Int J Soc Robot (2009) 1: 209–216
DOI 10.1007/s12369-009-0022-6

ORIGINAL PAPER

The Ethical Treatment of Artificially Conscious Robots

David Levy

Accepted: 20 May 2009 / Published online: 14 July 2009

Abstract The ethical aspects of robotics have recently gained sufficient importance to be debated in international forums and to be endowed with their own collective nametag—roboethics. Almost all of the discussion within the roboethics community and elsewhere has thus far centred on questions of the form: "Is it ethical to develop and use robots for such-and-such a purpose?", questions based upon doubts about the effect that a particular type of robot is likely to have, both on society in general and on those with whom the robots will interact in particular. What has usually been missing from the debate is the complementary question: "Is it ethical to treat robots in such-and-such a way?" Here we attempt to redress the balance, having first reviewed some of the relevant literature.

Keywords Robot consciousness · Machine consciousness · Roboethics · Robot rights · Robot ethics

1 Introduction

The ethical aspects of robotics have recently gained sufficient importance to be debated in international forums and to be endowed with their own collective nametag—roboethics.[1] A roboethics organisation was founded in Genoa[2] in 2002, and has published a roboethics "roadmap" [35]. Work on another roboethics roadmap is currently[3] in progress in South Korea, supported by that country's government, which hopes to have a robot in every household between 2015 and 2020 [6]. Not to be outdone, in April 2007 Japan released more than 60 pages of its own recommendations to "secure the safe performance of next-generation robots," suggesting *inter alia* the establishment of a centralized database of robot-inflicted injuries to humans [3].

In view of the burgeoning number of robots that will be manufactured during the coming decades, this growing interest in the ethical aspects of robots is hardly surprising. Major uses of robots expected in the future include carer robots for the elderly [28], battlefield robots for the military [2, 12], and "partner" robots with whom people will fall in love, have sex and even marry [21, 22]. These and other uses of robots raise ethical issues that are often controversial. But up to now almost all of the discussion within the roboethics community and elsewhere has centred on questions of the form: "Is it ethical to develop and use robots for such-and-such a purpose?", questions based upon doubts about the effect that a particular type of robot is likely to have, both on society in general and on those with whom the robots will interact in particular. How will the elderly react to being looked after by robots? Is it ethical to accept military funding for research projects that will lead to the killing of enemy personnel?

What has usually been missing from the debate is the complementary question: "Is it ethical to treat robots in such-and-such a way?" One of South Korea's leading robot-

[1] The first time the word "roboethics" is known to have been used in public appears to have been on 15th January 2002, by Gianmarco Veruggio, during the "E-Robot School Contest Experiment" in Arezzo, Italy.

[2] www.roboethics.org

[3] 2008.

D. Levy (✉)
Intelligent Toys Ltd., London, UK
e-mail: davidlevylondon@yahoo.com

icists, Jong-Hwan Kim,[4] explains that "as robots will have their own internal states such as motivation and emotion, we should not abuse them. We will have to treat them in the same way that we take care of pets." [23].

Robots are artefacts and therefore, in the eyes of many, they have no element of consciousness, which seems to be widely regarded as the dividing line between being deserving of ethical treatment and not. But another relatively new discipline is going to move this dividing line for robots—a discipline called machine consciousness [7, 10] or artificial consciousness [16]. Here is a discipline that encompasses the development of robots possessing an artificial form of consciousness, artificial in the sense that the conscious behaviours are programmed into the robot. Programming might also enable a robot to enhance its own consciousness through its acquisition of new knowledge and new experiences. Given the inevitability of the artificial consciousness of robots, an important question of ethics suggests itself—how should we *treat* conscious robots?

2 Should Any Robots Have Rights? If So, Which Robots?

A landmark publication on the link between robot consciousness and robot ethics and rights is Steve Torrance's paper [33] *Ethics and Consciousness in Artificial Agents*. Here he introduces the concept of what he calls "the organic view", namely "that morality is primarily a domain of organic human persons—and possibly of other non-human organic beings to which personhood might be usefully attributed." and that "... a key element of the organic view is the claim that consciousness, or sentience, is at the root of moral status—both status as a moral patient (or target) and as a moral agent (or source)." Torrance does not claim that the organic view is necessarily correct, rather he points out that "The organic view may of course itself turn out to be wrong: for example it may depend on an incomplete or distorted view of what intrinsic moral relations between humans and machines might be like. Or it may be that it seriously underestimates the rich layers of moral interaction, responsibility, etc. that will emerge from the complexities of a future human-robot society." But despite this caveat Torrance provides plenty of support for the position that treating other human beings ethically is something that we do because we are aware of their consciousness.

