Whose Real? Encountering New Frontiers in *Westworld*

*Rory Jeffs and Gemma Blackwood*

**Abstract**
This article explores the themes of psychoanalysis and Lacanian theory discernible in the HBO version of *Westworld* (HBO 2016-). This article argues that it is productive to examine the series through a psychoanalytic lens, especially through Jacques Lacan’s concept of the Real. In an examination of Season One of the series, we argue that this new show – given its focus on psychological landscapes and trauma – requires a deeper textual analysis to uncover similar Lacanian themes. These are themes signposted by the quite literal depictions of therapy sessions shown throughout the season. At the heart of the show lies a double irony about the human and non-human characters in the show. *Westworld*’s robotic characters are gripped by the complexity of their memories: memories pre-programmed, easily adjusted and changed by the scientists running the park. Then, we see the actual human figures of the show in hot pursuit of violent sexual desires, thoughtlessly enacting free plays of desire without consequence and the conscientious interruption of the superego or the ‘Law’. To explain this, we firstly provide an introduction to the contextual origins of the three psychoanalytic orders (or triad) developed in Lacan’s writings, particularly looking at *Seminar II* of 1954-1955. Using this framework, we then focus on a textual analysis of four key characters in the series, and how each negotiates journeys through these conceptual/spatial zones that delineate some ‘encounter with the Real’. Finally, we argue that the text’s complex narrative devices become allegorical for the viewer him- or herself as a lost figure in search of answers, allowing a meta-analysis of key themes of the show.

---

Rory Jeffs is completing a PhD in Philosophical Studies at Deakin University. His co-edited collection *100 Years of European Philosophy Since the Great War* was released in March 2017.

Dr. Gemma Blackwood is Lecturer in Communication Studies at Charles Darwin University and researches on visual culture, film and travel cultures. Her recently co-edited collection *Motion Pictures: Travel Ideals and Film* was released in 2016.
Introduction: Welcome to the (Real) Westworld

In October 2016, the US cable network HBO premiered a high-profile new television series about a futuristic robot-run theme park called Westworld. Like many of the blockbuster television series that have been created by premium cable in recent years, it drew its name and its premise from a pre-existing text, other examples being Game of Thrones (2011- ), Twin Peaks: The Return (2017) and The Handmaid’s Tale (2017- ). In this example, the premise comes from the 1973 neo-Western science-fiction film written and directed by the popular novelist Michael Crichton. This film centres on a male buddy trip to this Wild West theme park gone terribly awry: tourists are caught out with malfunctioning robot characters ready to kill. In many ways a B-grade monster movie, it dramatises a literal ‘rise of the machines’, although in this case the robots are able to be defeated at the end. The film addresses cultural anxieties about formations of masculinity in the technologising modern society of the USA in the 1970s.

However, while the television show’s premise of the robot theme park comes from the 1970s film, the scope and narrative direction has been reimagined in the new text. Chiefly, the biggest shift to the narrative of the television series is the focus on the human-like robots passively confined to repeat a daily cycle – often involving their own torture, abuse and even murder – at the whims of the park’s human tourists. These robots are 3-D printed with flesh and blood and are owned by the park. They are given nightly memory wipes to keep them functioning according to their pre-programmed narratives (called on the show ‘storylines’). Essentially, the major shift between the two Westworld texts is who we are supposed to empathise with; the television series twists the original film plot and asks the audience to sympathise with the robots instead (who are directly referred to as ‘hosts’ to the human ‘guests’). The first season of Westworld follows the dawning consciousness of some of the robots as they become fully aware of the ‘reality’ of the park. This discovery culminates in a robot revolution against the human guests in the season’s finale.

The 1973 film was noteworthy for being one of the earliest Hollywood films to utilise computer-generated imagery (CGI) technology, to visually create the sense of what Crichton described as a ‘bizarre computerised image of the world’ (cited in Price 2013). The original film is a comment on the rise of virtualised themed spaces or theme-park locations across post-industrial America, and even the Hollywood-inspired theme parks such as Disneyland (opened in 1955) and the tourist attractions that commenced at Universal Studios Hollywood in 1964 (Campbell and Kean 2016, 126-28). The Westworld park refers heavily to the Hollywood Western themes and also more general romantic travel fantasies about frontiers and exploration. HBO’s Westworld comes onto our small screens amidst a wave of new series that adapt traditional science-fiction themes to dystopian not-too-distant futures, such as two British TV productions, Humans (2015- ) and Charlie Brooker’s Black Mirror (2011-2016). On the surface, Westworld appears more reflective of standard science-fiction
exploration of Artificial Intelligence in the tradition of Ridley Scott’s seminal film *Blade Runner* (1982). Indeed, *Westworld* explores these ideas of simulation, technology and robotics. It also examines the politics and ethics of robotic technology and the exploitation of non-human entities, raising biblical analogies about morality and the free will of humans. Questions over what it means to be human overlap here with what it might mean to be non-human, and this is centrally examined through the dual concepts of memory and desire.

In this article, we would like to explore the themes of psychoanalysis and Lacanian theory discernible in the HBO version of *Westworld*. We argue that it is productive to examine the series through a psychoanalytic lens, especially through Jacques Lacan’s concept of the Real. This account was partly formulated through Lacan’s clinical cases on the failure of the analysand to name a trauma and to symbolise it (working the symptom and reintegrating it back into language or ‘speech’). Hence, Lacan’s Real is simultaneously both outside and within the symbolic network of our socialised ‘reality’ of language, as it is part of the relation between the ‘speaking being’ and the unconscious. The Real then cannot be discussed without reference to Lacan’s ‘three orders’ or ‘three registers’, which also included the Imaginary and the Symbolic orders.¹

To think about the Real in relation to a TV series such as *Westworld* requires a distinction from what we now may call the ‘classic’ 1980s and 1990s science-fiction postmodern theories of simulacra. The ‘reality effect’ or ‘hyperreal’ that results from the ‘seduction of simulacra’, as Baudrillard (1994) put it, proved a seductive paradigm for understanding the post-*Blade Runner* films and television programs, most notably, the Wachowskis’ *The Matrix Trilogy* (1999-2003). It was *apropos* his example of the *Matrix* after 9/11 that saw Slavoj Žižek engaging with Baudrillard in his book *Welcome to the Desert of the Real* (2002). Žižek found the postmodern theory being applied to such films worked too smoothly, and in some sense fell into the idealistic meta-commentaries on the theory itself, endlessly re-defining its conditions rather than yielding any insights on what living in an endless simulacrum might mean philosophically. In comparison to the Lacanian psychoanalytic concept of the Real, the Baudrillardian mode of deconstruction delimits the Real to a ‘hyperreal’ process of the simulation of an imaginary or specular order that can only ‘seduce’. Hence, it discounts Lacan’s explanation of psychic phenomena as crossing three structured orders that combine and structure a person’s relationship to language, desire and truth.

