Fischer, Pascal

'Is it perfume from a dress / that makes me so digress?' : the categorical differences between human experience and AI highlighted by literature

Article

'Is it perfume from a dress / that makes me so digress?' – the categorical differences between human experience and AI highlighted by literature

Sage

Pascal Fischer [ID]

Abstract

Since the term 'Artificial Intelligence' was coined, the respective research field has frequently emulated human mental faculties. Despite diverging viewpoints regarding the feasibility of achieving human-like cognition in machines, the very use of the word *intelligence* for complex computer systems evokes human consciousness. Likewise, there have been attempts to understand the human mind in terms of computers, exemplified by the computational theory of mind. By contrast, my article underscores the categorical difference between the mind and machines. Partly building upon arguments by David Gelernter and Bert Olivier, I focus on literary examples spanning from Shakespeare to T.S. Eliot that accentuate subjective experience, the intricate relationship between body and mind, and the anticipation of death as human characteristics beyond the reach of computational systems.

Introduction

In recent years, debates about 'Artificial Intelligence' (AI) and its relationship to the human mind have reached an unprecedented intensity. Among the reports that caught the attention of a wider public was an article published in the *Washington Post* in June 2022 about the Google engineer Blake Lemoine, who

Department of English and American Studies, University of Bamberg, Bamberg, Germany

Corresponding author:
Pascal Fischer, Department of English and American Studies, University of Bamberg, Bamberg, Germany.
Email: pascal.fischer@uni-bamberg.de

claimed that the company's LaMDA system would be sentient (Tiku 2022). The balanced article also gives ample space to sceptics of this view: The linguist Emily M. Bender, for instance, points to the fact that machines 'mindlessly generate words, but we haven't learned how to stop imagining a mind behind them' (Tiku). The controversy over Lemoine's claims received new currency when OpenAI launched their chatbot ChatGPT in November 2022 (Nezik 2023, 14). Millions of people have since had conversations with this system that appear to be close to human interaction. The illusion of a conscious interlocutor is partly due to the high quality of language created by the 'Large Language Model' but unrealistic expectations about the alleged human character of sophisticated computer programs are already aroused by the terms 'Artificial Intelligence' and 'Intelligent Systems', so widely used by scholars, corporations and popular media today. The language employed about technology could be misleading outside a narrow circle of experts, particularly if there is a vested interest in promoting and exalting certain concepts and expressions.

While inviting readers to ponder the similarities between literature and culture on the one hand and intelligent systems on the other, this special issue of *ISR* takes a critical stance towards the unqualified adoption of postulates promulgated by some proponents of AI. Starting from a discussion of the word *intelligence*, this paper emphasizes the categorical difference between the human mind and AI and claims that literary texts have a unique power to accentuate this difference. A comparison to computers may help us to understand what is unique about the human mind. Most people have at least an intuitive understanding that the mind works differently than a set of microchips. However, the very use of the term *intelligence* and the discussion about AI in the media have contributed to blurring the distinctions.

AI and human intelligence

AI has from the very beginning deliberately aligned itself with human intelligence. True, Alan Turing's 1950 paper 'Computing Machinery and Intelligence', which is frequently cited as a precursor to the research field, hardly strives for the entirety of human cognitive faculties and is mainly interested in the question of whether a machine can be 'linguistically indistinguishable from a human' – the well-known Turing test (Bringsjord and Govindarajulu 2022). However, the document that is regularly mentioned as having kick-started the discipline of AI, namely the 1955 proposal for the 1956 conference at Dartmouth College, New Hampshire, which is 'generally regarded by historians and computer scientists as the "birthplace" of AI' (Kline 2015, 153), clearly strives for all facets of human intelligence. The first paragraph defines the central research goal:

> The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. (McCarthy et al. 1955, 2)

The reference to machines forming 'concepts' clearly invokes consciousness as an integral part of human intelligence, which, in turn, served as the target of the technological endeavour. Recapitulating the early hours of AI, Stephanie Dick explains: 'human intelligence was the central exemplar around which early automation attempts were oriented' (2019).

Recent accounts of AI frequently continue to mention human cognition as the decisive reference point. David Alan Grier, for instance, defines AI as 'a field seeking to express human intelligence through machinery' (2014, 82). Some scholars have no doubts about the attainability of this goal. Kate Jeffery still writes in 2015: 'We'll sidestep discussions about whether machine intelligence can ever approximate human intelligence, because of course it can—we are just meat machines, less complicated or inimitable than we fondly imagine' (366). However, in the course of the vicissitudes of AI

history with its boom phases, setbacks and 'AI winters' (times of reduced funding), an approximation of 'man' has moved to the background.

As a matter of fact, the human-based understanding of intelligence has been rejected by a great number of scholars in the field. Russell and Norvig's highly influential textbook *Artificial Intelligence: A Modern Approach* (AIMA) distinguishes between those that define the goal of AI in terms of human capacities – 'systems that think like humans' – and those that settle for a more abstract rationality (2021, 19–23). The related distinction between Strong AI and Weak AI is also concerned with the similarity between machines and humans. While advocates of the former strive for levelling the distinctions between the human mind and a machine altogether, proponents of the more restrained version are content with a semblance of certain human mental faculties. Although there is hardly agreement on the definitions of these subcategories of AI, John Searle's explanation may serve as a guideline: 'The contrast is that according to Strong AI, the correct simulation really is a mind. According to Weak AI, the correct simulation is a model of the mind' (2009).