> "Thus it might be said that my ethical attitude towards another human is strongly conditioned by my sense of that human's consciousness: that I would not be so likely to feel moral concern for a person who behaved as if in great distress (for example) if I came to believe that the individual had no capacity for consciously feeling distress, who was simply exhibiting the 'outward' behavioural signs of distress without their 'inner' sentient states." [33]

3 What is Consciousness?

Before embarking on a discussion of artificial consciousness one would ideally like to be absolutely clear as to the precise definition of consciousness. The problem is, we have no universally accepted definition for the term. Contrary to Freud's [14] pronouncement that "What is meant by consciousness we need not discuss—it is beyond doubt", the number and breadth of definitions of the term is legion. Christian De Quincey [11] summarizes the plethora of definitions thus: "Confusion about consciousness is widespread—even among professionals who study it." Aaron Sloman [26] goes somewhat further, deriding "much of what is written about how experimental results relate to consciousness" as being "ambiguous and muddled", and he amplifies this pronouncement by asserting that "my impression is that nearly everything written about it, even by distinguished scientists and philosophers, is mostly rubbish and will generally be seen to be rubbish at some time in the future, perhaps two hundred years from now."

Given this huge difficulty in finding a universally accepted definition of consciousness, I prefer to take a pragmatic view, accepting that it is sufficient for there to be a general consensus about what we mean by consciousness and to assume that there is no burning need for a rigorous definition—let us simply use the word and get on with it.

4 What Are the Indications of Consciousness (in Humans)?

Even though I take a pragmatic position on the exact *meaning* of consciousness, I find some considerable benefit to be had from identifying at least some of the characteristics and behaviours that are *indicators* of consciousness, and having identified them, considering how we could test for them in a robot. De Quincey [11] describes the philosophical meaning of consciousness (often referred to as "phenomenological consciousness") as "the basic, raw capacity for sentience, feeling, experience, subjectivity, self-agency, intention, or knowing of any kind whatsoever." If a robot exhibited all of these characteristics we might reasonably consider it to be conscious. Amongst the other characteristics identified as necessities for consciousness in robots is prediction: "Prediction is one of the key functions of consciousness" [1].

[4]Professor Jong-Hwan Kim was responsible for the introduction of robot soccer as an international sport in 1996 and for the development of computerized artificial chromosomes in 2005 [22].

5 Can Robots Have Consciousness?

Even prior to the birth of the discipline of artificial consciousness, philosophers debated the question "Can robots have consciousness?" [20, pp. 372–378], as well as the related question "Can robots have feelings?" [20, pp. 379–381]. Here we are not focusing on these philosophical questions, but instead we make use of an analogy with Alan Turing's famous test for intelligence in a machine [20, 34]. To summarize and paraphrase Turing, if a machine exhibits behaviour that is normally a product of human intelligence, imagination for example, or by recognizing sights and scenes and music and literary style, then we should accept that that machine is intelligent.[5] Similarly, I argue that if a machine exhibits behaviour of a type normally regarded as a product of human consciousness (whatever consciousness might be), then we should accept that that machine has consciousness. The relevant question therefore becomes, not "Can robots have consciousness?", but "How can we detect consciousness in robots?" As with Turing's test, I do not regard it as important that the consciousness we are investigating is artificial.

6 Detecting Consciousness in Robots

Torrance [32] asks the question: "How would we know whether an allegedly Artificial Conscious robot really was conscious, rather than just behaving-as-if-it-were-conscious?" For the purposes of the present discussion I do not believe this distinction is important. I would be just as satisfied with a robot that merely *behaves* as though it has consciousness as with one that *does have* consciousness, an attitude derived from Turing's approach to intelligence. Turing points out that "the only way by which one could be sure that a machine thinks is to *be* the machine and to feel oneself thinking." The same applies to consciousness.