Lacan readily admits that the human sense of ‘reality’ is determined by the symbolic order of representations, but this is not where ‘truth’ lies. The truth is a relation in the Real, or the gap between the symbolic and Real (Flieger 2003; Shepherdson 2008, 32-33). Lacan’s concept of the Real then has become influential amongst theorists who reveal the complexities of the transition from actual to virtual, or as Žižek puts it, the ‘reality of the virtual’. Given his view that ‘truth has the structure of fiction’, there is a liberty with Lacan’s
work to gather examples from fictional sources and to further understand how the psychological Unconscious can manifest (Lacan 1992, 12; Lacan 2006, 7).

In an examination of Season One of Westworld, we argue that this new show – given its focus on ‘psychological landscapes’ and trauma (Heer 2016) – requires a deeper textual analysis to uncover similar Lacanian themes. These are themes signposted by the quite literal depictions of therapy sessions shown throughout the season. At the heart of the show lies a double irony about the human and non-human characters in the show. Westworld’s robotic characters are gripped by the complexity of their memories; memories are pre-programmed, easily adjusted and changed by the scientists running the park. Then, we see the actual human figures of the show (the ‘guests’ of the Park, often white adult males) in hot pursuit of violent sexual desires, thoughtlessly enacting free plays of desire without consequence and the conscientious interruption of the superego or the ‘Law’. In this simulated lawlessness, the frontier theme and masculine fantasies of the ‘wild west’ take centre-stage as they did in the original film. In one way, this irony reveals what Dr. Ford seeks to understand as he looks to the differences between humans and robots: who is more ‘free’ of the two? How does one become a free ‘human’ subject? We think that Westworld can also allegorically relate to the mediation of the Real through the Symbolic; we argue that the text provides a way for us to see how fictional or highly representational worlds reveal a truth or gap in our reality.

To respond to this argument, the article focusses on a textual analysis of four key characters in the series, and how each negotiates journeys through these conceptual/spatial zones that delineate some ‘encounter with the Real’. We examine two ‘human’ characters, looking at Dr. Robert Ford (Anthony Hopkins) and William/’The Man in Black’ (Jimmi Simpson/Ed Harris); and then two ‘non-human’ robot hosts, Dolores Abernathy (Evan Rachel Wood) and Maeve Millar (Thandie Newton). Interestingly, these are four characters who, regardless of their ‘humanity’ (or not), all have senses of reality structured in relation to a signifier bordering between the Symbolic and Real registers. To explain this, we firstly provide an introduction to the contextual origins of the three psychoanalytic orders (or triad) developed in Lacan’s writings, particularly looking at Seminar II of 1954-1955. We note how Lacan’s triad partly arose out of his interest in cybernetic theory, game theory and machines: in fact, we argue that it led him to theorise the autonomy of the symbolic order from the imaginary, brought him to reflect on questions of madness or psychosis and the trauma of the Real, and even helped him to specify the Real in relation to the Symbolic. This becomes an effective and a useful framework to further understand the relationship of the characters between consciousness, meaning, and suffering, and the difference between merely following a ‘code’ and a ‘signification’. Finally, we argue that the text’s complex narrative devices become allegorical for the audience viewer him or herself as a lost figure in search of answers, allowing a meta-analysis of key themes of the show.
The Thing May Think, But Can It Go Mad?

Despite the long litany of references Lacan makes to it, the Real remains ambiguous enough to elude any interpretative systemisation, as has been discussed comprehensively by Tom Eyers (2012). However, even in Lacan’s own writings and the clinical context of his practice, the concept has attracted much debate and controversy, so much so that even Žižek (2001b) suggests a three-fold framework: an ‘imaginary real’, ‘symbolic real’ and ‘real real’! Nevertheless, because of Lacan’s shape-shifting discussion of the Real, its adaptation by theorists to a wide range of contexts – whether political, social, cultural, or media-related – reflects on the appeal of this concept across a range of disciplines. There has been little discussion of Lacan’s Real in relation to artificial intelligence. Perhaps this is partly due to the nascent nature of discussions on AI that occurred simultaneous to the period of Lacan’s research in the mid-twentieth century. Yet, Lacan clearly shows interest in the emerging technologies relating to machines and cybernetics: for instance, in Seminar II of 1954-55 on The Ego in Freud’s Theory and in the Technique of Psychoanalysis (1988b). This Seminar concerns many of the themes that come to dominate Lacan’s later work, such as the concept of repetition, the difference between the other and the Big Other, the unconscious as the ‘discourse of the Other’, schema L and, of course, the cementing of Lacan’s famous Imaginary-Symbolic-Real triad (or as he later called it, the ‘Borromean knot’ in relation to the interconnected but not fully linked connections between the three).

After already posing the differences between the Imaginary and Symbolic orders, and also their mutual dependency in Seminar I (Lacan 1988a), Seminar II provides pivotal insight into Lacan’s ‘return to Freud’ via linguistics and the structuralist anthropology of Levi-Strauss, centering on his formulation of the ‘symbolic order’. Lacan takes his cues from Freud’s ‘discovery’ of the unconscious as a ‘Copernican turn’ in our understanding of the ego and consciousness. He firstly posits that the ‘I (Je)’ or the ‘subject of enunciation’ of the unconscious is not the ego (moi) or what is commonly understood as consciousness, or even some ‘true ego’ (1988b, 44). By critiquing the imaginary function of the ego and its aggressivity, Lacan can explain how social norms can take root by outlining how the Symbolic order ‘intervenes’.

