Recognition, collaboration and community: science fiction representations of robot carers in *Robot & Frank*, *Big Hero 6* and *Humans*

Yugin Teo

Faculty of Media & Communication, Bournemouth University, Poole, UK

Correspondence to
Dr Yugin Teo, Faculty of Media & Communication, Bournemouth University, Poole BH12 5BB, UK; yteo@bournemouth.ac.uk

ABSTRACT

In the 2010s, a small number of science fiction films and television series exploring the theme of the robot carer and how humans respond to them were released. This paper explores three works in this regard: the films *Robot & Frank* (dir. Jake Schreier, USA 2012), *Big Hero 6* (dir. Don Hall/Chris Williams, USA 2014) and the television series *Humans* (UK/USA, Channel 4/AMC, 2015–2018). Examining these works with some of the ethical issues currently being discussed in the use of robot technology in care work, this paper demonstrates how they align themselves with, but also challenge some of these ideas, and ultimately direct viewers to consider their own expectations of personalised healthcare. The essay begins by examining the fears of the care industry deploying robots to replace the work of human carers, followed by a discussion of the effectiveness of robots as carers as depicted in these fictional representations, and the final section considers the social environment that these robot carers are situated in, and how the robots become a reflection of human lives and a repository of memories of affective relations. These texts suggest alternate ways of thinking about human–robot interactions and care work, advocating for a more mutually dependent and reciprocal working relationship that might lead to a better quality of care.

INTRODUCTION

The robot companion, assistant or sidekick in popular culture has a benevolent history in science fiction film and television, with reliable, timely and oftentimes quirky characteristics. These companions and assistants include astromech droids such as R2-D2 and BB-8 from the *Star Wars* film franchise, the drones in *Silent Running* (dir. Douglas Trumbull, 1972 US), Marvin the Paranoid Android from *The Hitchhiker's Guide to the Galaxy* (dir. Garth Jennings, 2005 US/UK) feature film adaptation, or the android Dorian in the short-lived *Almost Human* (US, Fox, 2013–2014) television crime series. In the 2010s, there have been a small number of science fiction films and television series that explore the possibilities and implications of robots in caring roles. These films can be viewed as fictional responses to the increasing presence of robots within healthcare, and specifically in care work. According to Sharkey and Sharkey, the number of elderly people in the population of Japan, Europe and the USA 'is beginning to overtake the numbers of young people' who are able to care for them.[1] The field of robotics and artificial intelligence (AI)

increasingly becomes an area of serious consideration by governments for providing care for the elderly. As such, the role of robots in care work has been widely debated from ethical, psychological and technological viewpoints. While robots, and assistive robots in particular, are being used in care work, AI technology has not yet reached the level of development where robots can be considered 'real' companions with people.[2] These robots can only demonstrate limited conversational ability despite being interactive and showing simulated emotions, and they are as yet unable to experience human empathy.

The field of human–robot interaction (HRI), while developing quickly, has yet to reach the stage where a robot is fully capable of observing and interpreting human facial expressions and vocal deliveries. This is where fiction, and in particular science fiction film and television, can be a compelling trigger for discussion within this area of robotics and healthcare, extrapolating possibilities from current developments into near future settings. Science fiction has been noted by scientific communities for generating largely negative perspectives and dystopian futures regarding robots, including the familiar tropes of robot uprisings and threats to humankind.[3 4] The films and television series in question, *Robot & Frank* (*RF*), *Big Hero 6* (*BH6*) and *Humans*, continue the development of the science fiction robot assistant tradition while engaging with public discussions of robots in care work, presenting interesting and notable depictions of synthetic carers. *RF*, in particular, has been cited as one of the exceptions in its positive depiction of HRI on screen.[5]

RF, set sometime in the near future, tells the story of Frank (Frank Langella) a former jewel thief who has been living with dementia, and of his relationship with his robot carer and companion bought for him by his son Hunter (James Marsden). Frank soon discovers that he can manipulate Robot (performed by Rachael Ma and voiced by Peter Sarsgaard) into becoming his partner in crime. As Frank's confidence increases, he tries to get the attention of Jennifer (Susan Sarandon) who works in the local library. One of the subplots concerning Frank's memory loss is the ongoing tension between him and his son, and this threatens to alienate him from his family. Through his adventures with Robot, he gradually becomes reacquainted and reconciled with his family. (When referring to Robot in this paper, I have followed previously published research

papers on the film in using the pronoun 'he'). *BH6* is a Disney animated action adventure film, loosely adapted from a Marvel comic book series of the same name. In the fictional city of San Fransokyo, Hiro Hamada (voiced by Ryan Potter) is a teenage robotics prodigy whose older brother Tadashi (voiced by Daniel Henney) has developed a ground-breaking healthcare robot companion called Baymax (voiced by Scott Adsit). (In the comic book series, Baymax is a bodyguard.) When Tadashi is killed in a mysterious fire, Hiro, together with Baymax and Tadashi's friends and colleagues from the San Fransokyo Institute of Technology, join forces to investigate what really happened. Chris Atkeson from Carnegie Mellon's Robotics Institute, working in the field of soft robotics, acted as a consultant for the film and the design of Baymax.[6] *BH6* was critically and commercially well received and won the 2015 Academy Award for Best Animated Feature Film, helping to bring the topic of assistive robots in care to public consciousness in a unique way. *Humans* is a television series that was first broadcast in 2015 on Channel 4 in the UK and AMC in the USA, adapted from the successful Swedish science fiction television series *Real Humans*. *Humans* depicts a fictional near future where androids called 'synths' are used as helpers, carers, workers and servants. The series focuses on the Hawkins family and the arrival of a synth called Anita (Gemma Chan) into their lives. Anita (later on revealed as Mia) is different from other synths in that she is already sentient and was formerly part of a group of sentient synths that are being hunted down. The first season was successful both critically and commercially, and a second and a third season subsequently followed. This paper will be focusing solely on season 1 as the ethical discussions surrounding robot carers are mostly limited to this initial season.