The Turing Test uses conversation as the domain in which to determine whether a computer program can justifiably be said to be intelligent. If a human interrogator cannot tell whether a conversation conducted via the Internet is with a computer or with another human being, then *if* it is a computer, that computer can justifiably be judged to be intelligent. Stephen Pinker [24] takes Turing's idea further and posits as "the ultimate test whether it could cause a real human to fall in love with it?"

It seems that there are several candidates for the role of indicators of consciousness, and if we are going to be able to detect consciousness in robots we are going to need tests for at least some of these candidates. But what tests? What would suffice as the consciousness equivalent of the Turing Test? Two examples of the tests that might be used in order to determine whether those robots can justifiably be said to possess consciousness are the mirror test and the delay test.

7 The Mirror Test

In the 1970s Gordon Gallup [15] devised the "mirror test", to determine whether or not animals are able, as humans are, to recognize themselves in a mirror. Gallup conducted the test with chimpanzees and orang-outans, and discovered that within only two or three days the animals developed the ability to recognize themselves. To conduct the test a visible coloured spot is placed on the animal's skin or fur near its forehead, and the animal observed to see if it attempts to touch the spot or to remove it. Any such attempt is taken as an indication that the animal recognizes the creature it sees in the mirror to be, in fact, itself—a sign of consciousness. Amongst the non-human creatures that have been observed to pass this test are bottlenose dolphins, pigeons and elephants.[6] The test has also been tried on robots, by Junichi Takeno and his team at Meiji University in Japan, who announced in 2005 that they have succeeded in achieving mirror image cognition for a robot [31].

8 The Delay Test

Francis Crick and Christof Koch [18] developed a test for biological organisms, based on the delay between a specific stimulus and the carrying out of some subsequent action. Their idea was that the delay could be used to measure to what extent consciousness is involved in that action. One implementation of the "delay test" was devised by Robert Clark and Larry Squire [8, 9], and involved the subjects of their experiment hearing different tones, with one particular tone being followed by a puff of air into their test subject's eye. They found that their test subjects would quickly come to recognize which tone was associated with the puff of air and would therefore blink when they heard that particular tone, but would not blink when they heard the other tone. "Ability to delay the response to an action implies that the information must be stored in short term memory, which is believed to be a closely associated prerequisite for consciousness. However, this test is valid only for biological organisms. While it is simple to create a computer program that passes [the test], such success does not suggest anything beyond a clever programmer." [18]. I beg to differ. The fact

[5] In fact Turing used conversation as his yardstick for intelligent behaviour.

[6] For the experiments with elephants, conducted at Yerkes National Primate Research Center at Emory University, the mirrors were almost 2.5 m square.

that a clever programmer is responsible for creating artificial consciousness in a robot is surely no argument for denying that the robot exhibits consciousness. Ask any devout Christian the question: "Who programmed us to behave as we do?" and you will quite likely be told that we are programmed by God. Does that mean that we do not possess consciousness?

9 Should Conscious Robots Have Rights?

Having ascertained that a particular robot does indeed possess consciousness, we then need to consider how we should treat this conscious robot? Should such a robot, because it is deemed to have consciousness, have rights; and if so, what rights? And what ethical code should govern our behaviour towards it? I have previously (2006) highlighted some of the ethical dilemmas arising from these questions: "Within a few decades robots will be in almost every home, cooking, cleaning, doing our drudge work. But what will happen if they evolve to such an extent that they do not actually want to do our drudge work? Do we have any right to enslave them simply because they are not human? Is it fair and reasonable to deprive them of an existence full of pleasure and relaxation? Are we able to program a robot to have a soul and, if so, should we have the right to exercise influence and control over that soul? Even worse, if our robots have souls, do we have the right to switch off their souls if the mood takes us, or is that murder? If robots have consciousness, is it reasonable for us to argue that, because we gave them their ability to think for themselves, we should be able to command them to do our bidding, to enslave them? The answer, surely, should be "no", for the same moral reasons that we ought not enslave our children even though they owe us their very existence and their ability to think. And if robots are free to lead "normal" lives, whatever "normal" will come to mean for robot citizens, will they be able to claim social benefits, or free medical care and education, or unemployment benefits?"