Irrespective of the disagreements between these camps, both resort to the term *intelligence* to describe the capabilities of highly complex computer systems. They hence use a term that has been related to the human capacity for understanding since classical antiquity. Intelligence is, after all, derived from the Latin *intellegere*, which means 'to comprehend', 'to understand', and 'to realize' – all activities that are not only native to the human species but remain inextricably bound up with it. Whatever its additional semantic components may be, the core of *intelligence* is human cognition, which cannot be adequately characterized without reference to consciousness.

Intelligence and consciousness

The connection between human intelligence and consciousness is acknowledged by many AI scholars as well. In her popular study, *Artificial You: AI and the Future of Your Mind* (2019), the philosopher Susan Schneider writes: 'In the context of biological life, intelligence and consciousness seem to go hand-in-hand. Sophisticated biological intelligences tend to have complex and nuanced inner experiences' (16). This does not deter her from treating intelligence and consciousness as two distinct concepts as soon as she talks about computers. She speculates, for instance, about a possible development when 'the most intelligent systems of the future may not be conscious' (36), and even envisions the possibility that machines 'exhibit superior intelligence, but […] lack inner mental lives' (16). Even scholars who are much more sceptical about the possibility of computers approaching human mental abilities like David Gelernter, about whom more will be said later, use expressions like 'unconscious intelligence' in this context (2007).

To be sure, the word *intelligence* has been applied to higher species of non-human animals for whom a certain degree of consciousness is assumed to exist. To expand it further to plants or machines involves a metaphorical operation.

AI scholars have argued, with some justification, that we are simply dealing with different definitions of intelligence; the semantic range has shifted. Russell and Norvig, for instance, explain in their AIMA textbook: 'According to what we have called the standard model, AI is concerned mainly with rational action'. They are well aware that the 'rational agents' built in AI are 'intelligent in this sense' only (2021, 52).

Intelligence is a polysemous word. But, as the cognitive linguist George Lakoff explains, 'polysemy is not just a matter of listing meanings disjunctively, as dictionaries do' (1987, 316). The phenomenon has been rightly understood as a cognitive mechanism that relates items in a conceptual system to a prototypical meaning, frequently involving such processes as metaphor and metonymy (Lakoff 1987, 333–34; Gibbs and Matlock 2001: 213–15). In the sentence: 'This is the fruit of my work', the word *fruit* just

means 'product', but it is obvious that it is metaphorically related to the more prototypical meaning of seed-bearing plant. The name of the clothing company 'Fruit of the Loom' primarily refers to the 'product' of the loom, the garment. But the prototypical sense of *fruit* is evoked as well. The company even plays with the different levels of meaning by using grapes, currants and an apple as its logo.

A related mechanism is at work when we talk about Artificial Intelligence or Intelligent Systems. We may define the word in regard to 'learning', 'reasoning', 'problem solving' etc., and analytically treat intelligence and consciousness as two distinct concepts, but for many people, the word intelligence will suggest the idea of the conscious human mind at the same time. In terms of scientific methodology, this is not unproblematic. To quote the science historian and Darwin scholar Robert M. Young: 'One of the cardinal rules of modern science is to avoid explaining things in terms which draw on human intentions' (1993, 379). While intentionality may not be central to the word Artificial Intelligence, there is clearly an anthropomorphic element. The ambiguity about the exact meaning and possible connotations of *intelligence* in machines may be sought for by some. Kate Crawford, one of the severest critics of some practices in the field, comments: 'The nomenclature of AI is often embraced during funding application season, when venture capitalists come bearing checkbooks, or when researchers are seeking press attention for a new scientific result' (2021, 9).

Understanding humans in terms of machines

The proposition that machines resemble human intelligence finds a counterpart in the idea that human intelligence follows a machine-like logic. We do not only understand machines in terms of humans but also humans in terms of machines. The computational theory of mind claims that the mind functions like a computer or a software programme, even though some of its theorists avoid references to the computer and prefer to speak about 'computational structure' (Chalmers 2011, 326). Nonetheless, analogies between minds and computers abound in AI as well as in the philosophy of mind. For Ellen Ullman '[t]his circular idea – the mind is like a computer; study the computer to learn about the mind – has infected decades of thinking in the computer and cognitive sciences' (2017, 136–37). She quotes passages from Daniel Dennett's *Consciousness Explained* (1991)*, 'which is suffused with conflations between human sentience and computers' (Ullman 2017, 137), for instance when Dennett talks about 'the brain's computer' (Dennett 1991, 218).

Many scholars following 'computationalism' take for granted that the mind relates to the brain as software relates to computers. Since, in this view, the wet matter of the brain and microchips perform the same computational operations, the approach has been referred to as 'substrate independent' (Schneider 2019, 23–24). An extreme example of this would be Raymond Kurzweil's idea, submitted in his celebrated book *The Singularity is Near: When Humans Transcend Biology* (2005), that the content of the brain, including a 'person's entire personality, memory, skills, and history', could be uploaded to a 'powerful computational substrate' (199). In Kurzweil's estimation, such uploading could easily be possible by the late 2030s (200).

The computational theory of the mind has come under fire from different quarters. Even in 1972*, the philosopher Hubert L. Dreyfus dissected what he considered the naïve assumption of such an understanding of the human mind in his book *What Computers Can't Do* (e.g. Dreyfus 1972, 156). John Searle has been among the fiercest critics of this view, e.g. in his book *The Mystery of Consciousness* (1997). Jerry Fodor's *The Mind Doesn't Work That Way* (2000) does not entirely reject the computational theory of mind, to which he has himself contributed substantially, but claims that it only constitutes a part of a much more complex and poorly understood reality.

