The "uncanny valley" goes on and on

Robotics Meets the Humanities
IEEE ICRA2013 SCIENCE FORUM – Karlsruhe May 10, 2013 (afternoon)

Willard McCarty
10 May 2013
www.mccarty.org.uk/

> …you have built the bridge between science and art. The danger lies in one's impulse to cross it too impetuously, without a proper understanding of the country on the other side.
>
> Sir Herbert Read to D'Arcy Wentworth Thompson (undated, probably 1944), Univ. of St Andrews MS25480

[SLIDE 1] I begin by going back in time, just prior to the beginning of that half-century of "substantial sustained communal scientific enterprise" we call robotics (Mason 2012: 72), with a specific encounter of human and machine that spooked the human: **[SLIDE 2]** engineer George R. Stibitz's, in 1942, with the Anti-Aircraft Predictor designed by Norbert Wiener and Julian Bigelow. The Predictor was intended as a countermeasure against the fast German military aircraft of World War II, whose speed had rendered manual aiming and firing ineffective. As its name suggests, the Predictor calculated in real time the future location of a given enemy airplane, a task which "both in theory and in mechanism", Stibitz wrote in his memoirs half a century later, was "beyond the abilities of the human brain and nerves" (1993: 160). But when he witnessed the Predictor in operation, on 1 July 1942, he jotted in his diary, "the behavior of their instrument is positively uncanny" (Galison 1994: 243). Even long afterward he recalled "an impressive demonstration. It gave me the feeling of having had my mind read" (Stibitz 1993: 204).

I choose this encounter from among the many that robotics inherited because we do not have to infer how we think a person *might* have reacted under

those circumstances. We have it in the eyewitness's own words. We have it of an incident in historical time, and so can make historical inferences. And we have two versions: first, sparingly, known just by the "positively uncanny" sense of it, then with time rendered almost familiar, like "having my mind read". The first, the more immediate and inchoate version is the one to pay attention to.

As a scholar from the humanities, let me use it to put an historical question to you about robotics. Stibitz tells us in his memoirs that the Wiener-Bigelow Predictor was never used: Wiener's mathematical model had assumed observation of the plane for far too long to make the device practical in the field (Stibitz 1993: 205). Subsequently the rapid development of fire-control mechanisms and other automata through the 1940s rendered such encounters ordinary, no longer spooky, no longer even remarkable. So, the question is, what can we conclude from Stibitz's momentary reaction? Do we bolster our pride by comparing our far more advanced devices with that primitive one, feeling perhaps some affection for a simpler age? Or (I think far better) do we probe what that incident in 1942 shares with like ones now – incidents that roboticist Masahiro Mori three decades later famously called "Bukimi no Tani Genshō" (不気味の谷現象), "the uncanny valley phenomenon"?[1] A moment's reflection will tell you that we keep finding ourselves in that valley. Why is that? Is the uncanny valley a problem roboticists should strive to overcome, or should we study it together, as Mori suggested, "so that through robotics research" – in collaboration with the humanities, I would add – "we can begin to understand what makes us human" (2012/1970: 100)?

Professors Nakamura and Sandini have brought us together with a sense of urgency, in their belief that the humanities might play a crucial role in the development of robotics. I share this belief, though I know very little about

robotics. My sense of urgency comes from the complementary belief that robotics has a crucial role in the development of the increasingly digital humanities, indeed is already affecting these disciplines, though not as intelligently and practically as could be. My argument here is that the mutual benefit, of robotics for the humanities and the humanities for robotics, originates precisely where they meet, in Mori's uncanny valley.

Like many who watch television, my interest has been caught by recent and quite thoughtful explorations of encounter between humans and those plausible entities we are capable of imagining though not yet building – those whom we might call the children of *Blade Runner* (1982), which itself is a child of many other depictions and constructions of automata dating back to ancient times. Within the last year or so, for example, television series have given us the artificial humanoid that replaces a boyfriend killed in an automobile crash, in an episode from the British *Black Mirror* (2013); the "hubots" in the Swedish *Äkta Människor* ("Real Humans", 2012) **[SLIDE3]**; and a pilot for the cancelled but still indicative American series *Beautiful People* (2012). I cite these popular entertainments as evidence of a long-running anxiety well known to the artificial intelligentsia (and shared by some of its members).[2] But I see these entertainments in the light of intelligent responses to this anxiety from the humanities, for example the *On the Human* project of the U. S. National Humanities Center, which was convened from 2006 to 2009 to consider the perceived threat from the sciences to our idea of the human; the similar conference held at Stanford in 1987, "Humans, Animals, Machines: Boundaries and Projections";[3] and Katherine Hayles' influential book *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics* (2008).