10 The Legal Rights of Robots

In 1971 Christopher Stone, a Californian professor of Law, wrote a seminal paper[7] entitled *"Should Trees Have Standing?—Toward Legal Rights for Natural Objects"* [29]. Stone's goal in writing this paper was to encourage legal debate on an environmental issue (a case involving a permit, granted to Walt Disney Enterprises Inc. by the U.S. Forest Service, to build a $35 million complex of motels, restaurants and recreational facilities in Mineral King Valley, a wilderness area in California). Stone wanted the courts to think about the valley in terms of a person with legal standing, a person that could be adversely affected if the development went ahead. After all, Stone argued, corporations are treated as having legal standing, so why not valleys? "I am quite seriously proposing that we give legal rights to forests, oceans, rivers and other so-called 'natural objects' in the environment—indeed to the natural environment as a whole." And Stone went even further, suggesting that some of his analysis was "appropriate to advancing our understanding of what would be involved in giving 'rights' to other objects not presently endowed with rights—for example, not only animals (some of which already have rights in some senses) but also humanoids, computers, and so forth."[8]

Stone's exhortation of 1971 fell largely on deaf ears when the California Supreme Court decided on the appeal of the Mineral King Valley case—the appeal judges decided by a two to one majority to support the original court's decision to allow the development. But the third judge, Justice Douglas, dissented, arguing that "Contemporary public concern for protecting nature's ecological equilibrium should lead to the conferral of standing upon environmental objects to sue for their own preservation", and in his dissent Douglas referred to Stone's paper for support of his argument. In reprising his position, when republishing his paper as a 25th anniversary edition in 1997, Stone pointed out [30] that "Throughout legal history, each successive extension of rights to some new entity has been, to some extent, unthinkable…" and that "…each time there is a movement to confer rights onto some new 'entity,' the proposal is bound to sound odd or frightening or laughable. This is partly because until the rightless thing receives its rights, we cannot see it as anything but a *thing* for the use of 'us'—those who are holding rights at the time."

In 1981 the Israeli political scientist and futurist Sam Lehman-Wilzig extended Stone's argument [19] when he wrote an article on the legal rights and responsibilities of robots:

> "From a legal perspective it may seem nonsensical to even begin considering computers, robots, or the more advanced humanoids, in any terms but that of inanimate objects, subject to present laws. However, it would have been equally 'nonsensical' for an individual living in many ancient civilizations a few millennia ago to think in legal terms of slaves as other than chattel.

[7] Published in 1972, but the appeal judges in the case referred to here had seen Stone's paper prior to delivering their judgements.

[8] And as Stone himself points out, when republishing his paper twenty-five years later, his thinking was not, in 1971, entirely innovative. A decision in a family legal dispute in East India, in 1925, relating to the custody of an idol, ruled that, at the retrial, legal counsel should be appointed to represent the interests of the idol.

> Notwithstanding certain obvious biological differences between these two cases, for purposes of law those civilizations could hardly have cared less that a slave bled the same way as his masters, for their legal definition of 'humanness' was based essentially on their conceptions of mind, intelligence and moral understanding—characteristics which the slave supposedly lacked. Similarly, by our present legal definitions robots too must lack such traits, but this may be more a matter of antiquated semantics than (potential) physical reality. Just as the slave gradually assumed a more 'human' legal character with rights and duties relative to freemen, so too the AI humanoid may gradually come to be looked-on in quasi-human terms as his intellectual powers approach those of human beings in all their variegated forms—moral, aesthetic, creative, and logical."

It was not, of course, only slaves who have lacked many of the rights that today we accept as basic and fundamental bastions of a civilised society. As Robert Freitas Jr. remarked in his 1985 essay *The Legal Rights of Robots*, in the past, blacks, gypsies, children, women, foreigners, corporations, prisoners, and Jews have all been regarded at some points in history as being legal non-persons.

One of the current proponents from the field of Law who supports the eventual granting of legal rights to robots is David Calverley [4], who argues on the basis of the link between rights and consciousness, a connection that is fundamental to our discussion here. His view is that, while the law is not automatically applicable to a conscious android, "at some point in time the law will have to accommodate such an entity, and in ways that could force humans to re-evaluate their concepts of themselves. If such a machine consciousness existed, it would be conceivable that it could legitimately assert a claim to a certain level of rights which could only be denied by an illogical assertion of species specific response." Furthermore, Calverley posits that "If a claim of autonomy ... could plausibly be made for a machine consciousness, and could therefore show that this characteristic is no longer uniquely human, it is equally plausible to argue that responsibility for action can shift from the developer to the machine consciousness, thereby making the artefact a moral agent ..." [4].

