. A note on images: I cannot begin to speculate on how AI creates and extends images; however, I imagine, or *hallucinate*, that the image, just like with text, is broken down into tiny tokens and that platforms like DALL-E are similarly completing a sequence initiated by a prompt and trained on previous use.
9. I recognize that VALUE is a form of MEANING, but I am keeping them separate for the sake of clarity.
10. I appreciated that use of the word “number” here is vastly oversimplifying.
11. Although this is a human, qualitative, way of thinking about it, for AI it is a calculation, so it is true or false, it is a hard binary.
12. In writing about AI, it is impossible not to humanize it: the language, even of the source code, “self attention,” “affinity,” forces a positioning of the AI as feeling, anthropomorphising it even as you try to write about it in the abstract. It is similar to the solipsism and confusion poststructuralists experienced when trying to write about writing.
13. The author wishes to acknowledge Margaret Atwood in the composition of this title (negotiating with the dead) and the author’s own best friend, Shafqat Nasir, many of whose other best friends are books.
14. Freud and Strachey (1961).

References

Adams, D. (2020). *The Hitchhiker’s Guide to the Galaxy*. Pan.

Ashwin, P. (2009). *Continuum studies in education: Analysing teaching-learning interactions in higher education: Accounting for structure and agency*. Continuum International Publishing Group.

Bauman, Z. (2007). *Liquid times: Living in an age of uncertainty*. Polity Press.

Freud, S., & Strachey, J. (1961). A note upon the “mystic writing-pad.” In S. Freud (Ed.), *The ego and the Id and other works* (pp. 225–232). Hogarth Press.

McMullen, K. (1983). *Ghost dance*. Cornerstone Media.

Saussure, F. de et al. (1986) *Course in general linguistics*. Open Court.

Searle, J. (1980). Minds, brains and programmes. *The Behavioural and Brain Sciences*, 3, 417–457.

Sinclair, A., & Thomas, D. (1972). *Under milk wood*. Lorrimer.

Swinton, J. (2022). *Alan Turing’s Manchester*. The History Press.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59, 433–460.

Woolf, V. (2002). *On being ill*. Pimlico.

Woolf, V., Pearson, D. & Clays Ltd. (Firm). (2004). *Great Ideas, 18: A room of one’s own*. Penguin.

Author Biography

Eileen Pollard previously worked as a Senior Lecturer in English Literature at the University of Chester and then as a Learner Developer at Manchester Metropolitan University. She has over ten years experience teaching in HE and she joined the University Teaching Academy at Manchester Met in February 2022. She has published on teaching and learning practice, including in the *THE* and *Wonkhe*, as well as in journal articles, blogs and podcasts, and she is an Associate Editor at the *Journal of University Teaching & Learning Practice*.