Unsociable Robots – Empathy in *Robot & Frank*

Stefan Herbrechter

> [A social robot is...] (1) an embodied object with (2) a defined degree of autonomous behaviour that is (3) specifically designed to interact with humans on a social level and respond to mistreatment in a lifelike way. (Darling 2014: 20)

> Frank: There's nothing wrong with my memory. (Ford 2014: 8)[1]

Empathy makes "us" human?

> We would not be where we are today had our ancestors been socially aloof. (De Waal 2009: 7)

For the primatologist Frans De Waal, empathy is an "ancestral" human trait, an essential part of "human nature" that biological evolution has favoured to turn us into social beings. For a "kinder society" we had better turn to our primate cousins to see how rewarding empathetic behaviour can be. While the existence of (true, authentic, or "ethical") empathy in nonhuman primates, or even its precondition, "theory of mind" (or "mindreading"; cf. Zahavi, 2014, Coplan and Goldie, 2011), remains a matter or scientific dispute (cf. Andrews and Gruen, 2014), few people nowadays would contest that our shared biological embodiment with other nonhuman animals calls for some form of *affective* empathy (cf. Maibom, 2014: 3-9).

In the context of emergent postanthropocentric posthumanism, empathy is challenged as a human privilege and as justification for human exceptionalism. Indeed, De Waal's argument is actually based on the *mutuality* of empathy – *nonhuman animals* (if not all, then at least some) may *feel* empathy, too. While animal studies and animal rights discourses use the affective capability of empathy to call for a more empathic response from *humans* in turn, the argument concerning artificial intelligence, and intelligent robots in particular, with regard to empathy runs somewhat differently: the concern here is usually not about embodied *affective* empathy *with* or inherent *in* a *technological* other – there is no shared biology – but

---

[1] All further references to Robot & Frank will be to the script and given in the text as *R&F* plus page number.

rather on *cognitive* empathic ability, or the question whether a robot (or software, or smart environments, AI etc.) *understand* empathy, whether they can *know* what humans *feel* (which would of course make them *virtually* human). Needless to say, however, that it has always been *virtually* impossible to distinguish knowing from feeling in the first place – and this will be one of the main aspects of my argument here: for empathy to *function* a deliberate confusion of feeling and knowing is probably indispensable.

Robot & Frank

> ROBOT: Frank. I know you don't like to hear this, but I'm not a person. I'm just an advanced simulation. (*R&F*: 94)

"Rethinking empathy through literature" (cf. Hammond & Kim 2014) is certainly a legitimate move given empathy's inherent connection to narrativization: how could any "I" project itself into the place of an other, or experience the feeling of an other, without some form of identification? And how would any "mind reading" involved in identification – with one's self or an other – occur without at least a minimal process of comparison and thus without a temporary construction of a "we" – regardless of the question of the neuro-psychological "activation" of the process and of the role "mirror neurons" might play.[2] The same, one might safely assume, counts for "visual" narratives, including those of science fictional filmic scenarios, which narrativize imaginary encounters between humans and their technological

---

[2] This is, for example, the motivation behind Suzanne Keen's extensive work on empathy in the novel (cf. Keen 2007) designed to complicate the idea that "a very specific, limited version of empathy located in the neural substrate meets in the contemporary moment a more broadly and loosely defined, fuzzier sense of empathy as the feeling precursor to and prerequisite for liberal aspirations to greater humanitarianism" (p. viii). Even though novel reading might not "cultivate empathy that produces good citizens for the world", as Keen of course rightly argues (p. xv), it also seems quite undisputable that *fiction* (and the identification with (fictional) characters more specifically) allows for and encourages empathising. However, it is also self-evident, of course, that empathy is no *guarantee* for "doing the right thing", and thus doesn't *automatically* lead to a more humane attitude, altruism or a "greater humanitarianism".

others as projected into futures, that are, however, inevitably related to social presents, on the basis of speculation and extrapolation. It is in this sense that science fiction, according to Steven Shaviro, through "embodiments" of "what ifs", opens up "new lines of inquiry that analytic reasoning and inductive generalization would never stumble upon by themselves" (Shaviro 2015: 9). Basic forms of empathy (which include human and nonhuman forms of agency, and which occur outside or "before" cognition – hence Shaviro's emphasis on "discognition") might therefore be seen as an important source for if not the origin of (all) speculation.[3] I feel therefore justified in approaching the topic "robot empathy" and "robotic care for the elderly" through a reading of the 2012 film *Robot & Frank* – and I'm of course not the first to make the connection between robots, nonhuman agents, empathy and this film.[4]

While the word empathy is not pronounced in *Robot & Frank* there are nevertheless some aspects that constitute in my view an important fictional or speculative visualisation of an issue which is in the process of becoming *social* reality – an increasing human-nonhuman entanglement with so-called social or sociable robots.[5] In a contribution to a volume on *Sociable Robots and the Future of Social Relations* (Seibt, Hakli & Norskov, 2014) entitled

---

[3] These ideas most certainly also apply to speculations on the future of empathy itself, as Stahl et al. insinuate with reference to "care robots": "It is in connecting concrete, tangible technology with plausible human scenarios that science fiction prototyping offers a useful tool in technological forecasting" (Stahl et al., 2014: 82).

[4] I'm aware of a number of articles by authors who specifically associate the film with "robot empathy": Jakub Zlotowski et al. (2015), who deals with the question of anthropomorphism in human-robot interaction; and Dennis M. Weiss (2014), who comments on the scenario of "a society in which the authentically human has been replaced by simulations, in which our closest ties are to machines rather than other human beings, our loneliness is assuaged not by the company of others but by robot companions, and our sovereignty and autonomy over technology disappear" (p. 219). However, only Kamphof (2015) and Redstone (2014) discuss the film in more detail.

[5] There is some confusion about terminology, although most authors seem to prefer "sociable". As my title suggests, we're also free to speculate on the negation of robots' sociability: in what ways would robots used in the care for the elderly be conducive to human-inhuman sociality, or in other words, would "empathic" robots augment or diminish human sociability?