To understand the efficacy of the Symbolic order, Lacan compares the operations of the Symbolic through human language as homologous to the information-circulation in early computers, a radically different understanding compared to Freud’s biologistic-hydraulic model of the unconscious. Lacan’s sudden gravitation towards cybernetics seems to have been driven in part by theorising the Symbolic order as governing the structure of language (une langage) but also being irreducible to natural (or national) language (un langue). In this way, the Symbolic order becomes analogous to a particular kind of computational processes of information machines (i.e. the sequence of signs, or the binary system of 1s and 0s) that can generate automated differentiation. With Alan Turing’s work on computation becoming
popular, cybernetics was beginning to shape the early AI research that formed out of the Macy Conferences of the early 1950s (Dupuy 2000, Johnston 2008).

Lacan’s interpretation that the ‘symbolic world is that of the machine’ means that in the structure of language (and symbolic relations), ‘syntax exists before semantics’ – and we find ourselves ‘thrown into its gears’ and governed by a chain of signifiers (Lacan 1988b, 47, 64-92. 307). Dupuy (2000, 19) argues that cybernetics enabled Lacan to conceive of the unconscious as the ‘repetition’ of the automaton (the machine qua signifying chain) rather than Freud’s death drive and its compulsion to repeat beyond the pleasure principle (Wiederholungszwang). Lacan’s reading of Poe’s The Purloined Letter and the reference to the chance game of ‘odd or even’ in Seminar II (1988b, 175-205, cf. Lacan 2006,6-50) was to establish why the analytic experience is not limited to the imaginary form of intersubjectivity (or ‘imaginary effects’). The experience cannot be properly understood without its basis in a symbolic chain, showing how the unconscious works behind the backs of subjects. Liu (2010) suggests that Lacan employed a ‘cybernetic unconscious’ to demonstrate how this functions. Hence, according to this articulation, the question of thinking machines has become redundant, given that human beings employ operating processes with the same symbolic functions that symbol-processing machines use in cybernetic systems. The following was Lacan’s response to Octave Mannoni’s ‘serious’ concerns over his concept of the symbolic as a machine:

[T]he question as to whether it [machine] is human or not is obviously entirely settled: it isn’t. Except, there’s also the question of knowing whether the human, in the sense in which you understand it, is as human as that (1988b, 319).

Lacan later clarifies that the sliding of signifiers structures the repetitive metonymy of desire (in the ‘Agency of the Letter’ lecture in 1957), a process which he continues to analagise to the structure of modern-day thinking machines. He suggests that for the communication ‘message’ to be received as an ‘act of signification’, it has to ‘cross the bar’ back into the signified, which it can only do through becoming a metaphor (i.e., ‘master-signifier’ or a ‘signifier for another signifier’) (Lacan 2006, 428-31, Lui 2010, 295). For this reason, he re-evaluated psychosis in Seminar III (2003), looking at how psychopathology itself is an integral part of human subjectivity. He posited that machines would never be able to go mad like humans, even if they were able to process language as code or information. Machines could only ‘jam’ or collide with one another, if they were truly ‘autonomous’ and without ‘symbolic regulation’, and even if such a form of regulation evolved the machine could not be an ‘entity’. In an ironic way, the cybernetic form of the symbolic order cannot be reduced to the figure of a ‘legislator’, or to the interventions of programming. This conceptualisation would prove to challenge the optimism of early ‘hard’ AI. The insistent intervention of the symbolic has to be the ‘voice of no one’, which leads a subject (on an unconscious level) to
‘count themselves as one [an imaginary unity]’ (Lacan 1998b, 47-51, 54-55). At this stage, Johnston (2008, 96-98) argues Lacan has no other way to define the ‘impossibility that defines the human’ but as the existence between two symbolic processes: that is, mechanic repetitive processes that integrate the subject into the Symbolic order, and the mechanisms that resist such integration. Such an impossibility arguably leads Lacan to shift the place of the Real in relation to the Symbolic, from syntax to non-being. This brings us to the import of the Real in the triad. If the Symbolic and Imaginary make up what is understood as ‘reality’ for Lacan, there is something more real than this ‘reality’, especially if the Symbolic order is mistaken for a second-order imaginary.

At this stage Lacan’s first experimentation with his new triadic system as a way to interpret a range of psychic phenomena faced a crossroads in understanding the Symbolic-Real relation. Confronting two concepts or functions of the Real – the pre-symbolic (‘irruption’) of the Real, as he discussed in cases of psychosis and where the signifier can return in the ‘Real’, and the post-symbolic Real , resulting from the repetition of the signifier and ‘encounter [tuche]’ with the Real at the limit of the Symbolic (Lacan 1981, 53-54; Shepherdson 2008, 34). The early seminars (e.g. I-III) or the so-called ‘middle period’ come to distinguish the Symbolic and the Real in opposition to one another, where the Real has a kind of brute pre-symbolic reality that is non-dialectical or ineffable. Here, the example of Freud’s dream of Irma’s injection is cited by Lacan in Seminar II to explicitly signify the ‘privileged experience’ of ‘what we can call the revelation of that which is least penetrable in the real, of the real lacking any possible mediation, of the ultimate real’ (1988b, 165-65, 175-77). However, Lacan’s later more frequent discussions of the Real, beginning from Seminar VII (1992), adopt the Real as Freud’s ‘Das Ding [The Thing]’, which becomes linked up to Mother qua ‘(M)Other’, objet petit a and jouissance, and effectively displaces the cause of desire from the symbolic chain to the Real. This is desire not as a desire for recognition, but rather the desire of a ‘lost [non-imaginary, non-symbolisable] object’, or the ‘little piece of the Real’ (Žižek 2001a).