This paper focuses on how these three fictional representations of robot carers in popular film and television in the 2010s explore contemporary issues, concerns and possibilities regarding robot companions and their role in the future of care work. The examination of these films is partly framed through the sociology of expectations, a theoretical approach within science and technology that analyses change, and an alternative conceptualisation of the fictive novum from science fiction criticism.

The sociology of expectations, as conceptualised by Borup, Brown, Konrad and Van Lente, highlights the crucial role played by expectations on the production and public implementation of technology.[7] Narrative is one of the key factors within expectation theory, where shared fantasies concerning particular emerging technologies can theoretically lead to a 'rhetorical vision' of hope, driving forward phases of implementation and production.[8] The significance of technological expectations is dependent on the strength of the 'vision' and 'future-oriented abstractions' accompanying the novel and emerging technology in question.[9] The visions of the future depicted in the narratives examined in this paper represent the 'expectations, hopes, fears and promises' for assistive care technologies.[10] However, as these visions are depicted through science fictional texts, the elements of extrapolation and critical enquiry are also present and will be examined later on. The discussion is composed of two parts: the first part is concerned with the fears and expectations concerning robots in healthcare and care work, including replacement fears and ethical concerns as to their effectiveness as carers; the second part critically examines the robots' social environments as proposed by these narratives. Throughout my analysis of the case studies, I will be using a unique interdisciplinary approach that combines material from bioethics, assistive technology, nursing care, HRI, science fiction criticism and expectation studies.

FEARS AND EXPECTATIONS
Replacement fears

Amanda and Noel Sharkey emphasise the importance of embodiment when considering how robots are perceived, as they are able to perform tasks, follow people around and interact with people, unlike more computational devices.[11] This sense of embodiment is even more important when robots take on life-like appearances. The physical presence of a robot in a shared space can lead to the person being cared for feeling vulnerable and experiencing a loss of privacy. The robots in *RF*, *BH6* and *Humans* differ in their physical appearances but all have a unique presence within the mise-en-scène of each text. While the synths in *Humans* are designed to look and behave exactly like humans, both Robot in *RF* and Baymax in *BH6* are designed more practically and their physical appearances reflect this. Robot in *RF* resembles the ASIMO robot developed by Honda with its white utilitarian outer body structure and large black visor, and is the most realistic depiction of a robot companion of the works being examined. Baymax on the other hand is tall and has its mechanised parts encased in a durable and inflatable balloon-like white outer skin. Neither robot is designed to look human, and therefore, cannot be mistaken as human. The design of the fictional robots in *RF* and *BH6* are aligned with research on robots in healthcare where human-like appearances could cause more problems in user interaction. For example, in a paper examining a range of concerns regarding the use of robots in therapy for children with autism spectrum disorder, *Coeckelbergh et al* discuss the concerns from some survey respondents of robots appearing too human-like and being perceived by children 'as a friend'.[12] Robots that look like animals seem to be a 'safer' option, partly because human–pet relationships (as opposed to human–robot ones) are a more acceptable replacement for human–human relationships. This appears to be borne out of fears of robots eventually replacing the work of human therapists, but also a concern that a child might become too attached to a robot if it appeared human-like, causing distress when the robot is no longer present. Therefore, one of the key reasons for a robot's physical difference from humans is an intended design philosophy in order that they can be easily identified as robot carers, and not as life-like companions. The synths in *Humans*, on the other hand, are designed to look human and mimic human behaviour, and apart from the colour and appearance of their pupils, are difficult to distinguish from humans in a crowd. They are designed to be incorporated seamlessly into the household as synthetic helpers or carers. The physical appearance of robots can influence human perception of robot use in society, and this is very much linked with public perceptions of robotics, often informed through the media, and especially through popular science fiction, where robots are portrayed as unpredictable and pose a danger to humans. In Christoph Bartneck's paper on the representation of robots in the theatre and the media in general, he discusses the frequently used trope of the struggle between humans and robots in science fiction narratives.[13] He argues that in order for a story involving robots to be successful, tension needs to be generated within the narrative, and the human–robot conflict has arguably been a popular device for this reason, particularly in Western culture. Bartneck adds that any research on HRI needs to recognise the importance and influence of fictional media representations of robots, and the 'elicitation of the associated fears' of robots threatening the existence of humans.[14] He emphasises that current users of robots will be aware of such narratives, and any work on HRI should make it clear that the robot is a machine with specified abilities and intentions.

A key concern regarding the utilisation of robot technology within healthcare, and one perhaps associated with the issue of robots appearing too life-like, is the fear of replacement. Results from surveys taken of stakeholders within healthcare demonstrate that people are concerned about robots replacing qualified healthcare professionals such as nurses and therapists.[15] Coeckelbergh argues from a phenomenological viewpoint that human carers need to be perceived by those in care as being in charge and in control, and only in such settings will assistive robots be more comfortably accepted.[16] There is a case being made here for visibility and presence concerning robot helpers. Healthcare in general is seen as a 'human' activity and practice, and the therapist for that reason should be human, and while robots could be 'part of the therapeutic process', they should not be the therapist.[17] The therapist should remain in the room and supervise (and intervene as necessary). Coeckelbergh *et al* suggest that any 'perceived' replacement of the therapist by the robot is to be avoided. Coeckelbergh also makes the point that there is a preference for robots in care to be seen assisting the qualified human healthcare professional, rather than demonstrating agency in making medical decisions and in performing tasks usually assigned to humans.[18] He warns that as care work becomes more machine dependent and individualised, there is a danger of losing sight of the phenomenological aspect of human care, which is about human contact. This means not only physical care, but psychological and relational care as well. What Coeckelbergh refers to here is the concept of the social, and as social beings part of the process of receiving care is through having relationships with others, making the case for care work as an essentially human-related activity that should not be an area in which robots are deployed in decision-making roles. Coeckelbergh reiterates his call for a reconsideration of the future direction of care work, where more suitable roles (other than as robot carers) ought to be created for machines being utilised in this area of healthcare.