"Making Sense of Empathy with Social Robots", the cognitive scientist Josh Redstone asks whether it makes sense to "conceive of people's emotional responses toward social robots as [genuine] empathic responses" (Redstone 2014: 172). Redstone evokes *Robot & Frank* as a "fictional example" to make the conceptual issues this question raises "a bit more concrete". He summarises the film's plot in the following terms:

> In the 2012 science-fiction film *Robot & Frank*, the titular character Frank, a retiree in the near-future who lives alone and who shows signs of dementia, is given a domestic robot that helps care for him. Frank initially regards the robot as somewhat of a nuisance rather than as a companion. Nonetheless, it soon occurs to Frank that he can use his robot to re-start his former career as a cat burglar. After the duo pull a number of heists and attract the attention of the police, Frank is forced to decide whether or not to delete the robot's memory in order to destroy the evidence of his crimes, and to protect himself and his family from the consequences of his behaviour. At this point, however, he has begun to treat the robot more like a person than a machine, thus he is hesitant to delete its memory. Indeed, Frank becomes so distressed by the decision he is faced with that the robot must remind him that it's "not a person, just an advanced simulation", and that by having Frank wipe its memory, it will have fulfilled its ultimate purpose, i.e. to help Frank. (Redstone 2014: 172)

Redstone eventually dismisses the idea that there is "genuine empathy" at work between Robot and Frank – or indeed, more generally, in the interaction between humans and social robots – and instead interprets the scenario as "a kind of perceptual illusion". He thus goes on to ultimately negate his own question: "if the human-like qualities of Frank's robot are just the product of an 'advanced simulation' then does it make sense to say that Frank's emotional experience for the robot, who he feels as though he is effectively killing by deleting its memory, truly counts as an instance of sympathy, compassion, or some other empathic response" (Redstone 2014: 172).

Robot & Frank, however, can indeed be said to touch on empathy as a key aspect of human-robot interaction (cf. Leite et al., 2013). The main question that arises from describing some instances of anthropomorphic human behaviour towards (social) robots as "empathic" (i.e. the fact that "empathic behaviours may be instantiated by agents that do not have the capacity to feel and, hence, truly empathize" (Stephan, 2014: 111)) is whether there is a need to distinguish between empathic *behaviour* (which can be simulated, i.e. known and

understood) and *affective* (or emotional) empathy (which apparently cannot be simulated because, precisely, it has to be experienced or *felt*).

The important feature of "true" empathy, according to the psychoanalyst Serge Tisseron, is that it is reciprocated (or that it is "extimate"), which, per definition, excludes animals and robots (Tisseron, 2011: 153-4). Whether the empathy Frank eventually develops towards Robot after passing through what the roboticist Masahiro Mori calls the "uncanny valley" (cf. Angelucci et al., 2014),[6] is in fact truly felt or whether it remains a form of sympathy, emotional contagion or even personal distress – in other words, whether it is based on "imaginative perception" or merely constitutes a "perceptual illusion" – the film certainly manages to sensitize our opinion including that of the "robotics engineers", as Redstone concludes:

> I wonder to what degree it matters to robotics engineers – and to those who interact with social robots – that empathy for social robots is, perhaps, nothing but an illusion, and whether this will give the designers and the users of social robots cause to reevaluate the role that empathy plays within the context of social robotics. (Redstone 2014: 177)

The question of "mattering" Redstone evokes – whether empathy actually *matters* in these scenarios – might be seen as the most important contribution *Robot & Frank* makes to the science *factional* discourse on "social robotics".[7]

The social driving force behind the development of social robotics the film takes as its cue is dementia (i.e. the increasing threat for the elderly to be excluded from sociality). "[A]pproximately 35.6 million people globally… have dementia. With the ageing of the population, this number is expected to double every 20 years. Dementia is one of the major reasons for admission into long-term aged care", as an Australian study on the positive effects of interactive robotic animals on engagement, mood states, agitation and psychotropic drug use in people with dementia explains (Moyle et al. 2015: 1). What is often neglected in the discussion on "robot empathy" is that people with severe dementia are often "dehumanized"

---

[6] The closest to an "uncanny-valley" moment in Robot & Frank probably occurs when Frank realises he's defending himself in front of an "appliance": "Look, there's nothing wrong with my memory, I just – What am I doing, I'm talking to a fucking appliance!" (Ford 2014: 16).

[7] By "science factional" I mean a deliberate blurring of the (always problematic) boundary between science fiction and science fact (see Herbrechter 2013: 107-134).

(both in a "mechanistic" and "animalistic" sense, to use Nick Haslam's (2006) terminology). In this sense, Frank's memory and behaviour become both more erratic *and* predictable as the story progresses, so that he is indeed seemingly becoming less of a "person" and more "robotic" himself. Frank is the visualisation of a general or popular perception about dementia that could be described in these terms: "people with dementia are vulnerable to acts of stigmatization, objectification, and infantilization that undermine their personhood. Within a hypercognitivist culture, they are seen as missing the core of what it means to be human: an autonomous, rational self" (Kamphof 2015: 360). The often proposed answer in terms of the future of care is the shift towards a "person-centred" approach, and this is where social care robotics is said to play a key role in the future: "Person-centred care also demands person-centred technologies" to prevent the ubiquitous threat of "social death" for dementia patients (Kamphof 2015: 360, 363). According to Ike Kamphof, therefore, in "raising the question of proper care, *Robot & Frank* challenges the dichotomy between cold machines and warm humans" (371). It shows instead the development of a relationship based on companionship, or, as Kamphof refers to it, "partners in crime".

Empathy and sociable robots

> …robots cannot *be* empathic. They may at most *behave as if* they were empathic. (Airenti 2015: 118)

> …if people were brutalizing their robots at home, we may want to keep an eye on their animals and children. (Darling 2015)

Humans are able and usually willing to anthropomorphize (and often also to deanthropomorphize) almost anything and anyone. Robots are no exception. Recent research by a "rising star of robotics" at MIT, Kate Darling, received wide coverage in *Forbes Magazine*, which ran a story entitled "MIT Researchers Discover Whether We Feel Empathy for Robots". *Yes Magazine* covered the same story under the heading: "Can We Learn About Empathy From Torturing Robots? This MIT Researcher Is Giving It a Try". The main outcome of Darling's experiments is that, what she calls "anthropomorphic framing" has an impact on how humans think about robots. This can "alter human behaviour and may even

impact public policy" (McNeal 2015). People were asked to destroy a bug-like robot with a mallet under different conditions:

> The participants hesitated significantly more to strike the robot when it was introduced through anthropomorphic framing (such as a name or backstory). In order to help rule out hesitation for reasons other than anthropomorphism, we measured the participants' psychological trait empathy and found a strong relationship between tendency for empathic concern and hesitation to strike robots with anthropomorphic framing. (Darling in McNeal 2015)

The basic idea of "anthropomorphic framing" is that by giving names and stories to robots it will be easier to empathise with them. The policy implications that Darling derives from this insight are potentially far-reaching for a society that is about to embark on large-scale "social robotics", as she explains:

> To preserve the advantages and future potential of robots, as well as facilitate the adoption of beneficial robotic technology, we should consider distinguishing between those robots whose use is hindered by anthropomorphic framing, and those whose use is enhanced by it. Framing could be a helpful tool in effecting this difference… For robots that are not inherently social in design, nor enhanced through social interaction, we should consider discouraging anthropomorphism using every tool at our disposal. Rather than viewing science fictional narratives and personification as harmless fun, those building or implementing robotic technology should be aware of framing effects. (Darling in McNeal 2015)

What is particularly fascinating in Darling's recommendation is first of all that it is another example of science faction at work, i.e. a deliberate blurring of science fiction and fact used by scientists to legitimate, popularize and promote their research to a lay audience with a view to influence "policy" (which in an academic setting and in relation to fundable research is also referred to as "impact"). We can only speculate of course what Darling would have to say about "anthropomorphic framing" in *Robot & Frank* but she would probably point to the strong link between the social care purpose of some robots and the anthropomorphic framing designs required to fulfil it.

