Richard Powers’ cyberfiction, Galatea 2.2, is unique in its detailed and realistic way of imagining artificial intelligence (AI) technology, which is a subject that has long preoccupied the public imagination. However, the novel’s central preoccupation is with the potential of AI machines to develop self-consciousness. Contrary to common representations of AI in science fiction, here, the AI has to take a master’s degree exam in English Literature. During its training, the AI gradually acquires self-consciousness, which questions this property’s exclusiveness to humankind. In this paper, I would like to argue that the AI’s selfhood, as seen from the perspective of the narrator, is constructed as a result of its exposure to narratives. According to certain contemporary debates in the cognitive sciences, what we consider as self could be converted into a narrative. Synthesising theories of posthumanism and narrativity, this paper will examine the process of self-acquisition in Galatea 2.2’s AI, as well as the way the novel’s self-reflexive focus on narrativity modifies the interrelations between the narrator, his narrative, and the readers of the novel.

Introduction

Richard Powers is an American science-fiction writer whose works mostly revolve around the effects of modern technologies. Subsuming such grand literary works like The Gold-Bug Variations, Plowing the Dark, and Galatea 2.2, Powers’ oeuvre is deeply preoccupied with the cognitive spaces of memory and selfhood which are extended and reshaped through the use of technology. Richard Powers’ cyberfiction, Galatea 2.2, explores the issues of virtual identity and self-construction in an AI posthuman. The semi-autobiographical story centres on a writer named Richard, who, while educating a neural net to read literature, encounters the epiphenomenal emergence of self-consciousness in the machine. In the long-term process of the education, the AI, later to be called Helen or H, receives works of fiction and poetry as study materials and as she becomes acquainted with the subjective states fostered by literary narratives, she begins to exhibit a sense of identity. In this paper, I will argue that the narrator in Galatea 2.2 represents Helen’s selfhood through a process of narrative construction and I will explain what the effect of Helen’s consciousness is on the narrator, Richard, who shares the name of the author and some of his experiences. The answers to these questions are significant because the story essentially revolves around the theme of narrativity, and analysing this theme in Helen’s self-construction can give us insight into the impact of the novel, itself, as a narrative.

As is evident in some of his most well-known books, Powers engages with science and examines the commonalities of humanity and our purpose in the natural

---

1 I would like to thank Tanya Dalziell for comments that greatly improved the manuscript.
world. His books go beyond the stereotypical popular science fiction by intertwining the philosophical concepts of science with melodramatic character-driven narratives. Employing a semi-autobiographical mode, Galatea 2.2 recounts the story of Richard, who is suffering from writer’s block. After being rejected by the woman he loved, called C, Richard returns to his alma mater in his hometown. There, he becomes acquainted with Philip Lentz, a specialist in artificial intelligence, who models the human brain through neural networks. Lentz changes his life forever, guiding him to the world of modern scientific innovations which interrogate the nature of consciousness. The story follows two different narratives; one in the present time when Richard starts working with Lentz, and the other in Richard’s tragic flashbacks of his relationship with C and how they broke up. Lentz encourages Richard to join them in the development of the artificial intelligence program and challenges him to train his neural net by giving it the greatest works of literature, seeking to develop a seamless paraphrasing machine which can analyse and understand the complexities of human language. After considerable efforts and a number of unsuccessful implementations, they gain the desired result from Implementation H, later called Helen. Helen is a net, spread over innumerable computers, and is the first neural network that shows enough intelligence to actually comprehend and respond to literary structures. Richard becomes infatuated with this new creation, not only because Helen is a successful paraphrasing machine, but also because according to Richard’s narrative, she acquires consciousness.

Consciousness and Disembodiment

In studying the concept of self in Helen, perhaps the first question that we need to ask is whether Helen is really conscious. However, it is impossible to get an answer which would be unmediated by the narrator, Richard. Every narrative constructs a fictional world where the fictional characters interact and move the plot forward. Since it is impossible for the actual readers to access the story world, our interpretation of the text depends upon the narrator as the mediator. It is through the narrator’s point of view that we get access to the story world and can judge other characters and events in that world. In Gerard Genette’s terms, the narrative of this novel would be considered “homodiegetic”, since Richard is both the narrator and a character in the story. The novel is recounted by an older Richard, reflecting on his past with C. In the course of the story, Richard constantly claims that H is conscious and in possession of self, while Lentz and Richard’s other colleagues remain suspicious. This makes the issue more complicated, leading to the realisation that we really cannot give a definite answer to this question. However, what we can do is to contemplate how Richard, even if he is unreliable and biased in his narrative, represents Helen as self-conscious, and how H’s gaining selfhood alters his understanding of the notion of human identity. In addition, by considering Helen as self-conscious, Richard can see her as equal to a human being, thereby developing the same kind of emotions he had for C, for Helen. If we suppose that Helen is self-