While premeditated interventions, such as deploying rules that ensure a human therapist or carer is always visible and in supervision of care robots at all times, make logical sense from the point of view of offering reassurance to users, they run the risk of furthering the narrative of the robot as a possible danger to human life. This could also have implications for any positive medical or care interventions made by robots if the human therapist or carer exercises too much control, negating any benefits of having assistive robots in healthcare. Baymax, designed as a 'personal healthcare companion' in *BH6*, contradicts this call for non-agency in robots, and is fully capable of making internal bodily examinations, medical diagnoses and providing light-touch medical interventions without seeking the subject's permission. Hiro, whose patience gets tested by Baymax's attempts at intervention, eventually accepts Baymax's characteristics. Baymax's non-threatening design and rounded appearance, inspired by Atkeson's work on soft robotics mentioned earlier, certainly encourages Hiro's positive perception of the robot. However, it is the way in which Baymax appears to strike a satisfactory balance between giving autonomy to the human subject while providing protection and assistance, as suggested by Sharkey and Sharkey,[19] that earns Hiro's respect. The autonomy exhibited by Baymax, Robot in *RF* and the synths in *Humans* function as both fictive novum (typically found in science fiction narratives)[20] and emergent technology. The novum is typically an innovation extrapolated through the perceived 'cognitive effect' of believable technological advancements.[21] In the case of these narratives, the novum of autonomy in care robots is a development from existing assistive technologies already used within

aspects of healthcare, and therefore the autonomy exhibited by these robots is both science fiction and anticipated technology. Baymax, Robot and the synths do not necessarily point toward a future occupied by autonomous robots, but they do reflect back on the present and on the continuing developments in assistive robot technologies, as well as highlighting the ethical concerns regarding the utilisation of robots in distinctively human-related tasks, such as care work and other aspects of healthcare.

Effectiveness as carers

One of the challenges for the future application of robots in care work will be in assessing how effective and safe these robots can be in real-life operational scenarios, and how technological advancement, commercial manufacturing processes and marketing strategies can work in tandem to create robots that serve the patient beneficially. While both Robot in *RF* and Baymax in *BH6* initially exhibit traits of responsibility and competent levels of care towards Frank and Hiro respectively, end-user autonomy soon begins to take over as both Frank and Hiro hijack their robots' abilities for their own purposes. In Frank's case, it is to revisit his past vocation as a jewel thief, and for Hiro it is to investigate the cause of his brother's death and find the perpetrator.

The first and second of Isaac Asimov's Three Laws of Robotics state that:

1. A robot may not injure a human being, or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with/ the first law.[22]

It is useful to note that both Robot and Baymax, through their actions, exhibit an adherence to Asimov's first and second Laws. While both robots begin by allowing their users to use them for their own personal quests, they also disobey their users' orders later on and play key roles in saving them from impending disaster at pivotal moments. Towards the end of *RF*, as the Sheriff (Jeremy Sisto) and his men close in on Frank for stealing jewels from the home of Jake (Jeremy Strong), a wealthy developer in charge of the local library redevelopment project, Robot makes an intervention. He impresses on Frank the need to 'wipe (Robot's own) memory'. He reminds Frank once again that he is 'not a person', but 'just an advanced simulation'. Robot, being fully aware that Frank's return to crime is going get him arrested, advises Frank to erase his memory files so that there will be no evidence of his recent crimes. This, however, would lead to shutting down Robot permanently, effectively erasing the ghost of the robot that Frank had been developing a friendship with throughout the film's narrative. Frank eventually comes round to the realisation that there is no other option, and follows Robot's instruction to turn him off and erase his memory. This act of self-sacrifice fulfils Asimov's first two Laws, protecting the user from harm (imprisonment) and disobeying the user's orders as it would also lead to harm. While Robot was mostly compliant with playing along and being Frank's accomplice in the heist, the overriding directive of care and attentiveness to Frank's welfare is still paramount. This crucial scene will be examined in more detail in the latter part of this essay.

In *BH6*, Baymax shows clear hesitation and ultimately refuses to use violence when Hiro commands him to destroy his adversary Callaghan (James Cromwell) for causing his brother's death, as this would violate the programming originally created by his brother. Hiro forcefully replaces his brother's healthcare chip installed in Baymax with one of his own in order to get Baymax to hunt down Callaghan. Back in his garage later,

when Hiro tries to replace his brother's healthcare chip again, Baymax refuses to open the access port for Hiro, repeating to Hiro his original purpose to 'heal the sick and injured' and not to be a fighting machine. (This scene will also be examined more closely later on.) At the film's climax, Baymax, like Robot in *RF*, sacrifices himself so that Hiro will be safe. Hiro and Baymax risk their lives to enter the teleportation portal to rescue Abigail Callaghan, but soon find themselves stranded in the portal with no way of escaping before it closes. Baymax's thrusters are inoperative, leaving the one option of using his armour's rocket arm to fire Hiro and an unconscious Abigail back through the portal's opening to safety. This would leave Baymax behind as the portal collapses in on itself. Baymax's statement of care to Hiro, 'you are my patient… your health is my only concern' falls on deaf ears as Hiro initially refuses to accept his proposal of using his rocket arm. Baymax eventually convinces a reluctant Hiro to accept, and the poignant moment is encapsulated by Hiro's utterance of the task completion statement 'I am satisfied with my care' to Baymax. The diegetic sound of this scene becomes muted and the scene locks on to Baymax's point of view, lingering on a close up shot of Hiro's face as the rocket arm fires from Baymax's armour and transports Hiro and Abigail through the portal's opening. Through acts of self-sacrifice, and in also refusing to harm other humans when ordered to do so, Baymax, like Robot above, fulfils Asimov's first two Laws, reiterating the assistant robot's long-established directive of placing the care of all human life above its own.