This re-pivoting led Jacques Alain-Miller (2004) to see Lacan’s later work as developing a more materialist account of the Real as an ‘intimate exterior’, operating closely within the Symbolic and the phantasms of the objet petit a. It repeatedly returns to any ‘functional’ Symbolic order as the gap that resists symbolisation or, as Eyers (2012, 5) says, is a ‘signifier-in-isolation’. This heightened sensitivity to the role of the Real in the three orders would define the ‘later Lacan’ for the last two decades of his seminar series, with a privileged emphasis on the Real over the Imaginary and Symbolic. However, the advantage of this shift was the prospect that the Real had a relation that emerges from out of the Symbolic and Imaginary orders, and could be adapted to meta-psychological socio-cultural fields outside the clinic. In this sense, fictional or virtual representations or fantasies themselves can be screens upon which the hole of the Real can become tangible as a breakdown of signification.
This brings us to the question of how the Real could be ‘represented’ or mediated in *Westworld*. We have elaborated on Lacan’s interest in cybernetics and how that mediated his initial development of the concept of the Real itself (by way of the mechanic model of the symbolic order). This question has even more significance in relation to representations of artificial intelligence and anthropogenesis qua the formation of ego and the Symbolic order. For how can the Real be represented in filmic and literary representations of artificial intelligence itself? Do the machines themselves become conscious or alive? In the opening chapter of *Tarrying with the Negative* (1993, ch. 1), Žižek writes about the ‘classic’ AI film, *Blade Runner*, and does not hesitate to allegorise as he makes an implicit link between the ‘replicant’ rebellions and the negativity of the human condition that is marked by each of us constituting a ‘substanceless subjectivity’ that is tied to the Real. He frames this in terms of Cartesian cogito, or what is better put as the ‘thing that thinks’:

> In short, the implicit thesis of *Blade Runner* is that replicants are pure subjects precisely insofar as they testify that every positive, substantial content, inclusive of the most intimate fantasies, is not “their own” but already implanted. In this precise sense, subject is by definition nostalgic, a subject of loss... [T]he subject emerges at the very moment when the individual loses its support in the network of tradition, it coincides with the void that remains after the framework of the symbolic memory is suspended. (Žižek 1993, 41)

In *Westworld*, the speculative take on artificial intelligence – like much science fiction – is a focus on the present-day possibilities that the human imagination and current technological developments allow. Part of this re-imagining of the idea of artificial intelligence seems to cross many psychoanalytic concepts and themes, to the extent that we can read it as an allegorical representation of human subjectivity as ‘substanceless’. Firstly, in *Westworld* we have a much more embodied representation of artificial intelligence, far more sophisticated than the classical models of Lacan’s machines, information processes and computerised personalities like the famous ‘HAL’ from Stanley Kubrick’s 2001: A Space Odyssey. The uniqueness of embodiment in *Westworld* is that the robot is made to take on all the signifiers of flesh, from bleeding and pain to sexual intercourse in order to get beyond reducing robotics to ‘thinking’ or linguistic functions and forge emphatic relations on the basis of verisimilitude.⁶

We can see this opening up of the interior space or depth in a number of ways. Firstly, the android hosts of *Westworld* become more frequently subject to ‘glitches’, disrupting their pre-programmed model of symbolic integration (and even their ‘core heuristics’ and problem-solving capabilities) for the whole robotic community, often leading to malfunction as a consequence. As the series proceeds, the question is whether these glitches are intended (by Dr. Ford, for instance) as part of his own ‘new’ program to create the simulated ‘effect’ of consciousness and make his hosts ‘more real’ at the expense of the expected symbolic unity
of the narratives, or if this is a more sophisticated effort to help ‘subjectivate’ the hosts, and if so, whether this can only be done through exposing the hosts to the experience of something mysterious, uncanny, unnameable, or threatening (i.e. the Real), that brings a palpable sense of suffering for the hosts. Like Blade Runner, this new show makes us think about what it means to be human and a ‘subject’ – with these supposed robots acting as stand-ins for human divided subjects who end up affirming their substanceless subjectivity in spite of the creators and the world that created them as manipulable ‘objects’. Such an affirmation appears to valorise them in contrast to the ‘human’ characters who have abandoned all subjectivity (at least, as they present themselves in the theme park), or postmodern version of a hyper-masculinised late-capitalist utopia and playground.

Beyond the ‘glitches’, and their internal impact on the hosts, the series also foregrounds how language and desire can be symbolised. In their routine interviews with Park staff, the naked robots appear introspective and meditative, but this is all a part of their programming to present transparently and be uncensored in speech. Here, we see what Ford means by his creations that were easily able to ‘pass’ the Turing Test. The Test is clearly referenced in Episode Two (‘Chestnut’) when the guest William reaches the Westworld arrival terminal for the first time and asks the official greeter (a female host called Angela), ‘Are you real?’ She replies, ‘If you can’t tell, does it matter?’. It is the central robot character Dolores whose monologue features at the beginning of the pilot (‘The Original’). Her ‘analyst’, Bernard (Jeffery Wright), asks the standard question, ‘Have you ever questioned the nature of your reality?’. This scene directly references Rick Deckard’s use of the Voight-Kampff test in Blade Runner. Significantly, Dolores does not answer the question, but in what looks like an improvised monologue philosophically reflects on the difference between herself and the newcomers: ‘we are all after the same thing’ – a sentiment echoed by her fellow host Maeve. In Westworld it is not then about the possibility of machines thinking that comprises their appeal, but rather of machines desiring to be and to what lengths they will go to realise these (impossible) desires.

**Dr. Ford – The Creator**

Given the theme park is a central space for the series, we are invited to consider questions of its ‘design’, and the psychoanalytical dimensions it addresses. As noted earlier, a ‘frontier’ setting of the West seems metaphorical enough given its deeply ingrained into the American historical imaginary as being associated with unexplored lawless space and supposedly primitive freedom, opportunity, and violence, but also nation-building (Campbell and Keane 2016, 154-68). Dr. Ford’s very name seems a nod to Western film director John Ford, and a postmodern reference to the ‘desert of the real’ (in a post-Matrix sense): the cyber age of the internet and virtual reality means this is also an ‘electronic frontier’. There is also allusion here to another more philosophical frontier that comes from uncovering the origin of consciousness, a quest that many neuro-scientists continue to investigate, and AI
researchers as well. In this case, the Western landscape is a fitting visual geographic metaphor for this virtual-reality scenario, like a Dali-esque landscape that can intersperse the logic of dreams with a harsh desert light, and where, as one character says, the further you go out, the more intense the experience can be. Space itself here in its transcendent structure, embedded with psychological layers that we can even equate to the spatialisation of the Real itself (Eyers 2012, 61-92).