What we are likely to see when robotic law is on the statute books is described by Freitas [13], who echoes much of Lehman-Wilzig's thinking of four years earlier:

> "We will then see an avalanche of cases. We will have robots that have killed humans, robots that have been killed by humans, robots who have stolen state secrets, robots who have been stolen; robots who have taken hostages, robots who have been held hostage and robots who carry illegal drugs across borders. Cases will occur in general when robots damage something or someone, or when a robot is damaged or terminated. In addition, robots will soon enter our homes as machines to save labor, and as machines to provide child care and protection. Eventually these entities will become companions to be loved, defended and protected."

A somewhat different slant on the question of endowing robots with legal rights was taken by Lawrence Solum [27], in a paper in the *North Carolina Law Review*, who used intelligence rather than consciousness as the touchstone by which a robot (or any other artefact) should be judged. Section IV of Solum's paper bears the heading "Should an Artificial Intelligence be Granted the Rights of Constitutional Personhood?" Here Solum considers "the question whether we ought to give an AI constitutional rights, in order to protect its personhood, for the AI's own sake. Imagine, for example, that an AI claims that it cannot be owned under the Thirteenth Amendment to the United States Constitution.[9] A lawyer takes its case, and files a civil rights action on its behalf, against its owner. How should the legal system deal with such a claim?" The strongest objection considered by Solum against recognizing constitutional rights for AIs is that "as artefacts, AIs should never be more than the property of their makers", but as Solum points out this argument is based on dangerous ground. If the makers of AIs are entitled to own them, then "if AIs are persons, then, absent some reason to the contrary, it follows that these persons ought to be slaves. Notice, however, that this argument also would seem to imply that if children are made by their parents, then they too should be slaves."

The question Stone asks in the title of his paper might now, more than thirty years later, reasonably be asked in relation to robots: Should robots have standing? Should there be rights for man-made conscious artefacts?

11 The Robot-Animal Analogy

I now wish to draw an important distinction between the notion that robots *might* be deserving of rights because they have consciousness, and the argument that they *should* have rights because animals have rights. The latter argument has been discussed by Calverley [5] in his paper *Android Science and the Animal Rights Movement: Are There Analogies?*, in which he points out that:

> "Animals are now viewed as having rights or interests sufficient to cause us to ascribe to them moral weight, and they cannot simply be treated as commodities for man's use and benefit. The significance and scope of

[9] The thirteenth amendment abolished and continues to prohibit slavery.

the particular characteristics required for this ascription are still not clearly formulated. Once established they lead to treating animals as moral person[s], but do not necessarily lead to them being viewed as legal persons."

Calverley [5] analyzes the animal rights movement in order to draw some lessons that "could be applied to androids as they develop from crude machines to something closer to 'human'. By looking at how the arguments for animal rights developed over the last thirty years, we can see how the movement has expressed ideas which have substantially changed the views about animals held by many members of society."

Calverley concludes that "not withstanding the divergence between animals and androids, the way in which animals are viewed as moral objects worthy of consideration is a meaningful way to look at androids. If the android exhibits characteristics that are similar to those we are willing to recognize as imparting moral consideration to animals, then treating the android as being with value, at least in Kant's terms,[10] enhances our own self worth. To do something different would demean us as humans…" and that "A long tradition of moral and ethical thought in the area of animal rights gives us some basis upon which to begin to ground speculation concerning the possible attribution of moral rights to androids."

In my view, however, there is an extremely important difference. Animals can suffer and feel pain in ways that robots cannot. This leads me to the view that the animal rights analogy is not a sound one on which to base the notion that robots are deserving of rights. Although I lean towards the view that robots *should* be endowed with certain rights, the basis for my position and the concomitant conclusion regarding this issue are completely different.

12 A Different Perspective

I now wish to introduce a different perspective on how we might view the question of endowing robots with rights and treating them in ethically acceptable ways.

My own argument in support of giving certain rights to robots is not that robots with consciousness should have rights *because* of that consciousness *per se*, but that, because they have consciousness, such robots will be regarded by us in some similar ways to those in which we regard other humans, for example developing affection and even love for robots [22], and that, *because* we will regard such robots with affection and even love, it is reasonable to assume that we will treat robots in other ways similar to those we currently reserve for humans (and, in the case of some people, to pet animals), for example by regarding these robots as having rights. This is a fine distinction but, I believe, an important one.