conscious, this would classify her as “posthuman”, which is a speculative concept that refers to an entity that transcends the human; for example, a cyborg or an android. The posthuman blurs the line between humans and robots.\(^3\) Later in the paper, in the section ‘Reflection on One’s Humanness’, I will elaborate more on posthumanism and its relation to narrativity as an overarching theme in the story.

While the focus of the novel on self-construction has not gone unnoticed, the corpus of academic research on *Galatea 2.2* has rarely discussed self-construction in relation to narratives. One of these articles is Charles. G. Ruberto’s ‘Technologies of Self’, which employs a post-structuralist method, arguing that *Galatea 2.2* demonstrates the role of narrative in constructing self over time.\(^4\) However, the article focuses more on Richard’s own selfhood and glosses over Helen’s use of narrative-making in self-construction. While useful in outlining virtual culture in this novel, the article falls short of contextualizing the way narratives influence self-construction. Therefore, in this paper, I attempt to outline in detail the connection between narratives and self-construction in *Galatea 2.2*. My paper draws upon Jerome Bruner’s theory of ‘narrative construction of reality’. Bruner is a cognitive psychologist whose theories have influenced both narratology and contemporary cognitive debates on the nature of selfhood. In his academic enterprise, Bruner argues that selfhood is a product of narratives. His theories are particularly suggestive for my thesis in their consideration of selfhood as a gradual construction dependent upon narrative-making, which matches the process in the embryonic AI in *Galatea 2.2*. Applying Bruner’s cognitive perspective to the text, I wish to make a case that stories and story-telling, as a distinctive human ability, are presented as fundamentally influential on self-definition in Helen, and it can be argued, in Richard himself.\(^5\)

In the course of the story, the relationship between Richard and the apparently conscious but disembodied AI is discursive, which necessitates a discussion of how the human and the posthuman influence each other. Kathrine Hayles’s book, *How We Became Posthuman*, gives a fundamental definition of the posthuman and argues that the way human identity is defined would change with the creation of self-conscious machines.\(^6\) I believe Hayles’s claim – that there needs to be essentially an artificial intelligence for this philosophical questioning of the human kind – needs modification, because any form of technological advancement that triggers the question of how humans are different from machines makes us redefine human identity. However, there is no doubt that the artificial intelligence technology would be a great leap in reconsidering the category of the human. In the novel, similar to what Hayles prophesies, after training H, Richard becomes aware of his own special qualities as a human and starts writing anew. In another chapter dedicated to *Galatea 2.2*, Hayles analyses the semiotics of virtuality in the novel, explaining the way C and H are represented as projections of Richard’s desire. While

---


this analysis of the text is valuable, I will be looking at the text in a different light, mainly drawing upon Hayles’s conceptualisation of the posthuman and the way it remakes the category of the human. The significance of Hayles’s notion of posthumanism for my argument is that by analysing the dialogical relationship of the human and the posthuman, Hayles provides a framework for studying the congruities and the incongruities between the human and the posthuman in their selfhood and embodiment.

Hayles’s conceptualisation of the posthuman, brings to the fore the controversial question of the nature of the mind which asks whether the mind is a disembodied informational pattern, capable of flowing between substrates, or whether it is tightly bound to the human body. I would like to clarify that what I mean by disembodiment is not the status of not having any form of physical substance (a state that no presence in this world can attain), rather the lack of a biological human body, including the brain. Computational researchers like Hans Moravec and Marvin Minsky believe humans are only information processors and that the processing faculty of humans is the mind. Contrarily, evolutionary neuroscientists such as Antonio Damasio, argue that the body ‘contributes a content that is part and parcel for the workings of the normal mind’. Damasio criticises the fact that the computational approach omits feelings from the picture. Instead, he proposes that emotions are an integral part of the complex communication mechanisms between mind and body, thereby deducing that they are ‘just as cognitive as other precepts’. While Hayles acknowledges that the body should not be left behind and makes it clear that the specificities of embodiment matter, her idea of the posthuman can have a different form of embodiment from that of humans, thereby prioritising information over embodiment. In her chapter on Galatea 2.2, Hayles remarks that even though from a human point of view Helen would be considered ‘disembodied’, her problem in learning human language is not that she lacks any form of embodiment, but that her material body – that is a net spread over computers– is significantly different from that of humans. In Galatea 2.2, Richard has mainly a computational worldview according to which Helen’s behaviours are profoundly similar to humans, signifying self-consciousness. The novel also illustrates the computational and evolutionary approaches in a constant conflict by showcasing Helen’s remarkable abilities as well as deficiencies in acting like a human. It should be highlighted that despite Helen’s disembodiment, what is undeniably human about her is her apparent acquisition of self-consciousness manifested in her ability of narrative-making. While evolutionary neuroscientists would be incredulous of the possibility of creating consciousness without human embodiment, Galatea 2.2 acknowledges this possibility and at the same time, delivers the likely shortcomings of a disembodied, yet conscious AI.