Related to the controversy over the computational structure of the brain is the debate on human consciousness. Whereas David Chalmers holds that a set of mechanisms that are associated with consciousness – 'how the brain discriminates stimuli, integrates information and produces verbal reports' (1995a, 82) – can easily be described by the prevalent theories of cognitive science, he complains that many scholars have evaded the 'hard problem':

> The hard problem, in contrast, is the question of how physical processes in the brain give rise to subjective experience. This puzzle involves the inner aspect of thought and perception: the way things feel for the subject. When we see, for example, we experience visual sensations, such as that of vivid blue. Or think of the ineffable sound of a distant oboe, the agony of an intense pain, the sparkle of happiness or the meditative quality of a moment lost in thought. All are part of what I am calling consciousness. It is these phenomena that pose the real mystery of the mind. (1995a, 81)

Harkening back to Thomas Nagel's 'What is it like to be a bat?' (1974), 'one of the most famous questions ever asked in consciousness studies' (Blackmore 2010, 23), Chalmers captures the subjective aspect of experience with the formulation 'something it is like to be in that state' (1995b, 201). These private qualities of experience have also been called 'qualia' in philosophical discourse (Blackmore 2010, 26). While some scholars claim that it would someday be possible to find out how the brain creates qualia, others, Colin McGinn among them, hold that the human mind is 'cognitively closed' to resolve the problem of how the brain produces consciousness (1989, 352, *passim*). Naturally, McGinn is highly critical of the idea that an increase in processing power in computers 'will take us one jot closer to building a genuinely mental machine' (1999). This kind of criticism has recently been updated by Erik Larson, who writes in *The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do* (2021): 'AI culture has simplified ideas about people, while expanding ideas about technology' (3).

AI scepticism: David Gelernter

Among those who are sceptical of the possibility of AI ever attaining human-like consciousness, David Gelernter is an original voice. The professor of Computer Science at Yale, who has frequently been referred to as a 'polymath' (e.g. Friedersdorf 2017), has published on the relationship between the mind and machines since the 1990s. His debate with Ray Kurzweil at MIT in 2006 threw the disagreements between the techno-optimist camp and the sceptics into sharp relief (MIT World, 2006). Building on this verbal exchange, Gelernter published a piece in the *MIT Technology Review* under the title 'Artificial Intelligence is Lost in the Woods' (2007), which gives more reasons for his conviction that the development of a conscious computer is 'highly unlikely'.

He particularly draws attention to the fact that human thinking is characterized by a 'cognitive continuum' between focused and controlled states of mind on the one hand and free association or even hallucination on the other. His evolving thoughts on this continuum were later summarized in his 2016 book *The Tides of Mind: Uncovering the Spectrum of Consciousness*. Gelernter argues that scholars in AI have only singled out one capacity of human intelligence, namely high-focus rational thinking, and declared it the essence of the mind, to the detriment of all those characteristics that are further down the cognitive spectrum. But it is hardly AI alone that is to blame for the neglect of interest in 'the middle- and lower-spectrum phenomena', as most of the modern philosophical tradition is biased against studying the less controlled aspects of consciousness (2016, 146). Gelernter thus demands: 'To understand the mind, we must go over the ground *beyond* logic as carefully as we study logic and reasoning' (2016, 147). Fathoming the deeper layers of the mind is particularly revealing in regard to human creativity:

> We *can* understand creativity, or a great deal about it. Creativity has much to do with the dynamics of the spectrum and two of the spectrum's major transitions: the gradual emergence of emotion, and the unconscious mind's gradual taking over from consciousness, as we move down-spectrum. (149)

Creativity flows from our physical interaction with the world. Sometimes it is outside stimuli that induce certain ideas. The computational theory of the mind, and particularly the computer analogy, does not help to explain these processes.

Gelernter puts a somewhat different emphasis in an article for *Commentary* magazine with the title 'The Closing of the Scientific Mind' (2014). Sure enough, he also censures the tendency, particularly among his colleagues in AI, to downplay the importance of subjectivity for understanding the human mind. He criticizes the master analogy between mind and software as well as brain and computer on several counts, among them that you can transfer a programme from one computer to another, but you cannot transfer a mind from one brain to another. In contrast to a computer, only one 'programme' can run on a human brain. Whereas computers are 'transparent', i.e. you can read the state of the programme at any time, minds are 'opaque', i.e. you cannot know what someone is thinking unless the person tells you (21). The deeper problem he identifies about computationalists is that they treat 'the mind as if its purpose were merely to act and not to *be*. But the mind is for doing *and* being' (21). Emotions cannot be broken down into action, but should be understood as 'states of being' (22) – they are not information that could be processed. Consciousness does not only mean that we are aware of ourselves, but also that we *experience* our being. Gelernter cites the American philosopher Thomas Nagel's *Mind and Cosmos* (2012) as a major authority to support his conviction that science has not been able to sufficiently explain consciousness.

What Gelernter particularly insists on in his *Commentary* piece is that the mind is not only embodied by the brain but by the brain *and* the body, which are closely intertwined. Emotions are partly the result of bodily processes and, in turn, can have effects on the body. Mental states like emotions are frequently felt in the body. The physical and the mental resonate with each other.

These may not be revolutionary insights, but they bear emphasizing in an area that glosses over these fundamentals. Granted, Kurzweil does not neglect the body entirely, but he talks about it as if it were a machine with removable parts, for instance when he concludes that 'we'll be able to rapidly alter our physical manifestation at will' (2005, 310). There is very little awareness of body–mind interaction. What is more, the notion, frequently associated with Descartes, that a person's 'self' is separate and distinct from the body, is widespread in today's discourse, in spite of recent research in the fields of 'embodied cognition' and '4E cognition' (see Shapiro 2019; Newen, De Bruin and Gallagher 2018). The idea, proffered by transgender activists, that a woman may be trapped in the body of a man or vice versa betrays this 'Cartesian dualism' (Wilton 2000).