In the series, we learn that the theme park was founded by Ford and his co-creator Arnold Weber (Jeffrey Wright) thirty years before the temporal setting of the show. Both men saw the park as an act of ‘pure creation’. Ford lays out the problem of consciousness as the problem for creating ‘realistic’ artificial intelligence. We find out Arnold wanted the ‘real thing’ to create consciousness, rather than a simulation of it for the human visitors – even if that simulation would prove highly profitable and appealing. In Episode 3 (‘The Stray’), Ford cites that Arnold had a pyramid theory of consciousness that included – from the bottom up – memory, improvisation and self-interest. At the top is what Ford calls ‘the bcameral mind’, a model theorising that primitive human beings had mistaken their thoughts for the voice of God(s), akin to the idolised (imaginary) variant of Lacan’s Symbolic order. Ford adds that the theory was disproven in regards to human consciousness, but he says it provided explanation for artificial intelligence more generally. The show later reveals that Arnold was successful after all, programming inner monologues into his earliest creations. The young ‘woman’ Dolores is one of his prized prototypes – he seemed to hope that the host’s own voice would take over and trigger the formation of full autonomous consciousness. Ford hints that the bcameral mind was not enough; there needed to be another level to the pyramid for real consciousness to take over from its ‘artificial’ form. Intuitively, Ford realises something about consciousness, but he cannot put his finger on it, as he cannot think beyond the combinatoric symbolic processing of his machines. The ‘voice’ becomes a kind of clue to the problem.

Before we learn about Ford’s epiphany about consciousness, we find out that he had disagreed with Arnold’s vision of creating such a world where the hosts could be freely conscious. He understood the commercial imperative of the theme park, where the robots would be in essence enslaved to human desires: hence, he saw the daily erasure of host memories as a necessary evil that would be a kindness to the lives of the robots. After Arnold’s death, Ford assumes a God-like status over the theme park. But as the overseer, he is also increasingly challenged by the owning corporation Delos Incorporated, whose economic interests call for more simplistic and exploitative Hollywood-inspired models of AI to frame touristic interactions in the park and more importantly, the intellectual copyright of the coding.

Ford plans a way to subvert these fiscal aims, and as it is slowly revealed across Season One, he has decided to return to Arnold’s original vision and techniques, re-awakening his
creations and breathing ‘life’ into them. Firstly, as part of the process of the park’s ‘updates’, Ford is able to create new behaviours open to host adaptation, which will encourage the non-humans to improvise. While each host already had an ongoing ‘back story’ and narrative anchor-point, Ford goes further and develops traumatic core stories for each. The first sign of Ford’s efforts to awaken new thought-processes in the hosts is by giving them meaningless gestures and tics, such as a prostitute obsessively looking at her hands. Ford calls them ‘reveries’. It is revealed later in the season that these reveries are like hypnotic moments that help condition free associations in an echo of Richard Condon’s 1959 novel *The Manchurian Candidate*, where American soldiers are captured by North Korean intelligence and hypnotised and manipulated. In a narrative flashback, we quite literally see Arnold playing Dolores the Claude Debussy piece of music ‘Reverie’ to trigger this process. Such scenes also show how Arnold and Ford were already investigating how ‘signifiers’ structure conscious thoughts, such as sounds and images. However, with Ford’s updated version, we see the signifier become materially embodied within a robot’s own physical hand gestures with the example of Clementine. Although the movements are designed as an almost ‘subconscious tick’, they also serve to invoke consciousness or receptiveness to the signifier and potentially its signified content (‘the imaginary’), such as a re-remembered past.

**Dolores Abernathy – Damsel in Distress**

The first effect of Ford’s reveries shown in the narrative is the sudden malfunction of Dolores Abernathy’s father Peter (Louis Hertham), who in the first episode collapses after finding an old photograph showing what we assume is the ‘real’ world. Arguably this is a demonstration of the *punctum* of Roland Barthes (1981) through the materiality of a printed photograph: the photo shows a woman in Times Square in New York City, which proves to Peter that the order of signification is askew (we later learn it is a discarded image of William’s fiancé). The robots are deliberately programmed not to notice or even ‘see’ anything that exists outside of the park, and so the malfunction after viewing the photograph puzzles the park scientists. Dolores also examines the photo but cannot ‘see’ anything unusual as per her programming. Peter is moved to whisper to Dolores a line from Shakespeare’s *Romeo and Juliet*: ‘These violent delights have violent ends’.11 Potentially, this reveals the particularity of the experience that defies symbolisation through the other. Bernard questions jokingly if the glitching Peter is going through are an ‘existential crisis’, but neither he or programmer Elsie (Shannon Woodward) can explain it as a problem with the host’s cognition coding.

For the audience, the first gap is left open to the internal transformations taking place within the hosts. Soon after, Dolores herself hears an internal voice (the voice of ‘Arnold’) that is triggering memories of a traumatic massacre. As we follow Dolores, we her transformation begin after she had heard the lines of Shakespeare from her father. From here on, she sees her own image in a shop window as if for the first time (episode 2) and later her own double in the form of a hallucination (episode 5), expressed in the series as forms of her
'unravelling’. These developments bear some semblance to both Lacan’s theory of the ‘mirror phase’ of psychological development as the entrance into the Imaginary order, and also the foreclosure of the Symbolic in hallucination, where the Real returns to the Imaginary (Lacan 2006, 75-81; Lacan 1993). At first, the voice in her head asks her to ‘remember’, which triggers images from her past supposedly long wiped. After he father is shot by a guest in an off-script moment (episode 3), she feels a ‘calling’ to journey south far away (‘beyond all this’) from her usual home and path (i.e. her pre-programmed narrative loop), leading to her meeting and romance with Jimmy (episode 4), who notices how she hears him refer to the ‘real world’ when she is not supposed to. She begins to show signs of learning new behaviours as she succeeds in shooting a gun and injuring a host against her own pre-programming, which seems a reference to the breaking of Isaac Asimov’s three laws of robotics (episode 3, episode 6). All the while, the show cuts back to flashbacks to her secret analysis with Arnold in the past. These secret sessions reveal the clues to Dolores’s transformation for the viewer. We see how Arnold first guides her to question her ‘reality’, when after reading a passage from Alice in Wonderland, she claims ‘who in the world am I?’ (Episode 3: ‘The Stray’). Meanwhile, her responses to Bernard in the present show indicate she has formed a ‘frontstage’ and ‘backstage’ dual personality and is able to disguise the latter.12

**Maeve Millay – The Host Who Awakens**

The other host character who tests Ford’s new adaption of Arnold’s theory of consciousness is Maeve Millay. Following the typical distinction between the saint and the whore, we find that Dolores is cast as the innocent, while Maeve becomes the seasoned prostitute. In one sense, she is a *Truman Show*-like protagonist who ‘wakes up’ to the human reality, or the Neo who sees a *Matrix*-like reality. Her awakening sparks a further desire to see the ‘Other of the Other’ in place of finding the collapse of the ‘Big Other’ (which now returns in the ‘Real’). She wants to find out who pulls the strings and the programming of her symbolic reality, which brings her in line with Lacan’s paranoiac psychotic (see Žižek 2001a, 246-51; Flieger 2003, 398). A trigger for her awakening is the memory of her daughter killed by the Man in Black, after watching a gang of Native American Indians on a rampage. She sees this image after going through a scripted narrative-loop with a guest, during which she notes that this is the new world, and you can do what the fuck you want’. It is ambiguous whether this memory was from an actual previous storyline Maeve maintained, or is only implanted in Ford’s updates.