Despite having robots that demonstrate a form of agency described by Coeckelbergh as carrying out tasks usually assigned to humans,[23] neither *RF* nor *BH6* exhibit the more familiar science fiction narrative trope of robot uprisings. In fact, the robots in both films are generally portrayed as responsible and loyal carers. Both films seem to encourage the possibility of a future where robots are not only effective sole carers, but are capable of what Dean *et al* describe as 'extending the range and quality of care that humans can provide'.[24] It is worth examining how this possibility is conveyed in both films above. Robert Sparrow, in his article on the use of robots in aged care, theorises two key components of human welfare (and by extension, qualities that relate to aged care): recognition and respect.[25] Sparrow argues that robots are machines that lack both the 'interiority' and the capacity to enter into 'sets of affective relations' that are necessary for relationships between human beings. As a result of this, robot carers working within the parameters of a care home, for example, will not be able to provide a communal environment where both recognition and respect is genuinely experienced between users and robots. There seems to be an attempt to address these ethical and philosophical issues in both *RF* and *BH6*, as some form of recognition and respect occurs between user and robot (to different degrees). This might explain the positive (or at least sympathetic) depiction of HRI in contrast to the more familiar trope of robots rebelling against human owners. Narrative framing devices are used in both films to convey this sense of an affective relationship (as described by Sparrow) between robot and user: in *RF*, Frank's dementia forms the thematic basis for the narrative as his exploits with Robot gradually reveal his fractured relationship with members of his immediate family; in *BH6*, Tadashi's death lingers as an unsolved mystery and a focal point of Hiro's sense of loss, paving the way for Baymax's mission to provide Hiro with the care he needs for emotional and physical recovery. Frank's refusal to accept Robot as just an assistive machine goes against Robot's function as his carer. Robot reminds Frank that '(his) health supersedes [Robot's] other directives' as they stake out Jake's home for the

heist. Robot reminds Frank of his identity as a robot, one who is not alive, and who does not object if his memories are erased. Robot's character in the film is a foil to Frank's, a machine that has a perfectly working memory in contrast to a human with a failing memory. Frank relies on Robot for his daily needs while Robot relies on Frank's compliance and improved health for his objective as his carer to be fulfilled. Their unspoken yet mutual recognition of each other's dependence begins an uneasy working relationship, but towards the end of the film this becomes a more intimate and affective understanding that is tinged with mutual respect. The relationship helps Frank wade through the murky depths of his failing memory as he reconciles himself to his family, particularly his ex-wife Jennifer. As Frank's affective relationship with Robot develops, the narrative gradually reveals the links between Frank's past and present, for example, how the librarian he has shown a strong romantic interest in is actually his ex-wife. In a similar way, Hiro's relationship with Baymax starts off with uncertainty and an initial distancing effect of HRI, and quickly moves into seeing Baymax as a useful tool in his quest to find out who killed his brother. As Hiro's quest for the truth turns into a journey of self-discovery, his relationship with Baymax moves from one that sees Baymax as a machine to one as an equal member of the superhero team he assembles. There is, however, an interdependency that is interesting to observe within this relationship. Baymax, like Robot in *RF*, is obliged to provide his patient with the assistive care he needs, and the success of Baymax's mission of improving Hiro's health is dependent on Hiro accepting the care that he offers, and this then leads Baymax to accepting increasingly dangerous missions from Hiro on the pretext of appeasing him and improving his overall health and well-being. Both film narratives appear to be pointing towards interdependencies between robot and user, advocating for affective relationships between the two subjects in ways that might be mutually beneficial. There is an element of wish-fulfilment here that is tied to both science fiction and the sociology of expectations, and that is the anticipation and hope for a communal environment where mutual recognition and interdependency exist between robot and user. However, there is a risk here that these depictions of assistive technology in care work have gone far beyond what technology is capable of producing in the near future, and any current ethical issues concerning HRI and assistive robot care are being ignored.

Focusing on the synths in caring roles in season 1 of *Humans*, both Anita (who is owned by the Hawkins family) and Odi (who is retired robotics expert George Millican's carer) are shown to assist and enhance the lives of their owners. George (William Hurt) is suffering from memory loss and Odi (Will Tudor), despite malfunctioning, is still able to retrieve memories associated with George's life and appears to have a close familial relationship with him. With the threat of Odi's removal looming, George takes great pains in hiding Odi from the authorities, even at cost to himself. While the story arc involving George and Odi was not fully explored in the series, it does depict another concern that has been discussed in papers on robot deployment in care work, and that is the issue of attachment. Research on possible psychological effects of robot deployment in care work on children and the elderly has raised concerns regarding this issue. While attachment to the right individual can form a secure base for growth, a child's attachment to a robot carer could pose problems in their development.[26] At present, technology does not allow for robots to provide the right form of guidance and care for children, and it would be better for children to form attachments to the right human carers, with further work then supplemented by the deployment of a robot carer.[27] The Hawkins's

children Toby (Theo Stevenson) and Sophie (Pixie Davies) form individually strong attachments to their synth Anita, and this threatens to distance them from their mother Laura (Katherine Parkinson). This particular story arc explores the problems with robot agency within a household environment, as there are often complex dynamics at play between family members that a robot carer will have to negotiate, and lends credence to the theory that robots would ultimately be better deployed in supplementary roles to human carers than as sole carers with agency to make major decisions. Both Odi and Anita are loyal to George and the Hawkins family, respectively, despite Odi's continued malfunctions and Anita's return to sentience in the form of Mia later on in Season 1. Both of these synths uphold the established directive of protecting the lives of whom they serve, demonstrated later on when Anita steps in front of a fast approaching vehicle to protect Toby in episode 3. The complication comes when we meet George's new carer synth Vera (Rebecca Front) in episode 2, who had been issued to George to replace Odi. Vera did not have the intimate relationship and history that George had with Odi, and Vera's more severe demeanour leads to George's feelings of imprisonment within his own home while being cared for. While Vera does not obey George's requests in order to provide a more disciplined and efficient form of care, it creates an atmosphere of confinement that George wishes to break free from. The series *Humans* presents a unique contrast to *RF* and *BH6*, in the sense that there is a pervasive atmosphere of mistrust evident between humans and their assistive robots, and varying states of unease exhibited as humans allow synth technology to enter the privacy of their homes. The synths are generally viewed and treated as high-end appliances in the home and workplace, where there is no expectation of affective relationships in HRI in the context of the world depicted in the series. George and Odi's unique relationship represents one of the very few instances of complete trust being shown between human and robot, however this seems to be largely due to George's past as an expert who worked on early synth technology, and his familiarity with its capabilities and weaknesses help him to accept Odi as part of his everyday living experience. There is perhaps a warning here, through the fictional narrative, of a practical need for guidance on potential attachment issues within HRI, particularly in care work where human vulnerabilities are on display and subject to greater scrutiny.