Examples could be multiplied manyfold. But these few are enough to suggest the rich history of imaginative pondering and scholarly reflection on the

anxiety of uncanniness I will be exploring. For reasons I will get to, however, I want to back away from the notion of the posthuman condition we are supposedly in to take a much longer historical view on how our inventions and discoveries affect us. I would speak of the *human*, not posthuman, condition because the human, as I will explain, is not a stable state but an ongoing process.

Let me return to my chosen human-machine encounter, Stibitz with the Anti-Aircraft Predictor. I want to use this encounter as a springboard to get nearer to the beginnings of a very good conversation between us.

In his article on Wiener's wartime research, historian Peter Galison cites philosopher Stanley Cavell to get to the bottom of engineer Stibitz's reaction to mathematician Wiener's and engineer Bigelow's hardware. Galison quotes from Cavell's Tanner lecture, "The Uncanniness of the Ordinary" (1986), in which he takes up psychiatrist Sigmund Freud's analysis of the German Romantic author E. A. Hoffmann's story *Der Sandmann* (*The Sandman*, 1816). I will get to that story in a moment, but before I do, let us note what has just happened: we have travelled quickly, thanks to an historian, from philosophical reflections on a human reaction to a robotic device, briefly through the cultured mind of a psychologist, into an imaginative work of literature. In other words, we are just where we need to be, amidst a dense network of disciplines interrelated by a shared interest in the uncanny. But what can we make of it?

In *The Sandman* – forgive this all-too-brief synopsis of a rich and complex story – a young fellow named Nathaniel is turned away from his true-love Clara by looking through a magical glass at a strikingly beautiful automaton, Olympia, who warms to his sight as he gazes on her. **[SLIDE 4]** He becomes obsessed by that sight, attends a ball where he dances with her – and then is

driven insane by witnessing her disassembly. Later he is cured and restored to Clara's love. But then, finding the glass in his pocket quite by accident, he happens to glimpse Clara through it. His madness returns. He leaps to his death from a church steeple.

Cavell draws a provisional moral from this story: that by means of such a glass (imagined just prior to the emergence of photography in 1822), indeed by means of any instrument that represents the world, "no marks or features or criteria or rhetoric" can be discerned "to tell the difference between [the animate and the inanimate]". But Cavell does not claim "that the difference is unknowable or undecidable. "On the contrary," he says, "the difference is the basis of everything there is for human beings to know, or say decide (like deciding to live), and to decide on no basis beyond or beside or beneath ourselves." (1986: 89) Thus the uncanny: we see but cannot explain, even to ourselves, especially not scientifically.

I suppose that the practical engineer would regard the experience of the uncanny as a momentary stage in a process of development toward a life-improving outcome, such as a prosthetic hand. But Cavell's Hoffmann urges us to look at this process in reverse, from the accepted, perfectly ordinary outcome back to that moment of uncanniness. Here philosophy proves itself worth every moment of our attention: Cavell wants us to look even closer, "to the fantastic in what human beings will accustom themselves to" (1986: 84) – to "the uncanniness of the ordinary". He wants us in effect to question the life we are living, i.e. to question the shape and qualities of life itself, not to improve our lives as we are now living them.

A scary prospect. The crucial question is, what is the value of putting oneself at risk? You will recall that Mori suggests a better understanding of "what makes us human". But what does this amount to? One answer is suggested by

the Russian formalist critic, writer and pamphleteer Viktor Borisovich Shklovsky, who in a 1917 essay famous to literary scholars, "Art as Technique", argued that "as perception becomes habitual, it becomes automatic…. And so life is reckoned as nothing." Everything becomes the dull ordinary accepted everyday that conceals the not-automatic, i.e. the uncanny, from sight. The function of art, Shklovsky argued, is to *defamiliarize* the world, make objects *unfamiliar*, and so recover "the sensation of life" (1965/1917: 11-12), i.e. the primal uncanniness of the ordinary. In 1935, when television was so new that it piggybacked on radio, the art historian Rudolf Arnheim leapt for the moment: **[SLIDE5]** "The new gadget seems magical and mysterious. It arouses curiosity: How does it work? What does it do to us? To be sure, when the television sets will have appeared on the birthday tables and under the Christmas trees, curiosity will abate. Mystery asks for explanation only as long as it is new. Let us take advantage of the propitious moment" (1957/1935: 188).

What then happens? What does that moment yield?