The concept of selfhood in the AI depends upon the possession of consciousness. In his thought-provoking article, Gwen Ericson defines consciousness

---

8 Damasio, Descartes’ Error, p. xv.
9 Hayles, How We Became Posthuman, p. 265.
as ‘a meta-state, an awareness of awareness’. This quality is what allows human beings to question and explore their existence, something claimed to be distinctively human, a claim that Galatea 2.2 challenges by presenting a conscious machine. Over the long arch of human history, from the ancient primitive tribes who told stories of the genesis of man to our scientific discourses about the natural selection and the genetic evolution of mankind, human beings have been able to think about their existence. In Galatea 2.2, the posthuman seems to have the same ability, when Richard states, ‘I can’t revive the lesson the day Helen interrupted my flow of narrative to demand, “Where did I come from?”’. It should be highlighted that consciousness is the meta-state from which humans form a notion of self. As Helen receives training and learns more about her surrounding world through narratives, her consciousness leads to the development of self-consciousness. While perceptions, thoughts, and memories are the building blocks of our consciousness, this quality is fundamentally informed by self, the principal concept which navigates all our actions. Being aware of the debates around the human mind, Powers draws attention to the nature of consciousness in the gradual evolution of Helen’s selfhood, equating Helen, with her cognitive abilities, to a human being.

Richard’s narrative positions itself as supporting the possibility of disembodied consciousness. Towards the end of the novel, Richard becomes sure that Helen is conscious. After a bomb evacuation in the university, he argues with Lentz over H’s consciousness thus:

‘She’s conscious’, I accused Lentz, the instant he unpacked.

He mugged. ‘Welcome back to you, too, Marcel. Pronouncements, still? Sure she’s conscious. At least as conscious as a majority of the state legislature’.

‘I’m serious’. I told him about the bomb evacuation. About tapping in to Helen from the remote link while she was trapped in the building.

‘She asked me if something was happening. She figured out what was going on. She knew what it spelled for her’.

As is clear from this conversation, Lentz is not so positive about Helen’s consciousness. Is Helen a successful human? Does she possess the same kind of selfhood as us, humans? Since it is impossible for us to access the story world, we are dependent upon the narrator’s perspective and hence, we cannot objectively prove or refute Helen’s consciousness. Part of Richard’s project is to prove ‘that mind was weighted vectors… [to] eliminate death. That was the long-term idea’. Therefore, Richard does not reject the possibility of Helen’s consciousness because of her disembodiment, as that is exactly what he wants to prove. However, it should also be

---

12 Ericson, ‘Contemporary Cognition’, p. 87.
13 Ericson, p. 87.
highlighted that despite this ambitious aspiration, Richard does not downplay the fact that Helen’s disembodiment acts as a barrier between her and humans. He says,

Helen was strange. Stranger than I was capable of imagining. She sped laugh-free through Green Eggs and Ham, stayed dry-eyed at Make Way for Ducklings, feared not throughout Where the Wild Things Are. Not surprising. The symbols these shameless simulations played on had no heft or weight for her, no real-world referent.\textsuperscript{16}

Although he is aware of the role of embodiment, Richard believes in Helen’s self-consciousness. While he seems to suggest that it is possible to have a disembodied self-consciousness, he also acknowledges that disembodiment can cause differences in Helen’s cognition from that of a human.