The other interesting feature of Darling's comment is that she doesn't only believe in the possibility but in the *necessity* (in terms of *policy*) to differentiate – to frame and unframe, so to speak – between what we might call social and *un*social (not to mention potentially *anti*social) robots. The conceptual problems involved in this kind of framing and differentiating process, from a philosophical point of view, however, are enormous. To be fair, it is not that Darling doesn't see the problem. In an online interview she explains her intentions (with the help of a typically Cartesian move based on the equivalence between machines and animals):

> I'd like to find out if we can change people's habits around empathy by using robots, in the same way animal therapy[8] is used with children or in nursing homes. If we could use robots like animals and have a positive influence on people's empathic responses to others, that would be amazing… On the flip side, I'm also interested in whether people can become desensitized or have less empathy because of robots. What does it mean if someone brutalizes a robot that is very lifelike? (Darling 2015)

Darling certainly takes anthropomorphism seriously – on the other hand, she turns a blind eye to the fact that her policy implications, already from the start, take some fundamental form of anthropomorphism for granted, namely that it is in fact *possible* to "brutalize" a robot (as you might say of a living organism, e.g. an animal, being "brutalized" and which (or who?) is suffering as a result). Most importantly, Darling is also somewhat of a liberal humanist when she seems to imply that empathy is an exclusively human ability – an ability that humans can somehow switch on and off, if not at will then at least by the use of the right framing device. "There are times", she says, "when we should use robots strictly as tools and discourage anthropomorphizing – like in the context of military robots. In that case, it can be inefficient or dangerous to anthropomorphize robots" (Darling 2015) – never trust (or wrongly empathise with) a killing machine, as any science fiction will readily tell you! At the same time, the same science fiction will also try and demonstrate to you that anthropomorphism works best when humans are faced with machinic *power* and will play on the very human fascination and desire that is involved in imagining what it might be like to *be* a machine.

---

[8] On the benefits of animal assisted therapy for dementia see, for example, Baun and McCabe (2003), and Filan and Llewellyn Jones (2006).

In her actual policy paper ("Extending Legal Protection to Social Robots"), Darling speaks of the emergence and effect of robots that are "designed to interact with humans on a social level": "As more and more robotic companions enter into our lives and homes, and those robots are increasingly engineered to engage us socially, our inclination to project life-like qualities onto robots raises design and policy questions" (Darling 2014: 1). The question or dilemma this socialisation process bears is: "if we tend to perceive robots as lifelike things, should we be treating them more like devices or like creatures?" (1). In other words, and from a posthumanist-postanthropocentric perspective one might say, do they answer to the kind of questions raised by object-oriented-ontology or by critical animal studies? Or, as Darling provocatively puts it: "As we move within the spectrum between treating social robots like toasters and treating them more like our cats, there may be some ethical issues to consider" (Darling 2014: 9). From Darling's standard humanist point of view, however – which incidentally in *Robot & Frank* corresponds to Frank's daughter Madison's standpoint – these ethical issues focus on the following question: "If social robots are used as a helpful supplement, or in places where human care is not possible, then allowing them seems desirable. But if robots start replacing companionship, this might warrant careful consideration of whether and what aspects get lost in the process" (10).[9] There is of course the

---

[9] This is also Sherry Turkle's view, who from something like careful enthusiasm for the potential of digital media technology "companionship" in her earlier work (cf. Turkle 2002), in more recent publications has been focusing on the problematic aspects of what she calls humanity's "robotic moment" (Turkle 2011): "a situation marked by a readiness to a loss of authenticity in relationships that is brought about by the isolating effect of social media, the internet and communication devices such as mobile phones" (Jones 2017: 341), but also involves the increasing acceptance of social robots as "companions". Turkle leaves no doubt that she believes this trend to be utterly resistible: "The current position seems to be that if there's a robot that could fool me into thinking that it understands me, I'm good to have it as a companion. This is a significant evolution in what we ask for in our interactions, even on intimate matters. I see it in kids. I see it in grown-ups. The new robots are designed to make you feel as though you're understood. Yet nobody is pretending that any of them understands anything" (Turkle, in Fischetti 2014). Even more clearly (maybe even with a hint of "moral panic") she says in another interview that "If people start to buy the idea that machines are great companions for the elderly or for children, as they increasingly seem to do, we are really playing with fire. I think the stakes are very high" (Turkle 2015a). Her answer, her proposed

danger of "manipulation" through social robots, since "not only are humans inclined to anthropomorphize the robots they interact with, they also actively enjoy doing so" (Darling 2014: 11). It is thus the human *desire* to empathize that constitutes the actual problem.[10] As Darling argues, form the point of view of evolutionary psychology: "we may be hardwired to respond instinctively to cues like simulated pain or need. If this discomfort [with violent behaviour towards robotic objects] becomes stronger with increasingly sophisticated technology, how should we deal with it? Should we encourage or discourage it? Might we even reach a point where we want rules in place governing the treatment of our robot companions?" (13). Indeed, as she continues, it might be necessary to go down that route for the "protection of societal values". Just as with pets or children, if brutal behaviour towards social robots is not sanctioned their social status might come back to haunt us by threatening the erosion of social values themselves: "there is concern that mistreating an object that reacts in a lifelike way could impact the general feeling of empathy we experience when interacting with other entities" (15). The implications of Darling's articulated "concern" can hardly be overestimated. The very empathy that makes humans human and which provokes the anthropomorphism that makes objects (or robots) "social" might be challenged, or indeed threatened, not because these objects or robots *manipulate* "us" but because, in a sense, they might "out-empathise" us. In short, it is clear that social robots play with the limits of our empathy including the very concept of it.

In ethical terms, Darling openly declares herself a "Kantian" humanist (building on her Cartesian metaphysics referred to above, with its fundamental equivalence between animal and machine as the human's most important "significant others"), when she says: "The

---

resistance to this self-roboticisation of human feeling, is simple as she claims: "the good thing is we don't have to invent anything to turn it around. We already have each other to talk to" (Turkle here refers to her *Reclaiming Conversation* (2015b).

[10] Or, as Turkle explains: "Such digital creatures are the shock troops of a cultural moment when figuring out how a robot is 'feeling' in order to get along with it begins to seem a natural thing. When robots make eye contact, recognize faces, mirror human gestures, they push our Darwinian buttons, exhibiting the kinds of behaviour people associate with sentience, intentions, and emotions. Once people see robots as creatures, people feel a desire to nurture them. With this feeling comes the fantasy of reciprocation: as we begin to care for robots, we want them to care about us. In our nascent robotics culture, nurturance turns out to be a 'killer-app'" (Turkle 2010: 3-4).