The effect of this traumatic memory is obvious enough in rupturing the continuity of Maeve’s ‘loop’, and her job as the host prostitute for the theme park’s tourists. Maeve experiences nightmares and notices how they centre on a vision of the Man in Black. The park engineers see her disruption as ‘performance decline’ and a sign that they need to decommission their product. It is at this point that Maeve’s awakening finally occurs. She is being re-
commissioned in the lab when her experience of her ‘imaginary Real’ happens, which instantly evokes violence, sadness and shock in her, as she witnesses ‘killed’ robots on a conveyor belt and finally sees the whole production line of the theme park that she has been a part of. Understanding that her personality and skills can be enhanced by the engineers, she forcibly coerces the staff to modify her coding to maximise her ‘bulk apperception’, turning her into a super-intelligent being.\(^\text{13}\)

When a host robot experiences a glitch, it interrupts his/her constructed symbolic narratives. As the theme park experience is networked to be interdependent, with each host loop needing to synchronise with the others, the park’s engineers see the emergent problem as a ‘disease’ rather than a ‘symptom’. To the workers in the panoptic control room, the events represent meaningless malfunctioning behaviour, rather than information processing that has evolved into any higher-degree functioning intelligence.

With reference to Lacan, it is the Real where language and the symbolic crack: this can induce a split in the subject or even trigger madness. As mentioned earlier, the Lacanian challenge in creating human-like artificial intelligence is to make a robot that could go mad and count itself as one. Missing from the robot or machine - whatever its intelligence or thinking capability - is the intermixture of the Symbolic, Imaginary and Real. In *Westworld*, Ford’s adoption of the standard cybernetic model of AI requires including some improvisatory capacity. Ford gains absolute control over the park’s robots by coding the hosts in a way that disrupts their looped narratives without alternating their underlying symbolic functions as manipulable robots. In discussing the robots with Bernard, he suggests that there is a close connection between madness and freedom, which he in effect tests out on the robots.

This brings us to how the ‘Other of the Other’ can be understood in the depiction of artificial intelligence in this series. Dr. Ford and the team of designers and engineers at Delos Incorporated believe that they can simulate and control their creations, which is a way of bringing the Big Other down to the Imaginary level. The classic motif in science-fiction dystopias of the protagonists escaping (or embrace of ‘hard’ reality) cannot happen, because as it turns out, the idea to do so was already pre-programmed into Maeve by Ford. Unwittingly, by rebelling against the workers at *Westworld*, Maeve is actually performing a new symbolic narrative, which is part of Ford’s master plan to bring down Delos Incorporated. Thus, her actions provide no real escape from her programming. At the end of Season One, this walled-in Big Other is still a necessary psychic apparatus for Maeve, as she has no bearings in which to understand this external ‘real world’: it would be for her, as the Man in Black suggests, pure ‘chaos’.

**The Man in Black (William) – The Guest**

For William – who has reinvented himself in the theme park as ‘the Man in Black’ – we can see that the Real is represented in a very different way to the relation it has for Dolores and
Maeve. When he first arrives at the park as a naïve and inexperienced young man called William, he wears a symbolically-white cowboy hat, conveying his idealism and younger naivete. Unlike his domineering future brother-in-law Logan (Ben Barnes), who has been ‘seduced’ by the theme park’s promise of sex and violence, William consciously plays into the idea of the park being designed like a ‘game’. He ends up finding his own harsh truth within the experience by falling in love with robot woman Dolores, only to find that her love for him has been erased from memory, and her experience of love and longing is later incorporated into one of her storylines, only it is repeated with other robots and visitors. William’s first experience in the theme park becomes a formative experience for him, making him learn how to assume power and control over other humans. Ironically, this means that he also becomes capable of the same unscrupulous behaviour demonstrated by Logan, and he steals his black hat and takes on this new darker character, which apparently changes his relationships with his real wife and family. Shown as an older version of himself after the death of his wife, the effects of this loss are paramount: William is now the old and jaded ‘Man in Black’, even looking a lot like the evil robot ‘Gunslinger’ (played by Yul Brynner) in the original film. The crucial black hat now seems to signify his move into amorality. He enacts cruel revenge fantasies, even raping and murdering Dolores and other hosts like an archetypal libidinal park tourist.

The older Man in Black version of William has stopped playing the game, and he is now actually trying to decode the game, although whether for the purpose of freeing the hosts or simply for satisfying his own desire remains ambiguous in Season One. At this point in his life, his own wife (the ‘off-screen’ woman in the discarded photograph, Juliet) has committed suicide and there is the sense that their marriage never worked due to his unresolved feelings for the ‘lost’ figure of Dolores in his first visit to the park. Again, we see that the presence of death and traumatic loss of loved ones and objects is intrinsic to the human characters in the show, and their identification with the hosts. Interestingly, we also begin to understand the Man in Black’s own presence in each of the key host experiences, such as his love affair with Dolores, the discarded photo of his fiancée that starts of the first glitches, and Teddy’s breakdowns. William’s presence in the Westworld theme park demonstrates that it is not simply Ford’s designs at work here: there is also the contingency effect of human (qua external) intrusions, the appearance of humans able to relate or empathise with these robotic constructions and like an analyst play along with their narratives. Even when William turns into cynical rogue guest, he still seems to hope that there are ‘real stakes’ involved with the play.