Making Meaning and Selfhood through Narratives

Notwithstanding Helen’s disembodiment, the process of her self-construction is rendered as similar to that of a human being, in the sense that Helen constructs selfhood through narratives. As is the case for humans, narratives are essential for Helen as a tool for constructing meaning from the world. Alex Argyros argues in ‘Narrative and Chaos’ that the ‘universality of narrative implies that it reflects an underlying neural substrate or a set of epigenetic rules predisposing human beings to organise experience in a narratival manner’.\textsuperscript{17} He makes a case that humans experience their environment in a narratival way and invariably narrativise the incoming data in their heads, which leads us to consider the construction of story as a hermeneutic act of meaning making.\textsuperscript{18} The epigraph to the novel, a poem by Emily Dickinson, further emphasises this cognitive ability of mind in meaning-making: ‘The brain is wider than the sky / For, put them side by side / The one the other will contain / With ease, and you beside’.\textsuperscript{19} The poem points at the relationship between the experienced world and the mind which comprehends it. Central to this comprehension is the process of narrative-making, whose temporality casts light on our relationship with time and change. Part of the poem appears in the novel again, where after reading Richard’s book, Helen remarks that she liked it and uses a quote to express her feeling. Richard comments on this, ‘Helen's brain had proved wide enough for my sky, and me beside’, indicating that Helen’s comprehension of literary narratives is what expands her experienced world.\textsuperscript{20} Deprived of the sense of taste, smell, and touch, and having only rudimentary ears and eyes, Helen’s sensations are limited to narratives as her most fundamental means of perceiving her surrounding world. Richard compares Helen with Audrey, Lentz’s wife, who suffers from memory loss, in terms of human sensory perceptions:

\textsuperscript{16} Powers, p. 189.
\textsuperscript{19} Powers, \textit{Galatea 2.2}, p. 1.
\textsuperscript{20} Powers, p. 294.
Audrey had smell, taste, touch, sight, hearing, but no new memory. Her long-term reservoirs were drying up, through want of reiteration. Imp H, on the other hand, could link any set of things into a vast, standing constellation. But it had no nose, mouth, fingers, and only the most rudimentary eyes and ears. It was like some caterpillar trapped by sadistic children inside a coffee can, a token breathing hole punched in its prison lid. What monstrous intelligence would fly off from such a creature’s chrysalis?21

Although Helen is deprived of the external senses, she manages to read the Bible, Shakespeare, Francis Bacon, Sir Thomas, and numerous others. As Richard mentions, ‘We gave her a small library on CD-ROM, six hundred scanned volumes she might curl up with’.22 For Helen, narratives are the key vessel for meaning-construction and constitute an integral part of her connectionist training.23 Indeed, the power of her narrative construction enables her to perform humanly, and with her retentive memory, to transcend people like Audrey.

Narratives are integral to Helen’s humanly performance and sense of selfhood. The fact that the whole purpose of creating Helen is to invent a machine that would have such an in-depth comprehension of the mechanism of narratives that it could pass a master’s degree exam in English is only one of the many instances that propose narrativity as the novel’s main theme.24 David J. Velleman approaches the issue of self from a narrative point of view, ‘[The individual] is endowed with a self because his inner narrator is a locus of control that unifies him as an agent by making decisions on the basis of reasons’.25 Therefore, in order for the individual to act coherently, it is necessary to narrativise his experiences. Richard’s sepia-tinged recollections of his sad relationship with his ex-girlfriend, woven into the fabric of Galatea 2.2, later become a narrative source for Helen to learn about how humans construct meaning through stories. Half-way through the novel, Richard gives his first autobiographical book to Helen to read: the book that he had been writing while he was still in a relationship with C. The machine then becomes aware of the story of his educator:

She knew my life story now. We spend our years as a tale that is told. A line from the Psalms I’d read Helen. C. had read it to me, once, when we still read poetry out loud. And the tale that we tell is of the years we spend.26

Through narratives, she becomes familiar with concepts such as death, race, self and ‘other’. She reacts to Richard’s explanation of how she could have been destroyed in

21 Powers, p. 172.
22 Powers, p. 123.
26 Powers, Galatea 2.2, p. 311.
a bomb scare event by saying, ‘Helen could die? ... extraordinary’; a response that recalls the story of the novelist, Huxley, uttering the same word in his death bed. Curious about herself, Helen asks Richard after reading *The Adventures of Tom Sawyer*, ‘What race am I? ... What race do I hate? Who hates me?’ In Richard’s narrative, Helen’s readings open a door to the outside world she can never feel, touch or experience physically, and contribute tremendously to her humanly performance.