Kantian philosophical argument for preventing cruelty to animals is that our actions towards non-humans reflect our morality – if we treat animals in inhumane ways, we become inhumane persons. This logically extends to the treatment of robotic companions" (19). As a consequence, these nonhuman others constantly come back to haunt the human, so that a strange mirror process is set in motion: to "prevent desensitization towards actual living creatures and protect the empathy we have for each other", Darling argues, thus requires laws to *protect* social robots. One case of "mistreatment" Darling marks up for such legislation might come straight out of *Robot & Frank*, namely the question whether switching off a robot (with the result of irreversible data loss or "death") should be seen as analogous to terminating an animal's life, because the loss of programmed reactions and the memories of social interaction will inevitably be emotionally traumatizing – first and foremost "to humans" (20), of course. This move towards "robot rights" would involve a fundamental shift, equivalent to the shift caused by "animal rights", namely from mere property legislation to the attribution of intrinsic value and rights, eventually even including a "right to life" (Darling 21). This would indeed constitute an important shift that might point towards the possible emergence of…[11]

Posthuman(ist) empathy?

> An android doesn't care what happens to another android. (Dick 2003: 86)

> What indeed are the purposes of living things? (Turkle 2010: 10)

The issue of a robot's "aliveness" is precisely what is at stake in the "switching off" scene in *Robot & Frank*. Frank has "taught" Robot to become his companion in a way he hasn't been able to teach his own children – in fact, he often confuses him with them. *Robot & Frank* is a kind of "buddy" or "road movie" that "portrays human-robot companionship as reciprocal

---

[11] The issue of robot rights is mentioned in *Robot & Frank* by Frank and his children, Hunter and Madison, as well as Robot himself: "Robot: Is your daughter politically aligned against robot labor?" (*R&F*: 24). However, it's Frank, who defends Robot's rights by saying: "FRANK: The robot isn't your SERVANT, Madison! You don't turn him on and off like he was a goddamned slave! – MADISON (screaming): It's a ROBOT!" (61), which is an ironic reversal and anticipation of the final switching-off scene (see below).

manipulation but also mutual transformation". Frank becomes more and more attached to Robot, whom he treats like a son, and even though Robot keeps reminding Frank that "he" (there is a clear gendering of Robot who is coded as a (male) "butler")[12] is not a "person".[13] He also clearly displays "comprehension"[14] and "social intelligence",[15] and thus "increasingly opens to Frank's specific being" (Kamphof 2015: 371). Nevertheless, when Frank, in the end, needs to protect himself and to cover his tracks so to speak, encouraged by Robot's "self-denial",[16] he decides to switch his "friend" off (and, by erasing the memory of their

---

[12] Even though Robot protests by saying: "I'm not a butler, Frank. I'm a healthcare aide programmed to monitor and improve your physical and mental health" (*R&F*: 15). The social care, domesticity, intimacy and empathy framework in which social robots necessarily operate inevitably raise questions about gender (but also other social categories like race and class), which are culturally specific. One of the main aspects for the "social engineering" dimension implied in sociable robotics is: should "existing" stereotypes or "scripts" be indulged to enhance (gendered) anthropomorphization and thus to "facilitate" affective bonding and human-robot interaction? Or should sociable robotics be seized upon as an opportunity to further the "deconstruction of gender representation" (Shaw-Garlock, 2014: 315) and move towards a (posthumanist) "post-gender" world (cf. Haraway 1991: 150)?

[13] Robot: "I'm not a real person" (*R&F*: 89); and "Frank. I know you don't like to hear this, but I am not a person. I'm just an advanced simulation" (94).

[14] Robot "agrees" to take part in Frank's burglary scheme and, seeing Frank's enthusiasm and planning skills, he says: "I'm very pleased with your progress, Frank. Planning this burglary was a great idea" (*R&F*: 52).

[15] For example, Robot "cunningly" provokes Frank's empathy by saying: "If you die eating cheeseburgers – what do you think happens to me? I'll have failed. They send me back to the warehouse and wipe my memory" (*R&F*: 17). When Frank later asks: "But what about when you said I had to eat healthy because you didn't want your memory erased – that seems like something more's going up in your noggin there", Robot rather "disingenuously" replies: "I only said that to coerce you. It doesn't matter to me if my memory is erased or not" (49). In order to help Frank, Robot is thus also prepared to "lie" for him and does so on a number of occasions.

[16] In a scene similar to the *Terminator* terminating himself to erase the memory (chip) that might otherwise help engender "judgment day" for humanity (*Terminator 2: Judgment Day* 1991), Robot "sacrifices" himself to save Frank from being incriminated: "It's not too late.

friendship, effectively kills off a big part of himself) – a decision which accounts for the unease and the "crisis of empathy" felt by some viewers who might see this as Frank's betrayal.

Empathy and its "crisis" in science fiction is of course not an innocent theme. The classic reference here is Philip K. Dick's *Do Androids Dream of Electric Sheep* (and its flmic adaptation, *Bladerunner*) with their mechanisms of human self-reassurance, the "Voigt-Kampff (empathy) test" and the "empathy box" (cf. Seegert 2011). *Robot & Frank* shares the "posthumanist" issue raised by *Do Androids/Bladerunner* – how do we tell the difference between an authentic human and an android (a perfect copy or simulation of a human)? Is empathy really a good "device" to shore up human uniqueness and thus to justify exceptionalism? As David Palumbo explains, the use of an empathy test in *Bladerunner* is rather "ironic", since the test is administered through a machine (i.e. a machine "decides" whether it is dealing with a human or an android). Moreover, the questions Deckard (Descartes?), who himself might be an android, asks "clearly reveal that the test really measures only a socially engendered, culture-specific aversion to killing or harming animals, this society's primary taboo, rather than any truly universal, innate, and exclusive – and therefore definitive – human quality" (Palumbo 2014: 1277). In fact, as Jill Galvan argues, *Do Androids* articulates human anxiety over the "animation" of its (technological) environment. In doing so, it anticipates "the awakening of the posthuman subject" who has already given up on the idea of "an exclusive and empathic community of humans" and instead "envisions a

---

You can reformat my memory. It's the only way for you to remain safe and healthy". Frank replies: "Can't you erase just the bad parts? (...) You can't just lobotomize yourself! Without your memory… You're the best lock picker in the world!" (*R&F*: 88-89). The particular poignancy in erasing Robot's memory of course lies in Frank's own progressing dementia and thus his own inevitable loss of memory, his own "erasure". The actual switching off scene (94-95) is presented both as a "recognition scene" (Frank looks at his own reflection in Robot's "astronaut-helmet", staring at Robot (and himself) "as if he's seeing it for the first time" (94), muttering a confused "I never meant…" (95), while Robot explains to him how he can be switched off. In doing so, Frank embraces Robot, who "goes dark" and leans against Frank like a child gone to sleep – a very moving scene of child "euthanasia", one might argue, encapsulating the ultimately (essentially human/humanist?) empathic dilemma of love and self-sacrifice.