A MIRROR TO OURSELVES: RECOGNITION, COMMUNITY AND THE SOCIAL

While the robots depicted in the works in question are understandably the centre of attention, they also function metaphorically as a mirror held up to the human characters, causing them to remember and confront what is missing in their own lives. These science fiction texts continue the narrative tradition of the companion robot in fiction: characters that not only point to a plausible future in terms of technological advancement, but that are also representative of the past as they reflect the actions of their human users.

The robots in all three works are not only signifiers of loss but are also repositories of memories. In the case of Baymax in *BH6*, it is actually a physical representation of memory, in the form of Hiro's brother Tadashi's healthcare chip. This artefact, complete with his brother's handwriting on the surface, and together with Baymax's robot body, consolidate various aspects of Tadashi's personality, including both his ambition and kindness. Whenever the chip is removed from Baymax, Tadashi's healthcare protocols along with videos of his time working on

Baymax are removed as well. Hiro is forced to confront the stark reality of his brother's death when Baymax refuses to open the chip access port at his request. Baymax says to Hiro that 'Tadashi is here' (in Baymax), and plays from a screen embedded in his chest a montage of video clips of Tadashi building Baymax. This scene marks the first time that Tadashi has appeared on screen since his death, and Hiro realises that he has allowed himself to be consumed by his desire for revenge and lose sight of the sense of kindness and humanity that was so integral to Tadashi and his work. Ultimately, Hiro's attachment to Baymax is really his attachment to the memory of his brother, a memory that continues to endure through Baymax. In season 1 episode 1 of *Humans*, George's carer Odi, despite being an outdated model, is of great significance to him as he functions as a repository of memories of George's past, allowing George to have occasional remembrances of his own life and of his deceased wife. Anita's return to sentience in the form of Mia in season 1 episode 6 simultaneously marks the return of her memories as she is finally reunited with members of her original family of synths (Leo and Max) who were all created by the enigmatic David Elster (Stephen Boxer).

Robot in *RF* serves a dual mnemonic purpose: first as a constant reminder of Frank's dementia (by being Frank's carer), and second as a repository of new memories being created in the present. Robot's constant companionship and close physical proximity to Frank serve as reminders of both his dementia and his estrangement from his family. Robot's dual mnemonic function is manifest most acutely in the climactic moment when Frank shuts him down, in the scene briefly examined earlier. Towards the end of the film, as the Sheriff is calling for Frank to come out of his house, Frank resolutely hatches yet another plan to elude the authorities, but is stopped by Robot who jolts Frank back into the reality of his predicament and offers him the only viable solution that will get him out of trouble with the law, and that is to destroy all evidence of their exploits by shutting Robot down and wiping his memory. Frank is understandably resistant to the idea, however changes his mind when he hears Robot repeat his own earlier mantra back to him about 'planning (his) next job' dealing in 'diamonds and jewels' and 'lifting that high end stuff', bringing stark and tangible proof to Frank of Robot's function as 'just an advanced simulation'. Frank comes to a sense of recognition, not of Robot's status as his robot companion, but of his own complicity in manipulating Robot's directives to his own ends where supposedly 'no one gets hurt'. In being 'an advanced simulation', Robot is able to act as a mirror to Frank's own self-deception by reflecting his lived experiences, including his misuse of robotic technology. There is a sense in this scene of a parental figure suddenly recognising their exploitation of a child's naivety. Amelia DeFalco's insightful and detailed analysis of this poignant scene illuminates the 'haptic intimacy' that occurs between the two characters before Frank shuts Robot down.[28] Robot adheres to Asimov's first two Laws by not allowing Frank to come to harm through his own inaction, while going against Frank's wishes of not discussing the idea of wiping Robot's memory. When Frank eventually pushes the button and Robot's memory is erased, Robot gently collapses into his arms, and a sense of loss is once again palpable: a recent friend and companion has now disappeared for good, and from the point of view of caregiving, the robot carer that Frank grew fond of and successfully interacted with is gone. Towards the end of the film Frank, living in a care home, spots what he thinks is his robot before realising that there are many identical robots to the one he had working in that care home. This points to a key aspect of HRI that deals with the need for individuals to build

emotional ties with their robot carers. However, the impersonal nature of the robot as a machine-generated and functional object places it at odds with this very need. In his examination of the theme of recognition, the philosopher and critic Paul Ricoeur argues that 'for recognition to be fully beneficial to an individual it needs to be mutual'.[29] In other words, mutual recognition taps into the desire of individuals to be recognised within a community; and this might be a useful way of examining the social aspect of interaction between robot carers and humans. It is important that patients come to recognise their carers (human or robot) as significant parts of their lives, and not just a product of the corporate world of modern healthcare. This will allow for more productive interactions that will potentially result in better overall care. In the fictional realm, when Frank overcomes his prejudice and objections to having a robot carer, and genuinely interacts with Robot, his day-to-day outlook begins to change as well. The act of recognition in fictional narratives featuring assistive healthcare robots such as *RF* asks important questions about how robots are viewed by individuals, and subsequently, how individuals view themselves.

Social interaction

The philosophical concept of the social is an important consideration when thinking about the roles of the robots in these three fictional texts. As social beings, humans rely on those around them not just for interaction and learning, but also to testify to one another of their existence in the world. Maurice Halbwachs describes humans as social beings by nature, and that more often than not memories are always collective.[30] Humans rely on one another to remember events, as well as remembering each other.