That Gelernter's approach has met with limited response in debates about AI may partly be due to his unorthodox ideas in other scientific areas and by his right-wing political stance: He has given his imprimatur to aspects of intelligent design, is a sceptic of anthropogenic global warming and has spoken his mind against liberal academia.

AI scepticism: Bert Olivier

The South African Philosopher Bert Olivier is among the few scholars who have integrated Gelernter's publications on the mind into their reflections on AI. In a number of papers, Olivier welcomes Gelernter's contributions to the AI debate (see also Olivier 2018), but argues that they do not go far enough in their criticism of 'computationalism'. Olivier identifies as a shortcoming of Gelernter's approach that the latter does not include the 'capacities of making ethical (and aesthetic) judgments'

(2017, 7) in his view of the incompatibility between AI and human intelligence. The philosopher insists that moral attributes must be counted among those characteristics that distinguish the cognitive operations of humans from those of machines.

Olivier also takes the embodiment of the human mind into account, but is less interested in mind–body interaction regarding immediate emotions. He rather connects our physical existence to aspects of ethical action as well as desire (2017, 12). Drawing on Heidegger's philosophy, Olivier explains that the distinctive mode of being human (Dasein) is characterized by their capacity for 'care' – care for oneself and for others. 'Care' and the related concept of 'concern' usually involve an orientation to the future. All sorts of human 'projects', including repairing a chair, presenting ourselves on social media and designing houses, are geared towards a future goal. Citing Heidegger, Olivier insists that humans are constantly 'ahead-of-themselves' in their endeavours (2017, 14).

Most importantly, our consciousness is inextricably linked to our anticipation of death – whatever the specific attitude towards it may look like. It is our 'Sein zum Tode', or 'being-towards-death', which determines our whole lives (15). This creates a fundamental difference to machines, which do not die, because they do not have an organic body. In summary, Olivier states:

> It is supremely doubtful whether any instance of AI is capable of this 'anticipation', let alone existential anxiety in the face of certain, though unspecified, death, because if it ceases to exist, for whatever reason, its cessation cannot possibly be synonymous with the multi-faceted death of a time-bound, inescapably mortal human being […]. (16)

The existential relationship each person adopts towards his or her finite life stories should indeed be seen as a quintessential characteristic of the human mind.

Literary perspectives on the mind

Recently, science-fiction scholars have shown how important the comparison between human minds and computers has been in that genre. Stephen Cave, Kanta Dihal, and Sarah Dillon's *AI Narratives: A History of Imaginative Thinking about Intelligent Machines* (2020) illustrates the enormous contribution fictional texts have made to the idea of machine intelligence from antiquity to the present. While many novels and short stories play with the convergence between humans and computers, others elaborate on the difference. Cave's own article (2020) in the collection zooms in on novels that have questioned certain assumptions about the possibility of uploading a mind to a computer chip, among them Greg Egan's *Permutation City* (1994), Robert Sawyer's *Mindscan* (2005), and Cory Doctorow's *Walkaway* (2017).

There is no doubt that these narratives help to crystallize the peculiarities of the human mind. At the same time, we do not need AI narratives to accentuate what distinguishes the human mind from computers: Other kinds of literature may serve the same purpose. Gelernter's insistence on the importance of human subjectivity and the mind–body interaction, as well as Olivier's Heideggerian conception of human transience, should now stand at the centre of an analysis of the potential of literary works to elucidate the categorical difference between human creativity and AI.

It is particularly Gelernter's approach that lends itself to a literary perspective, as he illustrates his ideas by reference to literary texts. John Keats, Jane Austen, Henry James, Leo Tolstoy, William Wordsworth and others are quoted frequently in 'The Closing of the Scientific Mind' and *The Tides of Mind*. For Gelernter 'these "subjective humanists" can tell us, far more accurately than any scientist, what things are like inside the sealed room of the mind' (2014, 22). He, for instance, cites Keats' 'Ode to a Nightingale' (1819) to illustrate that mental life does not only consist of computational operations

to achieve results, but also of moments in which people experience affections for their own sake. After quoting 'I cannot see what flowers are at my feet, / Nor what soft incense hangs upon the boughs … Darkling I listen…' Gelernter comments caustically: 'That was drafted by the computer known as John Keats' (2014, 21).

In the *MIT Technology Review* article, Gelernter uses a poem to exemplify how human creativity is employed in the invention or discovery of analogies. Wondering about the *tertium comparationis* in the famous question 'Shall I compare thee to a summer's day?' at the beginning of Shakespeare's Sonnet 18, Gelernter comes to the conclusion that the person addressed and the summer's day conjure the same feelings: 'The lady and the summer's day made the poet *feel* the same sort of way'. A computer, which has not access to the feeling of a summer's day, must of necessity stand outside this experience. What the example is furthermore supposed to show is that we go through an abundance of emotions for which no simple terms exist; 'happy' or 'elated' are much too unspecific. Cognitive operations like metaphors and similes can help to evoke these feelings nevertheless.

Literary scholars may find fault with the fact that Gelernter refers to the lyric thou as a 'lady'. In the context of Shakespeare's sonnet sequence, the addressee should, of course, be identified as a young man. Sure enough, the literary terminology of the computer specialist is also far from unimpeachable, not least when he conflates poet and speaker. One may add that Gelernter does not contextualize the analogy-formation of the verses in the Petrarchan tradition the poem engages with. But this is no reason to dismiss his invitation to productively juxtapose the multi-layered character of mental operations in literary works with the reduced complexity of computer 'intelligence'.