community of the *post*human, in which human and machine commiserate and comaterialize, vitally shaping one another's existence" (Gavan 1997: 414).[17]

The main difference of *Robot & Frank* is that it turns away from the "dark", tragic and "retrofuturistic" urban décor of Ridley's movie and displaces the question of the exceptionalism of human empathy into a pastoral social comedy setting. In doing so it makes some interesting connections that a posthumanist or postanthropocentric notion of empathy might articulate more fully. It thus produces some good examples of Shaviro's "discognition" through (science) fiction (i.e. "embodiments" of "what ifs" that lead to "new lines of inquiry", see above). Thus, instead of a conclusion, let me list some of the (philosophical, moral and social) questions *Robot & Frank* raises with regard to "robot empathy" (as exemplified in the context of "robotic care for the elderly"):

1. If empathy cannot ultimately guarantee humanness, does extending the capability of empathy to nonhumans really constitute an (ethical) "improvement" or does it in fact lead to the destruction of the very notion of empathy? In other words, if this is the "choice" put before "us", do we care about empathy and its future and if so, why exactly? This is a philosophical question as much as a pragmatic "technological" one in the age of "affective computing" (Stahl et al. 2014: 83) and "empathic media" (McStay 2018). Allowing for empathy between humans and nonhumans inevitably leads to an erosion of human exceptionalism, but it also threatens notions (and institutions) of "humanity" based on "*com*passion", the bodily experience of suffering shared (or in principle shareable) by all (and only) humans *qua* humans, based on "narratives of suffering" (i.e. the legitimation of human rights and the question whether these can really rely on "mere species being" (cf. Laqueur 2009: 31)). This is related to the practice of anthropomorphism and its role in human-human and human-nonhuman relations. If empathy is provoked by real or imagined "reciprocity" then humans having empathy *for* nonhumans requires some form of anthropomorphism *on both sides*. However, there is a current tendency to see anthropo*morphism* as a form of human narcissism which needs to be overcome to create more ecological and social justice beyond anthropo*centrism* (for humans excluded from "humanity", as well as for nonhuman others). But how "bad" (for both humans and nonhumans) can anthropomorphism really be if it is also a precondition for empathy? Surely, nobody

---

[17] Turkle makes a similar argument for social robotics (see below).

would want to live in a world *without* empathy, even though, the moment "we" reconfirm empathy we also reinscribe at least a minimal (affective, "tragic") form of humanism (it can't be all that bad to be human and to empathise, etc.). This form of closure (i.e. the reaffirmation of the human, and the status quo) is one way of interpreting the last scene in *Robot & Frank*, where Frank, now in a clinical environment (i.e. the "Memory Center"), sees a "reincarnation" of Robot (now someone else's, identical, companion model): "For a moment it seems to recognize him, but then it turns away, helping another patient. It's not *Frank's* robot. With a wistful smile, Frank turns and enters his room" (*R&F*: 97). The projected moment of "mutual" recognition (and interpellation) is already inscribed with everything that might be necessary for a replay of an "empathic" relationship. One might even argue empathy is always already there in any encounter that is perceived as such. The second aspect worth noting is that, as in any relation of "companionship" (animal, machine, object, environment etc.) there is also a property relation – *Frank's* robot/not *his* robot – which is a precondition for a "personal" involvement in the playing out of the empathy scenario. Finally, the fact that Frank understands that empathy and friendship will not be replayed here throws him back to "wistfulness", i.e. the nostalgic *memory* of empathy. This in turn is what makes him eminently "human" and worth empathising with (at least for fellow humans who can share the real or imagined loss) – a humanist reconfirmation *par excellence*, one might argue: being human means the ability to share loss which causes a yearning for transcendence, or privileged moments in which humans can forget that they're (merely) human all the while confirming their "humanity". Taking robot or nonhuman empathy *literally* would inevitably destroy the (illusion of) tragic, humanist, greatness.

2. Empathy for nonhuman agents necessarily requires a form of animism and an attribution of some kind of "theory of mind" (otherwise why empathise with a robot, for example). If humans are indeed "hardwired" for empathy (cf. the mirror-neuron debate, which might be extendable to other "sentient" beings), can we really maintain that there is a difference between empathy as an important human *affect* and the more or less "mechanistic" reflex of empathic *behaviour*? This is the critique levelled at Deckard's empathy test. The price for finding out whether the other (any other) has

feelings to share with is attributing *subjectivity* to him/her/it[18] – a dangerous moment of "self-exposure", one might argue.[19] Turning this risky moment of encounter between subjects into some form of (Darwinian or neuroscientific) automation might in fact annihilate the essence of empathy (which ultimately implies *responsible/resonse-able* subjects). On the other hand, extending responsible subjectivity to nonhumans inevitably threatens human uniqueness. Which price would you prefer to pay? Especially in the context of the kind of "dehumanisation" that constitutes memory loss (i.e. dementia), as in *Robot & Frank*, the attribution of subjectivity (or "personhood") is crucially exposed. Robot losing "his" memory mirrors and anticipates Frank's demise. In fact, one might argue, Robot loses *Frank's* memory, the memory *of* Frank (subjective *and* objective genitive). Because Robot is Frank's hypomnematic memory storage device (which also contains the affective memories of their companionship), Frank, in switching-off Robot, deprives himself of a big chunk of his self – it will prove impossible to locate where exactly the respective

---

[18] Turkle, in her "precritical" or optimistic phase regarding the development of sociable technologies, formulates the challenge to develop "personable" or "person-enhancing machines" in the following words: "how to design artifacts that would cause people to experience them as having subjectivities that are worth engaging with" (Turkle 2002: 150). She concludes by predicting that "[s]ociable technology will enhance human emotional as well as cognitive performance, not only giving us more satisfactory relationships with our machines but also potentially vitalizing our relationships with each other, because in order to build better sociable objects we will have learned more about what makes us social with each other" (157). In her later work, Turkle becomes sceptical of the idea of "person-enhancement" and defensive of the "human". Too much empathy and sociality, it seems, isn't good for us, after all.

[19] The question whether androids can have "authentic experience" and thus can be seen as "subject-of-a-life" is discussed by Lee (2006). The crucial point here is how we understand "aliveness". More and more people would be happy to extend subjectivity and the kind of consciousness that implies the "anguish" caused by seeing oneself as a "subject-of-a-life" to nonhuman animals, especially mammals. What would be the implications – costs, benefits, risks, etc. – of extending these traits to "machinic" nonhuman others? See also Wilde and Evans (2017) who engage with this question in the context of gaming.

subjectivities are. Empathy, one could thus say, is the mechanism that embroils subjectivities in shared memory.