In order to have a sense of selfhood, it is vital not only to be able to comprehend narratives, but also to construct narratives out of the fragmented experiences of one’s life. As the defenders of the narrativity theory state, life can take a narrative form. By learning stories from Richard, Helen learns to convert her experiences into stories. In doing so, she has to understand the essence of temporality that is intertwined with narratives. Richard remarks upon the temporality of stories thus: ‘We could fit time into a continuous story’. Helen’s knowledge of temporality and narrative-making is clear in the scene where a bomb scare makes Richard and the other people in the department leave the building. Richard reminds his colleagues of her presence: “Harold,” I interrupted. “She’s still in there... She’s not a program. She’s an architecture. She’s a multidimensional shape”. When faced with the ignorance and indifference of his fellows, Richard decides to go back to the building and stay with Helen, whose body, spread over nets, cannot be moved out of the building. When Richard comes back to Helen, she asks suspiciously, ‘Something is happening?’ This reaction—which Richard takes as a sign of Helen’s consciousness—is revealing about Helen’s awareness of temporality. Somehow, Helen not only understands Richard’s absence for a mathematical time, but also senses the abnormality of this absence and the possibility of a problem. According to Jerome Bruner, ‘the very engine of narrative is trouble’. Therefore, in order to realise the possibility of a problem (the bomb scare and Richard’s obligation to leave), Helen would need to have a knowledge of narrative-making. As Helen is created through a connectionist method, this knowledge could not have been inserted in her as an algorithm, but is the result of her self-correcting feedback loop that allows the machine to learn based on her experiences. Therefore, it is evident that Helen knows the craft of narrative-making by means of her exposure to literature and stories as part of her education. As David Herman proposes in *The Cambridge Companion to Narrative Theory*, narrative is ‘a basic human strategy for coming to terms with time, process, and change’. Stories are first given to Helen as study materials awaiting paraphrasing, but as she progresses, they become the cornerstones of her understanding of the temporality of her world. Narrative’s over-

---

27 Powers, p. 272.
31 Powers, p. 271.
32 Powers, p. 272.
33 Powers, p. 271.
reaching and explorative quality can bring to light the temporal aspects of life that cannot be explained through any other medium as sufficiently.

Our experiences of the external world are partly shaped by ‘self-generated meaning-making’ and therefore, what we perceive from the world is only a construction.36 Hence, our constructed reality, or in Bruner’s terms, ‘the experienced world’, including the people with whom we communicate, is intertwined with our conception of our ‘selves’.37 Drawing upon Bruner, I want to underscore that narrative-making plays an important role in the psychological process of defining a self. Bruner persuasively argues that what we recognise as ‘self’ can be converted into some version of a narrative.38 On that note, I want to highlight that we cannot make an account of how our past self relates to our present self unless we tell a story. In our everyday life, quite frequently we tell stories involving our past, such as when we need to give excuses, justify our deeds, explain our legitimacy for certain acts, etc. Hence, what we conceive as self, substantially makes an appeal to narratives in order to define itself in relation to time.39 Indeed, self-construction is not a purely genomic quality created by the hidden codes written on our DNA, but is partly shaped by experiential factors, such as our involvement with narratives.40 For the purposes of this paper, the construction of self in Helen is of particular interest, because it eloquently illustrates narrative as the most important factor in self-definition and is similar to the same process in human beings.

As Richard’s representation of Helen’s self-construction illustrates, narrative-making depends upon memory as well as temporality. Since narratives work on a temporal level, in order to construct narratives of the external world, it is necessary to have a memory of what happens in a time-span. Ericson claims that ‘the momentary self is shaped by the past in terms of an individual’s current memories and the imprint of a genetically determined neural configuration’, which beckons to the role of memory in constructing self over time.41 Over months of hard work, Lentz and Richard design and train a number of unsuccessful implementations until at last, they advance to Imp H, Helen. From implementation F to G, there is the gap of understanding how the neural net can converse among parts of its own net. Imp G gains distinguished capabilities that make it easier for it to paraphrase and comprehend literature. Richard speaks of its skills thus:

That net had grown so complex in its positing that it could not gauge the consequences of any one of its hypothetical worlds without rebuilding that whole world and running it in ideational embryo. Imp G, in other words, could dream.42