As social beings relying on human interaction, the introduction of robots into care work seems counterintuitive, as machines are at present unable to satisfactorily replicate and replace such forms of communication. Coeckelbergh argues that an over-reliance on 'modernity' might lead to people becoming more solitary beings; as care-receivers it is easy to fall victim to the sense of alienation found in contemporary healthcare practices, and lose the sense of vulnerability and interdependency that is required for good care.[31] Coeckelbergh calls for the roles of robots in healthcare to be re-defined and that robots in care be supplementary to the work of human care workers, as regular social interaction is essential for human health.[32] Perhaps the solution to this is one that is dependent on time and the advancement of AI research, which is trying to find the keys to understanding human consciousness and developing this for robots in future, allowing for authentic conversations between robot carers and users.[33] Setting aside any obvious apprehensions about robots turning sentient, such an advancement would dramatically alter the way robots relate to and interact with people, especially the people they are caring for. Robots require social interaction too in order for them to learn and interpret information related to care work and to build on experience. This calls for advanced 'social learning' abilities so that robots can learn from human carers as well as other robot carers.[34] This is particularly true in looking after people with dementia, as their identities can change with their illness. Carers often have to respond to these shifting identities and make adjustments in the way they interact with them.[35] While we are still a number of years away from witnessing more advanced levels of social interaction from robots, it is interesting to observe companion robots in care work depicted fictionally on screen with communicative abilities that can respond to the complexities of human interaction, whether it is the one-to-one interaction between Frank and Robot, or interactions within a surrogate family such as that between Hiro, Baymax and the rest of the Big Hero 6 team, and the Hawkins family with the synths in *Humans*.

At a deeper level of engagement, these fictional narratives question people's expectations of care work and how they view their relationship with care technology. Interaction between humans is likely to be very different to those between humans and robots, and if humans envision a future world that incorporates robots into everyday public and private spheres, then the ways in which they interact within their social and technological environments will need to undergo fundamental changes. These future environments will feature both humans and robots sharing the same domestic, work and leisure spaces, necessitating a complex negotiation of communication strategies when a subject interacts with both humans and robots in the same space. Both films suggest that a new understanding (and perhaps a new language) is needed for productive and fulfilling HRI, and this refers not just to changes from the creators and technologists who work with robots, but also changes from users, patients and clients as well. The concept of the social has its origins in the need for mutual recognition among humans, and future developments within HRI need to consider these essential elements moving forward. This is not necessarily a veering towards philosophical definitions of whether robot carers should be viewed as equals, but an acceptance that the robots are part of what DeFalco calls the 'posthuman interdependency',[36] where there is a 'relational and interdependent' connection between humans and robots that focus on the productive and life-affirming communication between patient and carer, even if the carer is non-human.

Carme Torras, in her paper on Asimov's short story 'Robbie' from *I, Robot*, discusses the benefits of science fiction as a tool for public engagement with robot technology used in healthcare.[37] Torras, who is both a scientist as well as a science fiction writer, describes how the genre is becoming more accepted within the scientific community, and allows for opportunities for non-scientists to be involved in the development of robot technology. She believes that generating engagement regarding the use of robot technology in care helps people to consider and clarify what their relationship with robotics is, and what it means to be human. She argues that robots at home ought to be treated as appliances, rather than 'emotional surrogates'.[38] It is interesting that the opening title sequence to each episode of *Humans* and the end credit sequence of *RF* show montages of real-life robots in demonstration, particularly in the case of *RF* where the robots are shown to be demonstrating their abilities in care work, thereby bringing the audience into the world of robotics and simultaneously generating space for conversations about robots in healthcare and at home. Before season 1 of *Humans* was first broadcast in the UK, a very convincing and innovative advertising campaign from a fake company called Persona Synthetics began advertising synths as domestic helpers, kicking off the debate ahead of episode 1 being aired.

These fictional narratives speculatively explore the human relationship with technology through alternate universes, asking the question of whether it would be possible to coexist harmoniously with robotic technology. Sparrow and DeFalco, via differing perspectives, have pointed toward the necessary changes needed in order for this to be possible. DeFalco's argument, supported by her analysis of *RF*, calls for a more critical posthumanist future, where the desire for embodied technology is tempered with an understanding of the human need for interdependence.[39] The first season of *Humans* certainly points towards such a future, as the synths discover the universal code that would allow for them to be given sentience. While DeFalco

examines the subtle elements of subversion within *RF*, *Humans* thrives on subversion and challenging preconceived ideas about living with robots within a community. The series goes against the grain set out by the other two texts concerning affective relationships and interdependencies between user and robot. Sparrow's discussion of the need for recognition and respect in HRI simultaneously points to the impossibility of robots being able to exhibit these human qualities,[40] due to their lack of interiority and inability to build affective relationships within a community as mentioned above. The fictional texts examined here, particularly in the case of *Humans*, suggest that these ethical components can only be possible if robots are capable of attaining high levels of social competence, in order for them to experience and express an understanding and appreciation of human social relations and community. The development of social competence in robots and communal living with assistive robotic technology are intriguing questions to explore with regards to the future of HRI, but those are beyond the scope of this essay.

These science fictional narratives featuring robot carers ultimately lead to questions concerning expectations of care work, particularly from the point of view of the patient. Viewers are being asked what kind of relationship they ultimately wish to have with technology, both within their communities and in their domestic lives. From the point of view of the patient, the ideal carer would be one that could operate at all hours of the day and night, exhibit empathy, with the ability to anticipate behavioural fluctuations and participate in social relations within a community of patients, human carers, and other robot carers. Fictional representations of robot carers like those in *RF*, *BH6* and *Humans* suggest a rethinking of people's relationships with and expectations of assistive technology is necessary going forward. Certainly one thing that can be gleaned from applying principles on sociotechnical change such as the sociology of expectations, is that these visions of the future are inherently selective and contradictory.[41] They are selective through privileging certain narratives of HRI and robots in assistive care over others and ignoring historical issues, and they are contradictory due to the nature of shared visions that have to cater to a wide demographic. So while there might be a desire to see robots exhibit a certain level of autonomy operating within a caring role and having mutually affective relationships with their users, the robot carer seems best placed not to replace a human carer, but to assist the carer in varying degrees and in different aspects of care, depending on both need and functionality. The relatively widespread usage of PARO the therapeutic robo-pet, designed by AIST in Japan, across a range of care settings is evidence that establishing some level of affective relationship between robots and patients is beneficial, and so is using robotic technology as part of an overall care system (as opposed to complete care replacement). Interdependency in future relationships between the human carer and robot carer might contribute to better individual care while also causing complications in working alongside one another. Interdependent relationships between patient and robot carer in the near future might yield the sense of mutual recognition and a sense of the social within a community of care that robotic technology currently lacks. The alignment with Asimov's first two Laws exhibited by some of these fictional robot carers points the way for assistive technology to operate within a 'relational and interdependent' basis with its users (to reapply a phrase coined by DeFalco),[42] in order for greater opportunities for effective care to take place.