A more serious objection may be that AI chatbots that are based on a 'Large Language Model' are also capable of producing literary texts. After reconstructing the long tradition of computer-generated poetry, Martin Paul Eve talks about GPT-3, the system underlying the current free version of ChatGPT, and the author concedes that 'its outputs are virtually indistinguishable from high-quality human-authored text' (2022, 57). At the same time, he betrays some doubts as to the literary aspirations of these texts.

A decidedly sceptical position – to say the least – is taken by Angus Fletcher, a trained neuroscientist and literary scholar, whose book *Wonderworks: Literary Invention and the Science of Stories* (2021) has reminded readers of the power of literature to expand the capacities and abilities of the human brain. In no uncertain terms, Fletcher dismisses the ability of computers to develop stories. Not only does he claim that '*no* computer AI has *ever* learned to write a story' (2022, 126), but he also explains why this will not be possible based on the current technology: Apart from being a writer, an author must also be a reader of his own stories in order to improve his or her storytelling. Computers, however, cannot understand the texts they produce and therefore cannot improve (126). This is very much in line with what Gelernter recently said about ChatGPT texts in a *Washington Post* article: 'Its lack of consciousness, and its consequent lack of intuition or feeling, limits Chat's ability to judge the quality of its own work'. The result is an 'officious' and bureaucratic-sounding language, leading to a 'deluge of bad prose' (2023).

But the question of whether chatbots will be able to produce sophisticated literature in the foreseeable future is not particularly relevant to the claim of this paper that literary texts have a distinctive ability to underscore the difference between the human mind and AI. The language stitched together by 'Large Language Models' like GPT-3 is, after all, based on utterances that have been made by humans. If we accept the premise 'the author is dead' from a literary studies perspective, we must also abstract from the author in our concentration on the literary texts in those cases in which AI has produced these texts. Even these may be able to accentuate the difference between the human mind and AI – paradoxical as it may sound.

There can be no doubt that it is the strength of many literary texts to capture subjective experience, accentuate the nexus between the body and the mind, and clarify processes of creativity. It should furthermore be emphasized that the transience of our existence has been one of the major concerns of literature since antiquity. Admittedly, there are realist writers that are more interested in the physical reality of a city or the social relations between the classes, but in some literary periods, the inner workings of the human mind take centre stage. In the following, a few examples will be given to illustrate how literary texts can be used to highlight some peculiarities of the human mind that could not be replicated by an artificial system.

We see a great interest in states of consciousness and complex emotions already in the early modern period. When Shakespeare's Hamlet in his first soliloquy wishes to have his 'solid flesh…melt / Thaw, and resolve itself into a dew' (I.ii.129–30), we can grasp that his suicidal desperation is felt by him physically. The image of the 'unweeded garden' (I.ii.135), Hamlet uses a few verses further down to describe the world around him is indicative of his trepidation, but may also be read as a metaphor for his unorderly state of mind.

Critics have identified 'interiority' as a central characteristic of the Metaphysical Poetry of the seventeenth century (e.g. David Reid 2000, 4–7). In George Herbert's poem 'Denial' (1633), the speaker suffers profound spiritual anguish on account of his alleged distance to God. In keeping with the ingenuity of this 'school' of poetry, Herbert uses a number of creative metaphors and similes to convey the speaker's distress:

> My bent thoughts, like a brittle bow,
> Did fly asunder:
> Each took his way; some would to pleasures go,
> Some to the wars and thunder
> Of alarms. (2001 [1633], 134)

In stark contrast to the idea of rational intentionality, which dominates concepts of the mind in AI, thought here appears to be unmanageable and unruly. The speaker analogizes his strained mind to a flexed bow, which conjures the idea of an archer but may also refer to a violinist. The conceit of the string instrument in fact runs through much of the poem, for instance, when the speaker calls his soul 'untuned, unstrung' (135). Following the religious convictions of the Anglican priest George Herbert, the human soul is in need of divine intervention to function harmoniously. In the poem, this inspiration is the prerequisite for the creative activity of writing poetry. Whereas the first five stanzas end with an unrhymed verse and thus 'deny' closure, God finally grants his grace, leading to a rhymed couplet at the end of the last stanza. The rhyme structure does not only symbolize mental strain and finally – with the 'mending' of the rhyme – its reduction on the part of speaker, but also reproduces the transition from tension to relief for the reader, who correspondingly has a cognitive as well as an affective experience. The poem illuminates various facets of human consciousness and conveys them comprehensibly to the recipient. While traditional interpretations of these verses elaborate on the religious aspects, particularly man's dependency on God, using artificial forms of 'intelligence' as a standard of comparison can spotlight the complexities of the human mind.

Metaphysical Poetry also amply illustrates the mechanism of analogy discovery which is at the heart of creativity. Gelernter emphasizes that humans do not just observe their thoughts, but also feel them. Reading about these processes in *The Tides of Mind* (2016, 115–122), one is reminded of T.S. Eliot's 1921 article on the Metaphysical Poets, in which he comments on the ability of John Donne, the most prominent representative of the group, to connect rationality with emotion: 'A thought to Donne was an experience; it modified his sensibility. When a poet's mind is perfectly equipped for its work, it is

constantly amalgamating disparate experience…' (Eliot 1932 [1921], 273). True, Eliot sets out to explain what was so remarkable about the Metaphysicals and also to depreciate later generations of poets, who were allegedly unable to 'feel their thought as immediately as the odour of a rose' (273). He nevertheless describes a human ability in general: In contrast to some AI theories, thought is not just a computational operation to solve problems but is experienced as part of consciousness.