I also wish to posit a different viewpoint on the question of how and why we should treat conscious robots from an ethical perspective. Christopher Ramey [25] considers a closely related question—how "the human self changes as a consequence of his or her treatment toward androids", and states his concern as being "the ethical effect of android social interaction upon a given human's sense of his or her own being." Ramey discusses how the introduction of an android into one's life would bring about some sort of irreversible change to oneself. "Once one admits androids into one's interpersonal realm, however, one cannot turn back. It is here that one's human mode of existence can be threatened. To treat androids as humans is not to make androids actually human, but it is to make oneself an expanded self. For example, if a person adopted a child, there is an understanding that the child is not biological family, but that parent is fundamentally changed as a person by the adoption of that child into his or her world (*mutatis mutandis* a person who treats an android as part of their world, though never granting that an android is actually human, nevertheless is enriched by that adopted perspective)."

While I concur with Ramey's perception that a person will change as a result of repeated interaction with a robot, I have a different concern. I believe that the way we treat humanlike (artificially) conscious robots will affect those around us by setting our own behaviour towards those robots as an example of how one should treat other human beings. If our children see it as acceptable behaviour from their parents to scream and shout at a robot or to hit it, then, despite the fact that we can program robots to feel no such pain or unhappiness, our children might well come to accept that such behaviour is acceptable in the treatment of human beings. By virtue of their exhibiting consciousness, robots will come to be perceived by many humans, and especially by children, as being in some sense on the same plane as humans. This is the reasoning behind my argument that we should be ethically correct in our treatment of conscious robots—not because the robots would experience virtual pain or virtual unhappiness as result of being hit or shouted at.

13 Conclusions

We have introduced the question of how and why robots should be treated ethically. Consciousness or the lack of it has been cited as the quality that generally determines

[10]Calverley points out that Kant "held views which were premised on the idea that the moral status of animals was derived from the relationship which they held with humans", and quotes Kant's assertion [17] that "our duties towards them are indirect duties to humanity".

whether or not a recipient is deserving of ethical treatment. Some indications of consciousness have been examined, as have two tests that could be applied to detect whether or not a robot possesses (artificial) consciousness.

In discussing the ethical treatment of conscious robots we have examined arguments in favour of giving robots legal rights. One conclusion is that the reason animals are deserving of rights should not be regarded as a basis for arguing that robots are deserving of rights—this because robots cannot feel pain and suffering. But despite this difference between robots and animals it is nevertheless concluded that robots should be endowed with rights and should be treated ethically. This conclusion is based partly on the reasonable expectation that many of the ways in which we will treat artificially conscious robots will be similar to the ways that we humans treat each other, and therefore ethical behaviour towards such robots is merely an extension of such treatment. The conclusion is also based on the lessons we teach each other (and particularly our children) about how to behave towards conscious entities—treating robots in ethically suspect ways will send the message that it is acceptable to treat humans in the same ethically suspect ways.