CONCLUSION

These three works of science fiction, released between 2012 and 2015, engage with some of the possibilities of the future of HRI within care work. The narratives consider how robots might work in more collaborative and interdependent relationships with patients to generate good care as well as enhanced social interaction. The texts challenge assumptions about HRI in care work, as well as the dystopian narrative and replacement fear of robots being the sole carers of vulnerable patients. They demonstrate that there is merit in considering the importance of affective relations between a robot carer and a patient. In the course of this essay, I considered how robot design and autonomy has an effect on patient responses to care work, how the effectiveness of robots as carers is often dependent on how well they operate within the human social environment, and finally how a more collaborative and affective relationship between the robot carer and patient stems from the mutual recognition of being part of a community of patients and carers. The texts examine, through fiction, people's values and expectations concerning the utilisation of robots as assistive technology in care work. Using a wide-ranging and interdisciplinary approach that includes assistive robot design, HRI, science fiction and expectation studies, this study examined how individual expectations of the use of technology in care work may be challenged through fictional representations. The texts examined seem to propose and advocate a communal environment where the assistive robot might benefit from a mutually affective relationship with its user. While it has so far been largely accepted that robotic technology is better suited to assisting human carers rather than fully replacing them, the potency of the 'rhetorical vision'[43] and technological expectation of autonomous robot carers will mean that future science fictional narratives depicting care robots may continue to make similar proposals as presented in these texts. Though these examples represent some of the long-term issues that will continue to affect the representation of robot carers in science fiction, by analysing them closely they can help people reconsider their own relationship with robotic technology, and how they may go on to discover new and mutually beneficial pathways toward better care.

Patient and public involvement statement

This research was done without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy.

ORCID iD
Yugin Teo http://orcid.org/0000-0003-4101-6298

NOTES

1. Amanda Sharkey and Noel Sharkey (2012), "Granny and the Robots: Ethical Issues in Robot Care for the Elderly," *Ethics and Information Technology* 14, no. 1 (March 2012): 27–40, https://doi.org/10.1007/s10676-010-9234-6

2. Theodore A Metzler, Lundy M Lewis, and Linda C Pope (2016), "Could Robots Become Authentic Companions in Nursing Care?" *Nursing Philosophy* 17, no. 1 (Jan 2016): 36–48, https://doi.org/10.1111/nup.12101

3. Christoph Bartneck (2013), "Robots in The Theatre and The Media," Proceedings of the Design & Semantics of Form & Movement (DeSForM2013), Wuxi: 64–70.

4. Mark Coeckelbergh et al. (2016), "A survey of expectations about the role of robots in robot-assisted therapy for children with ASD: Ethical acceptability, trust, sociability, appearance, and attachment," *Science and Engineering Ethics* 22, no. 1 (Feb 2016): 47–65, https://doi.org/10.1007/s11948-015-9649-x

5. Bartneck, "Robots in The Theatre and The Media," Proceedings of the Design & Semantics of Form & Movement (DeSForM2013), Wuxi (2013): 64–70, https://pdfs.semanticscholar.org/32a6/be0e8b19ec1a20438a38a83b5d158b8f85ac.pdf

6. Helmut Hauser (2015), "The real soft robots that inspired Baymax, with Chris Atkeson," *Robohub* (28 April 2015), https://robohub.org/the-real-soft-robots-that-inspired-baymax-with-chris-atkeson/

7. Benjamin K Sovacool and David J Hess (2017), "Ordering Theories: Typologies and Conceptual Frameworks for Sociotechnical Change," *Social Studies of Science* 47, no. 5: 703–50, https://doi.org/10.1177/0306312717709363

8. Sovacool and Hess, "Ordering Theories," 723.

9. Mads Borup et al. (2006), "The sociology of expectations in science and technology," *Technology Analysis & Strategic Management* 18, nos. 3–4 (2006): 285–86, https://doi.org/10.1080/09537320600777002

10. Sovacool and Hess, "Ordering Theories," 723.

11. Sharkey and Sharkey, "Granny and the Robots," 27–40.

12. Coeckelbergh et al., "Robot-Assisted Therapy for Children with ASD," 47–65.

13. Bartneck, "Robots in The Theatre and The Media," 64–70.

14. Ibid, 69.

15. Coeckelbergh *et al*., "Robot-Assisted Therapy for Children with ASD," 47–65.

16. Ibid.

17. Ibid.

18. Mark Coeckelbergh (2015), "Artificial Agents, Good Care, and Modernity," *Theoretical Medicine and Bioethics* 36, no. 4 (Aug 2015): 265–77, https://doi.org/10.1007/s11017-015-9331-y

19. Sharkey and Sharkey, "Granny and the Robots," 27–40.

20. Gavin Miller and Anna McFarlane (2016), "Science Fiction and the Medical Humanities," *Medical Humanities* 42, no. 4 (Dec 2016): 213–18, doi:10.1136/medhum-2016-0 11144.

21. Istvan Csicsery-Ronay (2008), *The Seven Beauties of Science Fiction* (Middletown, Connecticut: Wesleyan University Press), 74–75.

22. Isaac Asimov (2013), *I, Robot* (London: Harper Voyager).

23. Coeckelbergh, "Artificial Agents, Good Care, and Modernity," 265–77.

24. Dean Petters, Everett Waters, and Felix Schönbrodt (2010), "Strange Carers: Robots as Attachment Figures and Aids to Parenting," *Interaction Studies* 11, no. 2: 246–52, https://doi.org/10.1075/is.11.2.11pet

25. Robert Sparrow (2016), "Robots in Aged Care: A Dystopian Future?" *AI and Society* 31, no. 4 (Nov 2016): 445–54, doi.10.1007/s00146-015-0625-4.