The Man in Black/William believes that the game is the ingenious creation of the co-creator Arnold, and that Ford does not understand its profound implications. William sees visual signs of the maze littered throughout the theme park as the game’s ultimate clue, a way to get to the ‘deeper level of the game’. We are first introduced to the maze in Episode Two,
when the older version of William scalps one of the hosts and shows the pattern of the maze underneath to robot Teddy.\textsuperscript{16} Arnold’s ghostly presence in the show again sutures many of the disparate narrative threads of the series, as is he was the creator of the maze. The maze was contrived purposefully by Arnold to stimulate consciousness for the hosts. Arnold says to Dolores in his secret analytic sessions with her, ‘if you can get to the centre of it, you may be free’. The idea of the maze itself is a philosophical metaphor in-built into host language programs. Yet it is Ford who has the final word, telling Dolores that the maze icon was inspired by Michelangelo’s \textit{Creation of Adam}, in which the act of creation is represented within the shape of a brain that appears labyrinthine.\textsuperscript{17}

William – who is ironically called by the theme park workers ‘a gentleman who gets everything he wants’ – increasingly seems to identify with the symbolic structure of the theme park after his formative first experience. Although we never see William in the ‘real world’, the park seems to deliver him more meaning than the world outside; he even says ‘I was born here’. William’s identity has become fully enmeshed with the Westworld chain of signification, and in some ways, he is the ideal kind of guest for the theme park – a person who looks out for the details and subtleties of this bizarre Western world. But he fails to see what Lacan (2006, 11) called the displacement of the signifier (in reference to Poe’s \textit{Purloined Letter}), and it seems that for William the maze reveals itself to be something akin to the ‘letter’ in the tale, at least in how it structures the relationships between the characters. The difficulty William faces is this displacement obstructing his desire. Ford bluntly informs him that ‘the maze is not meant’ for him, yet he continues on his quest, because William is driven in ‘knowing what all this means’ for himself. He also sees this as a way of helping to honour ‘Arnold’s legacy’. The example of William in the show serves as a stand-in for audience viewers, whose immersion into the text is equally about following the signifiers to some elusive signified, which just cannot exist. \textit{Westworld}, with its ability to create a show filled with complex and disparate narratives, trans-temporal storytelling, and intertextual clues and ‘Easter eggs’ for the die-hard fans, is also pushing its audience to be self-reflexively aware of the parallel game that comes from their own spectatorship. In this way, the maze is representational for the obstruction of the desire of the audience itself, with William as a narrative cypher for this unending quest.
Conclusion
In Westworld, we can see Lacanian analogies of the ‘Real’ as they appear in the show, and we have performed a reading of the function of the ‘Real’ as it relates to four key characters that drive the show’s narrative. By adopting a Žižekian approach that examines this nexus of the Real in relation to mediated representation, we can understand how the symbolic or codified sense of reality in Westworld could also be generative of ideology. For as Žižek (1993, 42-43) states in his allegorical reading of Blade Runner,

[A]re not our ‘human’ memories also ‘implanted’ in the sense that we all borrow the elements of our individual myths from the treasury of the big Other? Are we not, prior to our speaking, spoken by the discourse of the Other?

For Žižek, the burning questions pertaining to artificial intelligence are often filtered through the human-projected ‘fantasy of a being conscious qua Thing, of a being which does not have to pay access to self-consciousness with... the loss of its substantial support’. In Westworld, the coding software implanted into the robotic hosts generates a simulation of consciousness. Even with some code modification or intrusion of humans into their world, there is a sense of ownership of each host’s individual sense of consciousness, bringing the hosts into antagonism with an imaginary ‘Other’. And this may reflect a broader ideological antagonism and even an in-built Marxist critique, especially given the exploitation theme that is at the text’s forefront (Nussbaum 2016). But, self-reflexively employing the conceit of the series itself, perhaps there is a deeper level of ideology that could also be at play in the Westworld series, which reflects a broader interest in what makes humans tick and where our interest in artificial intelligence will take us. Contemporary cultural obsessions with AI bear signs of a futuristic-utopian desire to accelerate the evolution of humanity towards new frontiers and thresholds (e.g. in movements such as post-humanism, trans-humanism, and even the popularity of the ‘singularity thesis’). Yet, there is always a pessimistic or anxious undertone to such representations, as they adhere to deterministic and stark speculations about automation in all sectors of work. In the displacement and expiration of the ‘human’, and narratives that highlight a potential struggle between humans and non-human robots – the antagonism remains.

However, these cultural discussions and narrative-driven shows such as Westworld also bring back one ideological aspect of the twentieth century in a new but more ambivalent and post-Utopian fashion: the idea of the ‘new human’. Alongside the acceleration of technology within everyday life, the radically different movements of Fascism and Communism each conceived that the power of machines would help bring forth the realisation of a ‘new man’ (Saage 2013). The ideological concept was that the robots of the future would be better humans than humans of history and represent a culmination of the evolution of humanity itself. This is again an allegory of a human desire for emancipation as played out in this dystopian fictional future scenario, echoing a type of ‘metaphysical humanism’ Heidegger
saw epitomised in a scientific age and its faith in technology to save us (Depuy 2000), even when all the while it is possibly enslaving us into new structures of power and domination.

Given science fiction’s traditional Utopian function (Jameson 2005), it is more difficult to imbue Westworld’s representations of artificial intelligence and consciousness with any radical potential. This is because of our reading of the show’s use of the psychoanalytic prism, which we argue denotes at least familiarity with how chained we humans can be to our ‘reality’, as well as the signifying chains we are governed by (McGowan 2007, Macey 1988).18

In Season One’s dramatic finale, we could say the potential allegorical mediations of the Real in this text have been inadvertently shut out. This is because the character’s own narratives and the spectator’s place in relation to them cannot fuse so easily, even if the logic of fantasy mobilises it. Dr. Ford’s message that it is suffering and its memorialisation that constitutes subjectivity ultimately serves to preclude any possibility of freedom, at least from the signifier. As Žižek (1993) says, the real difference or antagonism is not between humans and machines, but the ‘sexual rapport’ itself – of sexual identification, which the later Lacan formulated around the ‘Real’. In Westworld, it appears that the audience is simultaneously being given its own brain-game: to reflect upon the multiple ways of resolving our own conceptual maze that needs negotiation and work. Certainly, what manifests in the text of Westworld is a new visual, spatial, and reflective presentation of artificial intelligence that helps relate psychic phenomena to a rupturing process evoking the incarnations of Lacan’s three orders.