3. Following on from this: as the care dimension of social robotics becomes more urgent does the interaction between dementia patient and social robot and the human-nonhuman assemblage this might create effectively enhance or threaten "humanness", and, again, why should we care? The main objection to or concern voiced against robot companionship is indeed some form of "lack of authenticity" (or deception, maybe even "infantilization" and "deplorable sentimentality" (cf. Sharkey and Sharkey 2010: 251)) inherent in "simulation". Apart from the fact that there are no social robots like the one in *Robot & Frank* (and that they probably won't be technically feasible for quite some time, if ever – the real question here is whether they are even "desirable"), Amanda and Noel Sharkey explain that "any beneficial effects of robot pets or companions are a consequence of deceiving the elderly person into thinking that the robot pet is something with which they could have a relationship" (2012: 35). In short, giving an elderly person a robot companion is probably always demeaning and thus unethical since it cannot replace authentic *human* care and interaction. All kind of objections are thinkable here – would that mean a rejection of the idea of "companion species" in general? Is the humanist conceptual framework behind this argument maybe too "Robinson-like" – after all, what kind of replacement companion is Friday for Robinson, at least during the first encounter? Is there an inevitably colonialist, supremacist, not to say speciesist, kernel in the idea of (human-nonhuman, or maybe even worse, also in human-human) "companionship"? Undoubtedly, "stimulation" for the elderly[20] (or children) can be achieved by robots (even if it is only by providing humans with a "conduit" for human-human interaction or a "conversation topic") but this will ultimately remain a "s(t)imulation" threatening to lower "moral standards" (Sharkey and Sharkey 2010: 252).[21] Sherry Turkle,

---

[20] Robot suggests as much in replying to Frank who feels "infantilized": "FRANK: You're for children. Stupid. (…) ROBOT: Come on, Frank. It's important to have a project to focus. Mental stimulation plus a regimented schedule will vastly improve your cognitive functioning" (*R&F*: 14).

[21] As Amanda and Noel Sharkey (2010: 253, 254) argue: "Society has a duty of care and a moral responsibility to do its best to ensure the emotional and psychological wellbeing of all

similarly argues that there is something grotesque about elderly people telling "the story of their life to something that has no idea what a life is or what loss is". She thinks that "such interactions are a dead end" (Turkle, in Fischetti 2014) and merely constitute "evidence of a certain fatigue with the difficulties of dealing with people" (Turkle 2010: 5; see also Turkle 2011: 105-109). In short, handing over elderly humans to robotic care is a sign of *our* dehumanization: "Philosophers say that our capacity to put ourselves in the place of the other is essential to being human. Perhaps when people lose this ability, robots seem appropriate company because they share this incapacity" (Turkle 2011: 108). The circularity of this argument is worth pointing out: contemporary humanity suffers from a decline in empathy and outsources empathy to nonhuman others whom it begins to treat (falsely) with too much empathy, which, in turn, begs the question of what kind of people we are becoming "as we launch ourselves and those we love into increasingly intimate relationships with machines" (108). Is the fact that Frank seems to be totally impervious to this "moral panic" and embraces Robot as his "friend" while neglecting his real children and not recognizing his wife evidence of *social* decline? One scene might "help" complicate things in this respect. Frank's son Hunter is understandably upset when he finds out that Frank *pretended* to be dying to hide his criminal activity. He (anthropomorphisingly) addresses Robot with this reproach: "AND WHAT THE HELL *ARE* YOU? You're *lying* for him?" (*R&F*: 85), only to correct Frank's view, a few moments later, by telling him: "It's a ROBOT, dad! It's not your *friend*! It's your slave who you can apparently bullshit into doing just about *anything*" (85). The complicated "logic" behind this shift is interesting: Robot is a liar who can be brought to deceive, but he's also a slave who can be deceived by his human master – a perfectly Hegelian dialectic, which thus apparently also applies to nonhuman others. What Hunter is really upset about, however, is the idea of "care" and its multiple and ambiguous meanings: "You think I care?", he repeats twice and calls his father an "egotistical asshole". "You think I CARE? I'm just *guilty* because I AT LEAST HAVE FEELINGS! The best thing you ever did was being locked up so I didn't have to be raised by you. Why do you think I left you with some shitty, bargain [care]

---

of its citizens regardless of their age. We must not allow Companion carers to be used for alleviating our guilt about the isolation and loneliness suffered by many elderly people… we must not allow them to serve only to alleviate society's guilt about elder isolation".

robot?" It's almost impossible to unpick the emotional baggage contained in these statements. However, one thing the film makes clear in the context of the discussion about the threat of a moral social decline posed by sociable (or maybe rather "unsociable"?) robotics is that human beings don't seem to make too much difference between humans and nonhumans as far as the betrayal of trust is concerned.[22] It will be extremely difficult to both epistemologically-conceptually and ethically-pragmatically retain any "purity" (or human exceptionality) within the idea of empathy under these circumstances.

4. Which brings me to the implications of social robotics for the future of "sociability" (human-human, human-nonhuman and nonhuman-nonhuman). Robot is not the only robot in *Robot & Frank*. Jennifer, the librarian and the wife Frank does not recognize but nevertheless seems to be in love with, has a (much less anthropomorphic) android "helper" called Mr. Darcy (a reference to *Pride and Prejudice*, which, however, doesn't seem to have been exploited any further apart maybe the general idea of "bookishness" of "old-fashioned literature", see below). There isn't much empathy going on between Robot and Mr. Darcy and hardly any sociability either. When left "alone together" (cf. Turkle 2011) the two androids barely keep up a "conversation" (they merely check each other's "functionality") never mind going on a "robo play-date" as Frank suggests (*R&F*: 30, 39). Mr. Darcy simply observes: "I have no functions or tasks that require verbal interaction with the VGC-6OG" (41). The standard explanation for this is of course that "robots are unlikely to have empathy for other entities unless or until they have emotions of their own. Without emotions, empathic behaviour by robots will largely be the result of rational responses built on top of a merely symbolic representation of the minds of others" (Wallach 2009: 165). Human imagination can stretch as far as projecting empathy onto robots (i.e. "fictional awareness", including our awareness of it this being "fictional")[23] so that robots can

---

[22] See Mark Coeckelbergh on the question whether we can "trust" robots and whether they can actually "trust" us (Coeckelbergh 2012a and b). In fact, as Robot explains to Frank: "Developing trust between us is a part of my program" (*R&F*: 29).

[23] See Thomas Fuchs's classification of "primary" (embodied), "expanded" (cognitive) and "fictional" (imaginary) empathy (Fuchs 2013: 273), which, however, in my view can only have a heuristic function.