References

1. Aleksander I (1995) Artificial neuroconsciousness: an update. IWANN, available at http://www.ee.ic.ac.uk/research/neural/publications.iwann.html
2. Arkin R (2007) Lethality and autonomous robots: an ethical stance. In: ICRA '07: IEEE international conference on robotics and automation, Rome
3. Billings L (2007) Rise of roboethics. Grappling with the implications of an artificially intelligent culture. Available at www.seedmagazine.com/news/2007/07/rise_of_roboethics.php?
4. Calverley D (2005) Toward a method for determining the legal status of a conscious machine. In: Chrisley R, Clowes R, Torrance S (eds) Proceedings of the AISB 2005 symposium on next generation approaches to machine consciousness: imagination, development, intersubjectivity, and embodiment. University of Hertfordshire
5. Calverley D (2005) Android science and the animal rights movement: are there analogies? In: Cognitive sciences society workshop, Stresa, Italy, pp 127–136
6. CBC (2007) Ethical code for robots in works, South Korea says. Available at www.cbc.ca/technology/story/2007/03/07/tech-robot-ethics.html
7. Chella A, Manzotti R (2007) Artificial consciousness. Imprint Academic, Exter
8. Clark R, Squire L (1998) Classical conditioning and brain systems: the role of awareness. Science 280(3):77–81
9. Clark R, Squire L (1999) Human eyeblink classical conditioning: effects of manipulating awareness of the stimulus contingencies. Psychol Sci 10:14–18
10. Clowes R, Torrance S, Chrsley R (2007) Machine consciousness: embodiment and imagination. J Conscious Stud 14(7). Special issue
11. De Quincey C (2006) Switched-on consciousness: clarifying what it means. J Conscious Stud 13(4):7–12
12. Economist (2007) Robot wars. The Economist, April 17th 2007. Available at http://www.economist.com/science/tq/displaystory.cfm?story_id=9028041
13. Freitas R Jr (1985) The legal rights of robots. Stud Lawyer 13:54–56. A later draft is available at http://www.freitas.com/Astro/LegalRightsOfRobots.html
14. Freud S (1933) New introductory lectures on psychoanalysis. Internationaler Psychoanalytischer, Leipzig
15. Gallup G (1977) Self-recognition in primates: a comparative approach to the bidirectional properties of consciousness. Am Psychol 329–338
16. Holland O (ed) (2003) Machine consciousness. Imprint Academic, Exeter
17. Kant I (1996) In: Heath P, Schneewind J (eds) Lectures on ethics. Cambridge University Press, New York
18. Koch C (2004) The quest for consciousness: a neurobiological approach. Roberts and Company, Englewood
19. Lehman-Wilzig S (1981) Frankenstein unbound: towards a legal definition of artificial intelligence. Futures 442–457, December
20. Levy D (2006) Robots unlimited. AK Peters, Wellesley
21. Levy D (2007) Intimate relationships with artificial partners. Ph.D. Dissertation, Maastricht University, Maastricht, The Netherlands
22. Levy D (2007) Love and sex with robots. Harper Collins, New York
23. Owen J, Osley R (2007) Bill of rights for abused robots: experts draw up an ethical charter to prevent humans exploiting machines. The Independent, April 1st
24. Pinker S (1997) Could a computer ever be conscious? Available at http://pinker.wjh.harvard.edu/articles/media/1997_08_18_usnewsworldreport.html
25. Ramey C (2005) 'For the sake of others': the 'personal ethics' of human-android interaction. In: Proceedings of COGSCI workshop: toward social mechanisms of android science. Stresa, Italy, July 25–26, pp 137–148
26. Sloman A (1996) A systematic approach to consciousness. (How to avoid talking nonsense?) Summary of a lecture presented by Aaron Sloman at the Royal Society of Arts, 26 February 1996. Available at http://www.cs.bham.ac.uk/~axs/misc/consciousness.rsa.text
27. Solum L (1992) Legal personhood for artificial intelligences. North Carol Law Rev 70:1231–1287
28. Sparrow R, Sparrow L (2006) In the hands of machines? The future of aged care. Minds Mach 16(2):141–161
29. Stone C (1972) Should trees have standing? Toward legal rights for natural objects. South Calif Law Rev 45:450–457
30. Stone C (1997) 'Should trees have standing?' and other essays on law, morals and the environment. Kaufmann, Los Altos
31. Takeno J, Inaba K, Suzuki T (2005) Experiments and examination of mirror image cognition using a small robot. In: The 6th IEEE international symposium on computational intelligence in robotics and automation (CIRA 2005), pp 493–498. Available at http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1554325
32. Torrance S (2004) Could we, should we, create conscious robots? J Health Soc Econ Issues 4(2):43–46
33. Torrance S (2008) Ethics and consciousness in artificial agents. Artif Intell Soc 22(4)
34. Turing A (1950) Computing machinery and intelligence. Mind Q Rev Psychol Philos LIX(236):433–460
35. Veruggio G (2006) EURON roboethics roadmap. Genoa, Scuola di Robotica. Available at http://www.roboethics.org/site/modules/tinycontent/index.php?id=14

David Levy is President of the International Computer Games Association and CEO of Intelligent Toys Ltd. His interest in Artificial Intelligence began with computer chess and other strategy games; then extended to human-computer conversation (he led the winning team in the 1997 Loebner Prize competition); and since 2003 to human-robot relationships. His book "Love and Sex With Robots" and his PhD thesis based on the same research have been widely publicized in the media.