Complex psychological conditions are of paramount importance to the period of Romanticism. Even Edmund Burke's treatise on aesthetic theory *A Philosophical Enquiry into the Origin of Our Ideas of the Sublime and Beautiful* (1757), which helped pave the way to Romanticism, enlarges upon highly intricate and allegedly contradictory emotions that may simultaneously affect people's consciousness. For Richard Bourke, it is 'these mixed states that absorb most of Burke's attention in the *Enquiry*' (2015, 127).

In English Romanticism proper, many poems record moments of heightened awareness. In his little book on 'Innehalten', the German word for psychological states of pause, reflection and introspection, Christoph Bode discusses several Romantic poems that illustrate the fascination with the depths of consciousness in that period. It does not come as a surprise that John Keats's poems feature prominently here, among them the 'Ode to a Nightingale' (1819) cited by David Gelernter as evidence for the great distance between human perceptions and AI. Bode shows how the speaker comes closer to the bird – not by any change of his physical location but by intensifying his consciousness, leading to a purely mental 'rapture' (Bode 2017, 26–27). Equally relevant are Bode's comments on the Romantics' attempts to approximate those mental conditions that are not representable by language. Keats's sonnet 'On First Looking into Chapman's Homer' (1816) is very much aware of the fact that its speaker's amazement cannot be adequately expressed (22). The incommensurability of language to capture consciousness entirely becomes a central subject of a poem that has to end in silence. That the reader is nevertheless afforded glimpses of the speaker's inner life is due to a number of images that are supposed to substitute the experience and could best be regarded as external correlatives of a mental state (23).

Inspired by Sigmund Freud's insights, Modernist literature of the first decades of the twentieth century famously centres on narrative and poetical means to represent the workings of the human mind. James Joyce's novel *Ulysses* (1920) provides a variety of examples of the technique of *stream of consciousness*, most strikingly in the final episode known as Molly Bloom's soliloquy. Reading her rambling thoughts, one is reminded how little the human mind follows the idea of coherent logic undergirding the computational theory of mind. The passage, highly charged with physical desire and interlaced with smells and tastes, furthermore emphasizes the close connection between body and mind so essential for human consciousness.

T.S. Eliot's long poem 'The Love Song of J. Alfred Prufrock' (1915) constitutes a particularly pertinent case to illustrate Gelernter's and Olivier's arguments concerning the difference between the human mind and computers. The poem, which has been called an internal (or interior) dramatic monologue, presents a direct rendering of impressions and thoughts that come to the speaker's / Prufrock's mind in the course of a visit to fashionable people, or at least thoughts he has while imagining himself making that visit. Rather than being interested in structured reflection, Eliot is anxious to convey Prufrock's tense emotions and frayed nerves. With the first image provided by Prufrock, it becomes clear that he projects his psyche on his surroundings:

LET us go then, you and I
When the evening is spread out against the sky
Like a patient etherised upon a table; (Eliot 1990 [1915], 3)

One may elaborate on the contrast between the invitation to a journey and the promises of a night out here and the sobering image of the etherized patient there, but what is important is that the image is supposed to conjure up a particular emotion, an emotion that cannot be verbalized, quantified or digitized. It is not an action, but a 'state of being'. In his essay on Hamlet, Eliot later referred to such objects or situations in art as 'objective correlatives' of emotion (1932 [1919], 145). We are presented with an image, an event or a situation and the emotion is evoked in us as well. At least in theory. It is, of course, impossible to measure whether the objective correlatives always achieve their desired effects.

However, the poem also draws on several other strategies to provide access to Prufrock's states of consciousness – his qualia – particularly his self-doubts and insecurities towards women or a particular woman, but also his desire for the female. In a few brush strokes, Prufrock is introduced as a character with a personal narrative, whose former frustrations impinge upon his current state of mind. The phrase 'I have known them all already' (e.g. 5) is repeated in several variations in the course of the poem, thus reminding us that we are to a large extent the product of our past experiences and memories. At the same time, Prufrock is extremely self-conscious about his bodily decay and mortality. He imagines, for instance, women whispering behind his back about his thin hair and 'how his arms and legs are thin' (5). Death casts his long shadow over the poem's speaker:

> I have seen the moment of my greatness flicker,
> And I have seen the eternal Footman hold my coat,
> and snicker, (6)

Hammering home the message 'I grow old…I grow old' (8), the poem underscores that the human mind, whose operations we become privy to, is defined by the expectation of death. The poem exposes how strongly our constant awareness of human transitoriness affects our consciousness and may thus help to illustrate Olivier's insistence on the incomparability of the human mind with computer intelligence. Dominic Griffith has shown that 'The Love Song of J. Alfred Prufrock' lends itself to a Heideggerian reading. Concluding his analysis, Griffith writes: "The existentials that Heidegger discusses, which ontologically constitute Dasein, become the means of diagnosing the Prufrockian dilemma: how is it that one should *be*?' (2009, 123). While Prufrock ask himself frequently how to behave correctly in a particular situation ('do I dare to eat a peach?' [8]), the poem is indeed equally interested in fundamental questions of being – in contrast to doing – as they were discussed above.

It is with this image of old age, impending death, continuing longing and desire and the frustration of recurrent rejection that the poem ends – of course, again with an objective correlative that sets out to capture not a reflection but a sentiment:

> We have lingered in the chambers of the sea
> By sea-girls wreathed with seaweed red and brown
> Till human voices wake us, and we drown. (8)

The whole stream of consciousness that is the poem illustrates how the mind works – not in a directed rational manner, but rambling from one thing to another, from one emotion to another. We may even argue that the whole narrative of the poem is just a reverie, a dream-like state: The speaker mentions being woken up in the end. We may be dealing with one of the tides of mind described by Gelernter that are particularly conducive to human creativity.