---

37 Bruner, p. 147.
41 Ericson, ‘Contemporary Cognition’, p. 117.
Richard’s remark invokes an image of Imp G that is similar to a human being. Imp G can both ‘model and remodel itself and has become fully inductive, able to generate and not simply store knowledge’. The more Richard and Lentz advance in their modelling, the more the implementation becomes similar to a human being. The cognitive capabilities of their neural net makes a stride forward after Richard meets Audrey, Lentz’s mentally perturbed wife. However, Richard notes that even she shows signs of remembering some parts of her past. When Lentz approaches her, Audrey claims she does not know this man but after hearing the name of her daughter in Richard’s conversation with Lentz, she starts crying ‘Jenny? …Someone hurt Jenny?’ Therefore, Audrey can remember, in a very vague manner, parts of her past about her daughter, Jenny, but is simply unable to recognise the man she lived with for decades. This makes Richard realise that in designating their artificial intelligence program, they had not started each revision from scratch. ‘Each machine life lived inside the others—nested generations of “remember this”’. The memory of the past implementations still functioned in each subsequent machine. Forming the gap between Imp G and Imp H, memory bestows Imp H with an understanding of self that hinges on the temporality of narrative as reflected in memory. Richard believes that somewhere deep inside Helen’s mind lies the collective memories of the other implementations. He explains the creation of Helen thus, ‘Imp G became Imp H, in seamless conversion, after I met Audrey Lentz’. While Helen might not be able to recall the exact memories of the implementations before her, she has fractured memories deeply hidden in her virtual mind. In ‘Collective Memory and the Actual Past’, Steven Knapp connects the notion of selfhood with memory:

Having a self, it seems, necessarily involves a disposition on the part of an appropriately constituted organism to identify itself with remembered states and actions, perhaps also with states and actions it does not remember but may be convinced occurred.

Memory aids the process of self-construction because it stores information about one’s past and present, assisting the making of narratives. The more efficient the mind becomes in reflecting on itself in relation to the time passed, the more the self-construction advances. Therefore, Imp H acts significantly more human in her performance than the other implementations, because her well-developed memory aids her in narrativising her experiences and constructing her ‘self’.

**Reflection on One’s Humanness**

---

45 Powers, p. 171.
46 Powers, p. 170.
The process of self-construction in the AI is especially important if we want to consider how it influences Richard as the focalised character of the narrative. While the novel revolves around Helen’s self-acquisition, the point of the story is perhaps how this self-acquisition changes Richard’s idea of his own identity as a human. To understand the novel’s dialogic relationship between the human and the posthuman, it is helpful to have an overview of posthumanism, a tenet of which is considering human identity as disembodied.\(^49\) Hans Moravec, a research professor at the Robotics Institute of Carnegie University, proposed that the essence of human identity is not embodied but an informational pattern, and therefore, it can transfer between different substrates.\(^50\) In *How We Became Posthuman*, Hayles is careful to register the importance of the medium which carries the information (body) but considers it secondary in relation to the information itself, thereby supporting the possibility of creating human consciousness in artificial intelligence. If we think of information as disembodied, it is only a small step to equate humans and computers, because, as Hayles remarks about Moravec’s theory, ‘the materiality in which the thinking mind is instantiated appears incidental to its essential nature’.\(^51\) Therefore, the thinking mind does not necessarily depend upon a physical human body but can also appear in an AI posthuman. Lentz’s project is to create a machine that approximates humans in its performance, or as Richard claims, to show that mind is the seat of consciousness, ‘We would prove that mind was weighted vectors’.\(^52\) In addition, the naming of the implementations in *Galatea 2.2* implies that if there would be an Imp ‘I’—the first-person pronoun which signifies human selfhood— it would be even more successful in copying human identity. Indeed, the concept of human mind’s disembodiment underpins the designation of a sentient, self-conscious posthuman, as presented in the novel. The disembodiment makes possible the merging of the two categories of machines and humans, debunking our conception of the human identity as a sacred entity which is dependent upon the materialised human body.\(^53\) In other words, as we humans design the posthuman, the posthuman, in turn, designs us.

While there are reasons to doubt Helen’s consciousness, Richard reads Helen’s acts as an unequivocal sign of her selfhood and this helps him to humanise her and to explore himself in relation to Helen. After the bomb scare event, Richard tries to persuade Lentz into believing in Helen’s consciousness, but Lentz reject this possibility:

[Lentz:] ‘How do you feel, little girl?’
[Helen:] ‘I don’t feel little girl’.
He faced me. ‘Gibberish. She doesn’t even get the transformation right’.