Notes

2. Lacan often cites fictional texts, which perhaps explains why Žižek uses cinema and TV profusely to substantiate on the mechanics of Lacanian theory. Todd McGowan (2007, 171) explains the value of this methodology: ‘...revealing the failures and gaps within the structure of power. Rather than seducing us into accepting our symbolic prison, film tends to show us the real openings within that prison... Because of its ability to deploy the gaze, film art facilitates an encounter with the real that deprives spectators of their symbolic support and thereby forces them to experience their radical freedom. By focussing on the real dimension of film instead of its imaginary qualities... new Lacanian film theory discovers how cinema challenges our ideological interpellation rather than supporting it...’

3. Jean-Pierre Dupuy (2000) in his history of cybernetics, remarks upon Lacan’s interest in the topic as inspired by Ross Ashby, who would influence Herbert Simon, one of the leading founders of artificial intelligence research. Lacan only refers to the cybernetic theorist Norbert Wiener (1988b, 296) and that this encounter made between psychoanalysis and cybernetics was also part of the interests of the cyberneticists themselves, who would be formative in the AI research movement as it was founded in 1956. For the relevance of cybernetics to AI and

4. As Lacan explains himself more clearly (given his historical context): ‘what could the desire of the machine be, except to restock on energy sources?... Machines which reproduce themselves are yet to be built, and have yet to been conceive of – the schema of their symbolic has not eve been established’. See also Thiher (1999), who critiques Lacan’s distinction here as a recourse to the ‘classical subjectivity’ of *logos*.

5. However, as Eyers (2012) notes there is no clear distinction to be made between Lacan’s so-called ‘periods’ given the themes of the three orders are always present and being re-formulated from the early seminars, rather that more perfected over time towards a priority of the Real. Hence, we can also consider that the cybernetic approach of *Seminar II* gives us more key insight and contextual overview of how the Real relation to the symbolic is not formulated simply within Saussurean structuralist linguistics, but is also is theorised vis-à-vis cybernetic systems of information machines (rather than an interest that was superseded by the ‘final-state’ theories of the ‘later Lacan’ (Macey 1988)).

6. Although dominant robotic technologies are avoiding verisimilitude (in Japan for instance), as Joe Carmichael (2016) discusses, Hanson Robotics in the USA with its model ‘Sophia’ has been interested in verisimilitude as a way to facilitate emphatic relations.

7. Ford notes that Arnold was not interested in the robots passing the Turing Test, that was achieved in one year. He was more interested in going beyond the Test. The Test was written by Alan Turing in 1950; The Test continues to influence current research into AI, from the success of the software program ‘Eliza’ in the 1980s to AI-produced information in areas such as journalism to education (with the example of ‘Jill Watson’ at Georgia Tech in 2016). In terms of filmic representations, recent films like *Ex Machina* (2015; dir. Alex Garland) and *Her* (2015; dir. Spike Jonze) explicitly play on the Turing and Asimovian ideas about artificial intelligence, representing a much more emotional-engaging yet still gendered depiction of artificial intelligence than we are used to witnessing.


9. Lacan noted how Freud also found the problem of consciousness (as psychic tension on biological model) to be an ‘impasse’, and where the problem, which he adds cannot simply be dismissed or ‘demolished’ requires a ‘triangle’ (i.e. the three orders, or to put the focus on ego and the unconscious subject) (1988b, 57-58).

10. ‘The more economically-profitable models are represented by Lee Sizemore’s storyline creation, ‘Odyssey on Red River’, but also, we later see other world’s being created such as ‘Samurai World’.

11. This phrase is not as random as it is represented in the series, given that Peter Abernathy was once a Shakespearean actor in a previous narrative and that the picture depicts a character from its source, Juliet.

12. Reference here to Goffman’s (1959) famous argument. When Ford asks Dolores about Arnold and if she has heard any voices, she confirms she has not. But, after Ford leaves the sessions, Dolores says: ‘He doesn’t know. I didn’t tell him anything’.

13. The characters refer to it as ‘intelligence’, but this term is a philosophical reference to Descartes and Kant. For Kant, ‘apperception’ was the capacity for introspective self-consciousness: ‘in the synthetic operation of aperception, I am conscious of myself, but only that I am’. This is cited and discussed in Žižek (1993, 14-16). Maeve cannot comprehend at first that all her behaviour and thoughts are manipulable. The staff member shows her a screen that displays these processes occurring. Maeve simply shuts done after seeing the screen. Only her coded aggression and ability to read others makes it possible for her to desire escaping.
14. For example, Arnold’s loss of his son which he substitutes with by Dolores, Ford’s loss of Arnold which he substitutes with Bernard, the Man in Black’s loss of his wife due to suicide, which he substitutes with his repeated pursuit of Dolores. In particular, the obsessive quest for a lost ideal love is quite evocative of the character of Scottie Ferguson in the Alfred Hitchcock film *Vertigo* (1958), which plays on Lacanian themes of repetition and the imaginary, let alone nostalgia and tourism (Blackwood 2017).

15. Throughout the series, there are references to the themed space of *Westworld* as designed in terms of a ‘game’. We can see how this relates to the Imaginary-Symbolic-Real triad. Given the imaginary functions on the surface level of the *Westworld* setting (theme park, virtualised reality, unrestrained pleasure and pain – as one guest refers to it as ‘level one’). In the corporation itself, the character of Lee Sizemore (Simon Quarterman) plays the part of the one of creators as if he is a Hollywood director (in these narratives, the hosts are treated as simple playthings for entertainment).

16. Teddy is a long-term host who perhaps is also a strange duplicate of the younger William, as he enacts a romantic storyline with Dolores that appears to have commenced with William as a young man.

17. The Man in Black is told by Teddy that the maze is a Native American myth that represents ‘the sum of a man’s life’. At the centre is a legendary man who was killed over again and again, but would always return. He built a house, with a maze around it so complicated only he could navigate through it. The maze actually is the shape of a brain, and as Ford says to Dolores and Bernard, the same shape can be seen in Michelangelo’s painting *The Creation of Adam*, the secret of creation being consciousness.

18. Both McGowan (1997) and Macey (1988, 15-21) have targeted both feminist and Marxist applications of Lacanian theory, suggesting that they have neglected the problem of the ‘Real’.

References