Most crucially, feelings are triggered by the senses:

> Is it perfume from a dress
> That makes me so digress? (6)

These verses are spoken by Prufrock after waxing lyrical about female arms 'braceleted and white and bare / (But in the lamplight, downed with light brown hair!)' (5). We can easily imagine that this erotic vision is connected to the scent of perfume. In recent years, considerable research has been done on the 'odor-evoked autobiographical memory': an olfactory cue causes a spontaneous recollection of a past event (Hackländer, Janssen and Bermeitinger 2019). It is not a coincidence that this effect is called the 'Proust phenomenon', referencing a passage in the first volume of Marcel Proust's *À la recherche du temps perdu* (1913). It is simply the strength of literature to offer insights into the mechanisms of sensual stimuli, memories and emotions. This is how the mind works, but also how human creativity and intelligence work: None of this is unrelated to the body, none of it is unrelated to emotion, and none of it is unrelated to individual experience. And this is a fundamental difference to so-called computer intelligence.

Conclusion

Recent decades have seen a number of 'turns' and new perspectives in the interpretation of literary texts, among them the spatial turn, the ethical turn, the affective turn and ecocriticism. All of them have their virtues in 'turning' our attention to elements that had been insufficiently illuminated before. Certainly, comparing literary representations of the mind with AI cannot aspire to form a new 'school' in literary studies – after all, the discipline has always had an interest in human consciousness in literature – but it may help to highlight what is unique about the human mind and its operations. An analysis of the various literary techniques to capture experience may also sometimes arrive at the incommensurability of language to display the multi-layered complexity of the human mind comprehensively. But even then, the vast distance between conscious subjectivity and AI will be thrown into sharp relief.

ORCID iD

Pascal Fischer https://orcid.org/0009-0009-5800-8906

References

Blackmore, Susan. 2010. *Consciousness: An Introduction*. 2nd ed. London: Hodder Education.

Bode, Christoph. 2017. *Vom Innehalten. Anhand einiger Gedichte der englischen Romantik*. Gera: Garamond.

Bourke, Richard. 2015. *Empire & Revolution. The Political Life of Edmund Burke*. Princeton: Princeton University Press.

Bringsjord, Selmer, and Naveen Sundar Govindarajulu. 2022. "Artificial Intelligence." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford: Stanford University. https://plato.stanford.edu/archives/fall2022/entries/artificial-intelligence/.

Cave, Stephen. 2020. "AI: Artificial Immortality and Narratives of Mind Uploading." In *AI Narratives: A History of Imaginative Thinking about Intelligent Machines*, edited by Stephen Cave, Kanta Dihal, and Sarah Dillon, 309–32. Oxford: Oxford University Press.

Cave, Stephen, Kanta Dihal, and Sarah Dillon eds. 2020. *AI Narratives: A History of Imaginative Thinking About Intelligent Machines*. Oxford: Oxford University Press.

Chalmers, David J. 1995a. "The Puzzle of Conscious Experience." *Scientific American* 273 (6): 80–6. https://doi.org/10.1038/scientificamerican1295-80.

Chalmers, David J. 1995b. "Facing up to the Problem of Consciousness." *Journal of Consciousness Studies* 2 (3): 200–19. https://doi.org/10.1093/acprof:oso/9780195311105.003.0001.

Chalmers, David J. 2011. "A Computational Foundation for the Study of Cognition." *Journal of Cognitive Science* 12 (4): 325–59. https://doi.org/10.17791/jcs.2011.12.4.325.

Crawford, Kate. 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven: Yale University Press.

Dennett, Daniel C. 1991. *Consciousness Explained*. Boston: Little, Brown and Co.

Dick, Stephanie. 2019. "Artificial Intelligence." *Harvard Data Science Review* 1 (1), https://doi.org/10.1162/99608f92.92fe150c.

Dreyfus, Hubert. 1972. *What Computers Can't Do: A Critique of Artificial Reason*. New York: Harper & Row.

Eliot, T. S. 1932 [1919]. ""Hamlet." In *Selected Essays. 1917-1932*, 141–6. London: Faber and Faber.

Eliot, T. S. 1932 [1921]. ""The Metaphysical Poets." In *Selected Essays. 1917-1932*, 267–77. London: Faber and Faber.

Eliot, T. S. 1990 [1915]. ""The Love Song of J. Alfred Prufrock." In *The Waste Land and Other Poems*, 3–8. London: Faber and Faber.

Eve, Martin Paul. 2022. *The Digital Humanities and Literary Studies*. Oxford: Oxford University Press.

Fletcher, Angus. 2021. *Wonderworks: Literary Invention and the Science of Stories*. London: Swift Press.

Fletcher, Angus. 2022. "Why Computer AI Will Never do What We Imagine It Can." *Narrative* 30 (1): 114–37. https://doi.org/10.1353/nar.2022.0006.

Fodor, Jerry. 2000. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. London: MIT Press.

Friedersdorf, Conor. 2017. "'There's Enough Time to Change Everything': The Polymath and Computer Scientist David Gelernter's Wide-Ranging Ideas about American Life." *The Atlantic*, February 23. https://www.theatlantic.com/politics/archive/2017/02/theres-enough-time-to-change-everything/517209/.

Gelernter, David. 2007. "Artificial Intelligence Is Lost in the Woods." MIT Technology Review. https://www.technologyreview.com/2007/07/01/36908/artificial-intelligence-is-lost-in-the-woods/.