26. D. Petters, E. Waters, and F. Schönbrodt, "Strange Carers," 246–52.

27. Ibid.

28. Amelia DeFalco (2016), "Beyond Prosthetic Memory: Posthumanism, Embodiment, and Caregiving Robots," *Age, Culture, Humanities* no. 3, https://agriculturehumanities.org/WP/beyond-prosthetic-memory-posthumanism-embodiment-and-caregiving-robots/

29. Paul Ricoeur (2005), *The Course of Recognition*, trans. D. Pellauer (Cambridge, Mass.: Harvard University Press).

30. Maurice Halbwachs (1980), *The Collective Memory*, trans. F J Ditter, Jr. and V Y Ditter (New York: Harper and Row).

31. Coeckelbergh, "Artificial Agents, Good Care, and Modernity," 265–77.

32. Ibid.

33. Metzler, Lewis, and Pope, "Could Robots Become Authentic Companions in Nursing Care?" 36–48.

34. Petters, Waters, and Schönbrodt, "Strange Carers," 246–52.

35. David M R Orr and Yugin Teo (2015), "Carers' Responses to Shifting Identity in Dementia in *Iris* and *Away From Her*: Cultivating Stability or Embracing Change?" *Medical Humanities* 41, no. 2 (Dec 2015): 81–85, doi:10.1136/medhum-2014-010581.

36. DeFalco, "Beyond Prosthetic Memory."

37. Carme Torras (2010), "Robbie, the Pioneer Robot Nanny: Science Fiction Helps Develop Ethical Social Opinion," *Interaction Studies* 11, no. 2: 269–73, https://doi.org/10.1075/is.11.2.15tor

38. Torras, "Robbie, the Pioneer Robot Nanny," 269–73.

39. DeFalco, "Beyond Prosthetic Memory."

40. Sparrow, "Robots in Aged Care," 445–54.

41. Benjamin K Sovacool and M V Ramana (2015), "Back to the future: Small modular reactors, nuclear fantasies, and symbolic convergence." *Science, Technology, & Human Values* 40, no. 1: 99, https://doi.org/10.1177/0162243914542350

42. DeFalco, "Beyond Prosthetic Memory."

43. Sovacool and Ramana, "Back to the future," 100.

BIBLIOGRAPHY

Asimov, Isaac. *I, Robot*. London: Harper Voyager, 2013.

Bartneck, Christoph. "Robots in The Theatre and The Media." *Proceedings of the Design & Semantics of Form & Movement (DeSForM2013)*, 64–70. Wuxi, 2013. https://pdfs.semanticscholar.org/32a6/be0e8b19ec1a20438a38a83b5d158b8f85ac.pdf

Borup, Mads, Nik Brown, Kornelia Konrad, and Harro Van Lente. "The sociology of expectations in science and technology." *Technology Analysis & Strategic Management* 18, no. 3-4 (2006): 285–98.

Coeckelbergh, Mark. "Artificial agents, good care, and modernity." *Theoretical Medicine and Bioethics* 36, no. 4 (2015): 265–77.

Coeckelbergh, Mark, Cristina Pop, Ramona Simut, Andreea Peca, Sebastian Pintea, Daniel David, and Bram Vanderborght. "A survey of expectations about the role of robots in robot-assisted therapy for children with ASD: ethical acceptability, trust, sociability, appearance, and attachment." *Science and Engineering Ethics* 22, no. 1 (2016): 47–65.

Csicsery-Ronay, Istvan. *The Seven Beauties of Science Fiction*. Middletown, Connecticut: Wesleyan University Press, 2008.

DeFalco, Amelia. "Beyond prosthetic memory: posthumanism, embodiment, and caregiving robots." *Age, Culture, Humanities* 3 (2016).

Halbwachs, Maurice. *The Collective Memory (translated by F. J. Ditter, Jr and V. Y. Ditter)*. New York: Harper and Row, 1980.

Hauser, Helmut. "The Real Soft Robots That Inspired Baymax, with Chris Atkeson." *Robohub*, 2015. https://robohub.org/the-real-soft-robots-that-inspired-baymax-with-chris-atkeson/

Metzler, Theodore A, Lundy M Lewis, and Linda C Pope. "Could robots become authentic companions in nursing care?" *Nursing Philosophy* 17, no. 1 (2016), no. : 36–48.

Miller, Gavin, and Anna McFarlane. "Science fiction and the medical humanities." *Medical Humanities* 42, no. 4 (2016): 213–8.

Orr, David M R, and Yugin Teo. "Carers' responses to shifting identity in dementia in *Iris* and *Away From Her*: cultivating stability or embracing change?" *Medical Humanities* 41, no. 2 (2015): 81–5.

Petters, Dean, Everett Waters, and Felix Schönbrodt. "Strange carers: robots as attachment figures and aids to parenting." *Interaction Studies* 11, no. 2 (2010): 246–52.

Ricoeur, Paul. *The Course of Recognition (Translated by David Pellauer)*. Cambridge: Harvard University Press, 2005.

Sharkey, Amanda, and Noel Sharkey. "Granny and the robots: ethical issues in robot care for the elderly." *Ethics and Information Technology* 14, no. 1 (2012): 27–40.

Sovacool, Benjamin K, and David J Hess. "Ordering theories: Typologies and conceptual frameworks for sociotechnical change." *Social Studies of Science* 47, no. 5 (2017): 703–50.

Sovacool, Benjamin K, and M V Ramana. "Back to the future: small modular reactors, nuclear fantasies, and symbolic convergence." *Science, technology & human values* 40, no. 1 (2015): 96–125.

Sparrow, Robert. "Robots in aged care: a dystopian future?" *AI and Society* 31, no. 4 (2016): 445–54.

Torras, Carme. "*Robbie*, the pioneer robot nanny: Science fiction helps develop ethical social opinion." *Interaction Studies* 11, no. 2 (2010): 269–73.