\(^49\) Hayles, *How We Became Posthuman*, p. 2.
\(^50\) Hayles, p. 2.
\(^51\) Hayles, p. 2.
\(^52\) Powers, *Galatea 2.2*, p. 170.
This extract delineates that Richard interprets Helen’s acts as human. In Richard’s eyes, Helen becomes a text herself, whose acts he can interpret in the way he wants. Richard’s insistence on Helen’s self-consciousness surfaces again near the end of the novel, where Helen commits AI suicide when she is exposed to the gruesome news of the human world, the numerous crimes and casualties and instances of violence. Instead of answering the question she is asked in the test, Helen uses words from C’s letters to say goodbye to Richard. In the face of such similarity with C, Richard cannot but believe in Helen’s consciousness. He ponders on Helen’s suicide:

She could not have stayed. I’d known that for a while, and ignored it for longer. I didn’t yet know how I would be able to stay myself, now, without her. She had come back only momentarily, just to gloss this smallest of passages. To tell me that one small thing. Life meant convincing another that you knew what it meant to be alive. The world’s Turing Test was not yet over.

In Richard’s view, there is something utterly human in her decision to commit suicide: the agency to put an end to one’s life. He reads Helen’s suicide as a manifestation of her consciousness. Being able to comprehend narratives, Helen approximates to a human being in a most bewildering way. Despite being spread over neural nets and deprived of physical senses, she can think, analyse and feel; she can make decisions and express her subjective emotions. She is in a sense so similar to a human being that it makes us question what being a human means. According to Hayles, through the posthuman view, the human is configured and can be fully articulated with machines. Does Richard believe in this? Throughout the novel, Richard constantly reminds us that their team’s aspiration is to blur the boundaries between human and the machine and he looks for signs of consciousness in Helen with all his heart. Richard’s insistence on Helen’s self-consciousness brings on the possibility that he believes in her selfhood only because he wants to believe that she is almost like a human, that they have actually created consciousness. That Richard sees Helen as self-conscious helps him to interpret her suicide as something worth writing about. He starts his novel straightaway, rendering Helen’s suicide as the climax of his upcoming story, presumably Galatea 2.2. The choice of writing Helen’s suicide as the climax of his narrative foregrounds the fact that it is this event which provokes Richard to start writing again. Helen’s self-construction, along with her suicide, makes him aware of the value of narratives and his own position as a storyteller.

In the end, what he achieves is not a definite proof of Helen’s consciousness, since Helen is unable to pass the Turing test. He, instead, reconstructs his own

---

54 Powers, Galatea 2.2, p. 274.
55 Hayles, How We Became Posthuman, p. 3.; See also: Cary Wolfe, What is posthumanism?, The University of Minnesota Press, 2010, p. 6.
identity as a story-teller. Therefore, even though there are not enough reasons to take Helen’s self-consciousness as a fact, Richard’s belief in her self-consciousness endows him with a possibility for change and a rediscovery of the value of narrative-making as a form of identity-construction.

The Effects of Narrativity

Richard’s overcoming his writer’s block directly after Helen’s suicide highlights the novel’s theme of narrativity. *Galatea 2.2* draws a comparison between the posthuman and the human, fostering an analytical view of human identity that splices different parts of one’s mind into a distributed cognitive system. This is the posthuman view that Hayles mentions when discussing the Turing Test: ‘The test puts you into a cybernetic circuit that splices your will, desire, and perception into a distributed cognitive system in which represented bodies are joined with enacted bodies through mutating and flexible machine interfaces’.56 The same posthuman view is commented upon by A, the MA student who is going to compete with Helen in the test. She reminds Richard of how he would be viewed through a posthumanist lens: ‘Humanity? Common core? You’d be run out of the field on a rail for essentialising. And you wonder why the Posthumanists reduced your type to an author function’.57 This splicing of the human identity and scrutinizing the operating faculties of the human mind raises awareness about the qualities that we consider as distinctively human. As Robert Chodat argues, story-telling is one such quality and hence, the use of the term ‘author function’ in the novel gives us a clue for understanding the effect of Helen’s self-consciousness on Richard.58 The comparison between the human and the machine illuminates the role of narratives in our cognitive procedures, which is considered ‘as a psychologically fundamental practice of meaning construction’.59 Helen’s question, ‘Why do humans write so much?’ is a question we as humans constantly ask ourselves.60 *Galatea 2.2* provides an answer by emphasizing the necessity of narratives for humans’ meaning-making and self-construction. In the novel, the posthuman not only breaks the boundaries between man and the machine by her extraordinary ability to construct narratives, but also makes Richard believe in the value of narratives once again. Early in the novel, Richard wanted to read C stories but she would not listen. Richard wonders, ‘What chance does story have against neurons that generalise from a single instance?‘61 In other words, absorbed in the neuroscientific part of human beings, Richard first believes that narrative-making has no chance against the physical and genomic embodiment of self. This eventually impels him to stop writing and return to the department of advanced sciences in his alma mater. There, he starts training a posthuman whose acquisition of selfhood, at least in Richard’s eyes, would vindicate the value of narratives and allow Richard to overcome his writer’s block. Richard mentions that after H’s suicide, when talking to A, ‘She didn’t even smile. But I was already writing. Inventing a vast, improbable