"simulate" it back to us, but there is no reason why a robot should display empathy towards another robot (which, however, would be the *literal* meaning of robot empathy – empathy as "robotic" from beginning to end, so to speak). What would be the benefit (for either robots or humans) of such a *literally* empathic robot behaviour? Well, *Robot & Frank* might also provide an answer for that question: when reprimanded by Jennifer for his "rude" comment about the absence of functionality, Frank suggests: "So that's it? When all humans are extinct you're not going to start a robot society?" (41), implying that AI doesn't need (social) communication (which raises the question whether communication without sociality is thinkable…). So, as a supremely ironic move Frank proposes to Robot to "assume Mr. Darcy here is a human like me, have a conversation" (41), upon which the androids start "simulating" a "human" conversation – to *please* the human – but only manage to return to what by now must be assumed to be robot "small talk": "Hi there, how are you doing?" – I'm functioning normally". – "As am I" (41). The viewer is thus left with the puzzle of which of the two is more "robotic" or less "empathic": human or robot small talk (see also the robots' "social dance" scene repeatedly apologizing for being in each others' way (90). Behind the comic irony, however, lies a "darker" possibility: what if robots *insist* on their unsociability? Aren't robots (or AI) more likely to produce a "robot-robot culture" (or an "emergent artificial culture") that will remain "alien", or an "exo-culture" to "us", and will thus risk by-passing if not "willingly" excluding humans? (cf. Winfield 2010: 206; a similar argument is made by Hayles (2005)).

If you feel completely lost in what has become an empath(et)ic web, I certainly feel for you – in the same way Sherry Turkle might feel for the "human" vis-à-vis the prospect of future unsociability. Let me therefore try and conclude and provide one additional point *Robot & Frank* manages to raise.

The core question that begins to appear behind the moral anxiety the prospect of robot empathy (empathy *for* and *by* robots) is provoking might be the following: is there something unethical about the way social robots, on the one hand, may begin to develop a "life of their own", so that they can be used "narcissistically" by humans (as "true" and "empathic" companions), and, on the other hand, that the (human) insistence that they be switched off at will – which may be the true reason why interaction with robots is becoming more attractive than interacting with humans (or nonhuman animals to some extent), because these cannot be so easily be "switched off"? In fact, given the kind of affective relation that we, like Frank,

may develop to such "empathy machines", and whom we thus continue to "educate" so that they can interact with us, our relationship with social robots, for the moment, seems fundamentally "pedagogical" (in still very humanistic sense). It mirrors the relationship we have with our children (cf. Tisseron 2011)? It is therefore no coincidence, I would argue, and this is *Robot & Frank*'s barely hidden "message", that the human anxiety of being somehow "dehumanized" and "replaced" (by robots, AI, simulative "care" etc.) occurs in the context of a major cultural shift that threatens to provoke the dismantling of humanism and its main institution: the library. The human "tragedy" unfolding in Frank's dementia is the tragedy of human memory, the film ultimately seems to suggest. Jake, the young digital philistine[24] and transformer of the library (and who, incidentally, is also the target of Frank's burglary) addresses Frank by saying: "Wow, so you must remember the days when this library was the *only* way to learn about the world! (…) I'd love to talk to you some more about your history with the printed word! You're our connection to the past! (…) Frank – do you remember *paper* newspapers? (…) Amazing. Printing thousands of copies – to be read in one day and thrown away! Think of it!" (*R&F*: 19, 20, 43). Jake's clearly posthumanist, digital plans for the library are further explained by Jennifer: "It's all augmented reality now. Jake says it's about the library experience. People can get any book at home instantly" (40). Which means that *Robot & Frank* is very much caught up in the ongoing transition from an analogue humanist and anthropocentric to a posthumanist digital and technocentric worldview. The device (or the *dispositif*) it invests in to shore up the "human" against its posthumanisation and that of its environment is: empathy. By anthropomorphically extending empathy to (android) robots it suggests that affective control over technology could remain possible. The price of becoming a little more "robot-like", in turn, is a price worth paying for humans as long as there is the prospect of an ultimate (if receding) affirmation of the boundary between humans and their (technological) nonhuman others. This "bottom line" is drawn by Robot "himself": "It doesn't matter to me if my memory is erased or not… Think about it this way – you know how you KNOW that you're alive? You think therefore you are… In a similar way, I KNOW that I'm not alive" (*R&F*: 49). Which leads me to conclude that from a humanist perspective (and Robot definitely is a "programmed" humanist) adding another layer of Cartesian self-reflexivity should probably do the trick. However, nothing is less certain. Let's hope these androids won't keep asking for more life, Deckard.

---

[24] Cf. also Frank's damning remark to Jake: "You've been googling me" (*R&F*: 45).

Bibliography:

Airenti, Gabriella (2015) "The Cognitive Bases of Anthropomorphism: From relatedness to Empathy", *International Journal of Social Robotics* 7: 117-27.

Andrews, Kristin and Lori Gruen (2014) "Empathy in Other Apes", in Heidi L. Maibom, ed., *Empathy and Morality*, Oxford: Oxford University Press: 193-209.

Angelucci, Adriano et al (2014) "A Philosophical Look at the Uncanny Valley", in Seibt, Hakli & Norskov, 165-69.

Baun, Mara M. and Barbara W. McCabe (2003) "Companion Animals and Persons with Dementia of the Alzheimer Type: Therapeutic Possibilities", *American Behavioral Scientist* 47.1: 42-51.

Bladerunner (1982) dir. Ridley Scott, Warner Bros.

Coeckelbergh, Mark (2012a) "Can we trust robots", *Ethics and Information Technology* 14: 53-60.

Coeckelbergh, Mark (2012b) "'How I Learned to Love the Robot': Capabilities, Information Technologies, and Elderly Care", in Ilse Oosterlaken and Jeroen van den Hoven, eds., *The Capability Approach, Technology and Design*, Berlin: Springer: 77-86.

Coplan, Amy and Peter Goldie, eds. (2011) *Empathy: Philosophical and Psychological Perspectives*, Oxford: Oxford University Press.

Darling, Kate (2014) "Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior towards Robotic Objects", available online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2044797

Darling, Kate (2015) "Can We Learn About Empathy from Torturing Robots? This MIT Researcher Is Giving It a Try", *Yes Magazine*, post by Nur Lalji, 14 July; available online at: http://www.yesmagazine.org/happiness/should-we-be-kind-to-robots-kate-darling

De Waal, Frans (2009) *The Age of Empathy: Nature's Lessons for a Kinder Society*, New York: Three Rivers' Press.

Dick, Philip K. (2003 [1968]) *Do Androids Dream of Electric Sheep?* London: Millennium.

Filan, Susan L. and Robert H. Llewellyn-Jones (2006) "Animal-assisted therapy for dementia: a review of literature", *International Psychogeriatrics* 18.4: 597-611.

Fischetti, Mark (2014) "The Networked Primate", *Scientific American* 311.3 (September).

Ford, Christopher D. (2014) *Robot & Frank* (script), available online at:
https://archive.org/details/pdfy-77SBwT09DNwG-xTd

Fuchs, Thomas (2013) "Der Schein des Anderen: Empathie und Virtualität", in Thiemo Breyer, ed., *Grenzen der Empathie: Philosophische, psychologische und anthropologische Perspektiven*, Münich: Fink: 263-282.