Gelernter, David. 2014. "The Closing of the Scientific Mind." *Commentary* 2014: 17–25.

Gelernter, David. 2016. *The Tides of Mind: Uncovering the Spectrum of Consciousness*. New York: Liveright.

Gelernter, David. 2023. "The Wizards of AI Can't Give it a Brain, or Heart, or Consciousness." *The Washington Post*, April 24.

Gibbs, Raymond W., and Teenie Matlock. 2001. "Psycholinguistic Perspectives on Polysemy." In *Polysemy in Cognitive Linguistics. Selected Papers from the International Cognitive Linguistics Conference (July 14-19 1997, Amsterdam)*, edited by Hubert Cuyckens, 213–39. Amsterdam: John Benjamins.

Grier, David Alan. 2014. "Artificial Intelligence." In *The Oxford Encyclopedia of the History of American Science, Medicine, and Technology*, edited by Hugh Richard Slotten, 82–3. Oxford: Oxford University Press.

Griffiths, Dominic. 2009. "Daring to Disturb the Universe: Heidegger's Authenticity and The Love Song of J. Alfred Prufock." *Literator: Journal of Literary Criticism, Comparative Linguistics and Literary Studies* 30 (2): 107–26. https://doi.org/10.4102/lit.v30i2.81.

Hackländer, Ryan. P.M., Steve M.J. Janssen, and Christina Bermeitinger. 2019. "An in-Depth Review of the Methods, Findings, and Theories Associated with Odor-evoked Autobiographical Memory." *Psychonomic Bulletin & Review* 26 (2): 401–29. https://doi.org/10.3758/s13423-018-1545-3.

Herbert, George. 2001 [1633]. "'Denial' from *The Temple*." In *The Broadview Anthology of Seventeenth-Century Verse & Prose*. Vol. 1: Verse, edited by Alan Rudrum, Joseph Black, and Holly Faith Nelson, 134–5. Peterborough: Broadview Press.

Jeffery, Kate. 2015. "In Our Image." In *What to Think About Machines That Think: Today's Leading Thinkers on the Age of Machine Intelligence*, edited by John Brockman, 366–9. New York: Harper Perennial.

Kline, Ronald R. 2015. *The Cybernetics Moment: Or Why We Call our Age the Information Age*. Baltimore: John Hopkins University Press.

Kurzweil, Ray. 2005. *The Singularity is Near: When Humans Transcend Biology*. New York: Penguin.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago and London: University of Chicago Press.

Larson, Erik J. 2021. *The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do*. London: Belknap Press.

McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon. 1955. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf.

McGinn, Colin. 1989. "Can We Solve the Mind-Body Problem?" *Mind* 98 (391): 349–66. https://doi.org/10.1093/mind/xcviii.391.349.

McGinn, Colin. 1999. "Review Article "Hello, HAL: Three Books Examine the Future of Artificial Intelligence and Find the Human Brain is in Trouble."." *The New York Times*, January 3. https://archive.nytimes.com/www.nytimes.com/books/99/01/03/reviews/990103.03mcginnt.html.

MIT World. 2006. "Creativity: The Mind, Machines, and Mathematics: Public Debate" November 30. https://techtv.mit.edu/videos/16199-creativity-the-mind-machines-and-mathematics-public-debate.

Nagel, Thomas. 1974. "What is it Like to be a Bat?" *Philosophical Review* 83 (4): 435–50. https://doi.org/10.2307/2183914.

Nagel, Thomas. 2012. *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False*. Oxford: Oxford University Press.

Newen, Albert, Leon De Bruin, and Shaun Gallagher. Eds. 2018. *The Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press.

Nezik, Ann-Kathrin. 2023. "Hast du ein Bewusstsein? Ich denke schon, antwortet der Rechner." *Die Zeit*: 13–5.

Olivier, Bert. 2017. "Artificial Intelligence (AI) and Being Human: What is the Difference?" *Acta Academica* 49 (1): 2–21. https://doi.org/10.18820/24150479/aa49i1.1.

Olivier, Bert. 2018. "Explorations of the 'Transhuman' Dimension of Artificial Intelligence." In *Artificial Intelligence: Advances in Research and Applications*, edited by Luis Rabelo, Sayli Bhide, and Edgar Gutierre, 321–37. New York: Nova Science.

Reid, David. 2000. *The Metaphysical Poets*. Harlow: Longman.

Russell, Stuart, and Peter Norvig. 2021. *Artificial Intelligence: A Modern Approach*. 4th ed. Harlow: Pearson.

Schneider, Susan. 2019. *Artificial You: AI and the Future of Your Mind*. Princeton: Princeton University Press.

Searle, John. 1997. *The Mystery of Consciousness*. London: Granta.

Searle, John. 2009. "Chinese Room Argument." *Scholarpedia* 4 (8): 3100. https://doi.org/10.4249/scholarpedia.3100.

Shapiro, Lawrence. 2019. *Embodied Cognition*. 2nd ed. London: Routledge.

Tiku, Nitasha. 2022. "The Google engineer who thinks the company's AI has come to life." The Washington Post, June 11. https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/.

Ullman, Ellen. 2017. *Life in Code. A Personal History of Technology*. New York: Farrar, Straus and Giroux.

Wilton, Tamsin. 2000. "Out/Performing Our Selves: Sex, Gender and Cartesian Dualism." *Sexualities* 3 (2): 237–54. https://doi.org/10.1177/136346000003002008.

Young, Robert M. 1993. "Darwin's Metaphor and the Philosophy of Science." *Science as Culture* 3 (3): 375–403. https://doi.org/10.1080/09505439309526356.