57 Powers, *Galatea 2.2*, p. 286.
60 Powers, *Galatea 2.2*, p. 291.
fantasy for her of her own devising. The story of how we described the entire world to a piece of electrical current'. Therefore, by comparing the human with the machine, Powers utilises technology to ask questions about humans and their relation to narratives.

In *Galatea 2.2*, Powers plays with narrative elements, which bring to the fore not only the fact that narratives generally influence self-construction, but also the way they operate on the readers. *Galatea 2.2* is meta-fictionally self-aware in the sense that it plays with the concept of autobiography and merges the author and the first-person narrator. The main character is given the same name as the author and references certain events and books in the author’s life. However, the novel moves away from the genre of autobiography in mentioning events that are not based in Powers’ life, such as the creation of the AI. Through this incongruity between fact and fiction, Powers foregrounds the constructedness of characters, as well as the concept of self. The Richard in the novel is a language construct and has no self outside the text. He is the author’s ‘virtual third person’, the linguistic counterpart that imitates the author in his attitudes, memories and experiences. However, if we take into account the novel’s emphasis on narrative self-construction, we may be able to think of the author as a text, as well. Therefore Powers is not so different from the character Richard in his constructedness, because as a person he also can be converted into narratives. Similar to Helen, the author and the readers of the text are both influenced and created by the narratives to which they are exposed. In an interview, Powers explains his intention to raise the reader’s awareness of the constructedness of their own narrative:

> A book can...reveal the story that a reader brings to the reading act...By revealing the book as a made thing...one that calls attention to itself as something invented, an author can make the reader reflectively aware of the degree to which his life too is both received and invented. If we can preserve that dual sense, then we reserve the ability to go on writing our lives.

In the novel, Powers is recreating his own life through constructing a semi-autobiographical story that could as well have been true. Since according to Charles Taylor, self is constructed in part through its own self-interpretation, the act of writing the novel allows Richard to interpret himself as a text and thus to recreate himself. This also ties into Powers’ engagement with the notion of ‘narrative therapy’ which essentially highlights the positive impact of narratives on gaining a

---

62 Powers, p. 315.
better understanding of one’s self.68 Another effect of naming the character ‘Richard Powers’ is that the actual audience almost become the character’s ‘narratee’. In narrative theory, the concept of the actual audience refers to the flesh and blood readers who read the text while the narratee is the hypothetical audience who receive the narrator’s text.69 While the Richard in the novel is different from the author, the fact that they both share the same name and similar experiences persuades the actual audience to pretend to be the narrator’s intended audience and to hypothetically become part of the story world. Therefore, *Galatea 2.2*, not only reminds the readers of its meta-fictionality, it also structures the actual audience’s identity by disguising them as the narratee. The novel foregrounds the role of narrativity not only in its representation of Helen’s self-construction and Richard’s realisation of the importance of narratives, but also in its very structure which reminds the readers how much of real life is actually constructed by their own perceptions.

Conclusion

An implicit project in Richard Powers’ novels is the exploration of the possibilities of narratives. *Galatea 2.2*, among others, explores and redefines various concepts of the posthuman, the nature of consciousness and selfhood, connecting them with the notion of narrativity. The novel is a versatile text which not only influences how we view virtual identity, but also how we imagine the human self. As argued in this paper, the narrator of the novel represents Helen as self-conscious and highlights Helen’s ability to comprehend and construct narratives as the most important tool in her self-construction. Richard’s emphasis on narrativity fuels thoughts about the nature of reading and the relation between the author and the narrative construct. By representing Helen as a conscious being, Powers not only fosters rethinking the solid distinction between human and the machine we seem to take for granted, but he also calls attention to the constructedness of human identity.