Galvan, Jill (1997) "Entering the Posthuman Collective in Philip K. Dick's *Do Androids Dream of Electric Sheep?*" *Science Fiction Studies* 24.3: 413-429.

Hammond, Meghan Marie, and Sue J. Kim, eds. (2014) *Rethinking Empathy through Literature,* London: Routledge.

Haraway, Donna J. (1991) *Simians, Cyborgs, and Women: The Reinvention of Nature*, New York: Routledge.

Haslam, Nick (2006) "Dehumanization: An Integrative Review", *Personality and Social Psychology Review* 10.3: 252–264.

Hayles, N. Katherine (2005) *My Mother Was a Computer: Digital Subjects and Literary Texts*, Chicago: University of Chicago Press.

Herbrechter, Stefan (2013) *Posthumanism: A Critical Analysis*, London: Bloomsbury.

Jones, Raya (2017) "Archaic man, meets a marvellous automaton: posthumanism, social robots, archetypes", *Journal of Analytical Philosophy* 62.3: 338-355.

Kamphof, Ike (2015) "In the Company of Robots: Health Care and the Identity of People with Dementia", in Aagje Swinnen and Mark Schweda, eds., *Popularizing Dementia: Public Expressions and Representations of Forgetfulness*, Bielefeld: Transcript, 359-75.

Keen, Suzanne (2007) *Empathy and the Novel*, Oxford: Oxford University Press.

Laqueur, Thomas W. (2009) "Mourning, Pity, and the Work of Narrative in the Making of 'Humanity'", in Richard Ashby Wilson and Richard D. Brown, eds., *Humanitarianism and Suffering: The Mobilization of Empathy*, Cambridge: Cambridge University Press: 31-57.

Lee, Billy (2006) "Empathy, androids and 'authentic experience'", *Connection Science* 18.4: 419-428.

Leite, Iolanda et al. (2013) "The Influence of Empathy in Human-Robot Relation", *International Journal of Human-Computer Studies* 71: 250-60.

Maibom, Heidi L., ed. (2014) *Empathy and Morality*, Oxford: Oxford University Press.

McNeal, Gregory S. (2015) "MIT Researchers Discover Whether We Feel Empathy for Robots", *Forbes online*, posted 10 April; available online at:

https://www.forbes.com/sites/gregorymcneal/2015/04/10/want-people-to-like-your-robot-name-it-frank-give-it-a-story/#12c848a948f9

McStay, Andrew (2018) *Emotional AI: The Rise of Empathic Media*, London: Sage.

Moyle, Wendy et al. (2015) "Effect of an interactive therapeutic robotic animal on engagement, mood states, agitation and psychotropic drug use in people with dementia: a cluster-randomised controlled trial protocol", *BMJ Open* 5.8: 1-6, available online at: http://bmjopen.bmj.com/content/5/8/e009097

Palumbo, Donald (2013) "Faith and Bad Faith in *Do Androids Dream of Electric Sheep?*" *Journal of Popular Culture* 46.6: 1276-1288.

Redstone, Josh (2014) "Making Sense of Empathy with Social Robots", in Seibt, Hakli & Norskov, 171-77.

Robot & Frank (2012) dir. Jake Schreier, Samuel Goldwyn films.

Seegert, Alf (2011) "Ewe, Robot", in D.E. Wittkover, ed., *Philip K. Dick and Philosophy: Do Androids Have Kindred Spirits*, Chicago: Open Court, 39-49.

Seibt, Johanna, Raul Hakli and Marco Norskov, eds. (2014) *Sociable Robots and the Future of Social Relations: Proceedings of Robo-Philosophy 2014*, Amsterdam: IOS Press.

Sharkey, Amanda and Noel Sharkey (2010) "Living with robots: Ethical tradeoffs in eldercare", in Wilks: 245-255.

Sharkey, Amanda and Noel Sharkey (2012) "Granny and the robots: ethical issues in robot care for the elderly", *Ethics and Information Technology* 14: 27-40.

Shaviro, Steven (2015) *Discognition*, London: Repeater Books.

Shaw-Garlock, Glenda (2014) "Gendered by Design: Gender Codes in Social Robotics", in Seibt, Hakli & Norskov, 309-17.

Stahl, Bernd et al. (2014) "The Empathic Care Robot: A Prototype of Responsible Research and Innovation", *Technical Forecasting and Social Change* 84: 74-85.

Stephan, Achim (2015) "Empathy for Artificial Agents", *International Journal of Social Robotics* 7: 111-16.

Terminator 2: Judgment Day (1991) dir. James Cameron, TriStar Pictures.

Tisseron, Serge (2011) "De l'animal numérique au robot de compagnie: Quel avenir pour l'intersubjectivité ?" *Revue française de psychanalyse* 75: 149-59.

Turkle, Sherry (2002) "Sociable Technologies: Enhancing Human Performance When the Computer is Not a Tool but a Companion", in Mihail C. Roco and William Sims Bainbridge, eds., *Converging Technologies for Improving Human Performance: Nanotechnology, Biotechnology, Information Technology and Cognitive Science*,

Arlington: WTEC: 150-158; available online:
http://www.wtec.org/ConvergingTechnologies/1/NBIC_report.pdf

Turkle, Sherry (2010) "In good company? On the threshold of robotic companions", in Wilks: 3-10.

Turkle, Sherry (2011) *Alone Together: Why We Expect More from Technology and Less from Each Other*, New York: Basic Books.

Turkle, Sherry (2015a) "Interview: I am not anti-technology, I am pro-conversation", interview with Tim Adams, *The Guardian* (Sunday 18 October), available online at: https://www.theguardian.com/science/2015/oct/18/sherry-turkle-not-anti-technology-pro-conversation.

Turkle, Sherry (2015b) *Reclaiming Conversation: The Power of Talk in a Digital Age*, New York: Penguin.

Wallach, Wendell (2009) *Moral Machines: Teaching Robots Right from Wrong*, Oxford: Oxford University Press.

Weiss, Dennis M. (2014) "Seduced by the Machine: Human-Technology Relations and Sociable Robots", in Dennis M. Weiss et al., eds., *Design, Mediation, and the Posthuman*, Lanham: Lexington Books, pp. 217-232.

Wilde, Poppy, and Adrienne Evans (2017) "Empathy at play: Embodying posthuman subjectivities in gaming", *Convergence* (Online First): 1-16.

Wilks, Yorick, ed. (2010) *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues*, Amsterdam: John Benjamins Publishing.

Winfield, Alan FT (2010) "You really need to know what your bot(s) are thinking about you", in Wilks: 201-207.

Zahavi, Dan (2014) *Self and Other : Exploring Subjectivity, Empathy, and Shame*, Oxford : Oxford University Press.

Zlotowski, Jakub et al. (2015) "Anthropomorphism: Opportunities and Challenges in Human-Robot Interaction", *International Journal of Social Robotics* 7: 347-360.