Let us consider once again the replica of boyfriend Ash in *Black Mirror* **[SLIDE 6]** and the hubots of *Äkta Människor*. Their presence in those fictional worlds quickly leads to all manner of disturbing reactions. These, though on a direct, intimate level, lead us to the great anthropological question that Immanuel Kant said included all of philosophy: "What is man?" (1992/1800: 538). As you might expect, this is a question very much older than Kant. Its history has been traced, for example, by Roger Smith in *Being Human: Historical Knowledge and the Creation of Human Nature* (2007), which as I hinted earlier, demonstrates that the human is something we make and are perpetually remaking. In other words, to see this point in all its sharpness the most helpful arguments are those for which being human is continual invention, not something previously determined that we are given to be, not a

suit we are given to wear (and perhaps change with changing fashions), not a kind of received knowledge. In his brilliant analysis of our psycho-social condition, *Modernity and Self-Identity: Self and Society in the Late Modern Age* (1991), sociologist Anthony Giddens vividly depicts the perilously negotiated process of "going-on being" in our reflexive construction of self (1991: 39). But Giddens draws on the developmental psychologist D. W. Winnicott to argue further that going-on being is done by identifying and repudiating the "not-me" – e.g. the enemy whom we need in order to become who we are,[4] the not-human but eerily humanoid automaton that you roboticists are hard at work on. This discovery and repudiation of the not-me is part and parcel of how the autonomous self is formed and reformed (Giddens 1991: 42). In other words, the human is a problem perpetually being solved. Thus Smith once more: "The word 'human' denotes something coming into existence in historical processes" (2007: 7).

Now I need Freud once again, this time to bring in the sciences, and so robotics, as one of those historical processes. Just prior to publishing the essay Cavell draws on, "Das Unheimliche" ("The Uncanny") in 1919, Freud twice argued that scientific research had precipitated three great crises in human self-conception, or as he put it, three "great outrages" ("große Kränkungen") to human self-love: first by Copernican cosmology, which de-centred humankind; then by Darwinian evolution, which de-throned it by setting in motion discoveries of how intimately we belong to life; and finally by his own psychoanalysis, which showed we are not even masters of our own house.[5] Less often noticed is his suggestion (implicit in the German *Kränkung*, from *krank*, "ill, sick, diseased") that these dis-easings of the mind can be turned to therapeutic effect. We are apt to see only the physician here, but Freud was in fact showing his inheritance from the whole tradition of the physical sciences, which at least from Bacon and Galileo in the 17th Century had identified the

cognitively and morally curative function of science. Both Bacon and Galileo wrote of scientific method as a defence against what Bacon called "the sciences as one would"[6] – i.e. fanciful or capricious knowledge tainted by human weakness and corruption of mind, among which faults self-love is basic. In historian Alastair Crombie's words, science for them was a corrective, restorative force: "the moral enterprise of freedom for the enquiring mind … a therapeutic experience offering perhaps the greatest moral contribution of science to mankind."[7] We now know, contrary to earlier arguments pitting science against religion in its origins, that for these scientists such therapy was aimed at a reversal of the Fall of Man, specifically a restoration of cognitively diseased humankind to the unclouded intelligence Adam was said to have before the Fall (McCarty 2012: 9-11). The religious language has dropped away from the scientific literature, but much the same moral imperative can be traced forward into current arguments, especially in the biological and computational sciences, which school us to humility. Freud's argument is thus radically incomplete: outraging our self-regard in order to rid us of it is built into the scientific programme.

In a sense our machines and our bodies pose the identical threat, and for this reason embodied machines, i.e. humanoid robots, are the greatest outrage of all and, I think, hold a greater promise than disembodied AI. For this reason I turn to animals to flesh out, as it were, the compelling threat.

René Descartes, in his *Discourse on Method* (1637), came very close to equating bodies and machines by arguing that animals are essentially machines made of flesh (1998/1637: 31-2 [56-7]). He did this in order to establish our guarantee of uniqueness, of separateness from animals, in our capacity to reason. But simultaneously our likeness to the other known members of the *Hominidae* or "Great Apes" was being noticed all across Europe in consequence of the Age of Exploration. In 1641 the Dutch physician Dr Nicolaes Tulp (whom

Rembrandt painted giving an anatomy lesson and from whom Linnaeus and many others borrowed) published the first accurate description of an ape, **[SLIDE 7]** which he said the natives called an Orang-utang, "sive homo sylvestris" (or man of the woods). It was, he commented, so physiologically similar to humans "that it would be difficult to find one egg more like another".[8] A century later, in his *Histoire Naturelle* (1766), the Comte de Buffon expressed the ambivalence perfectly: "This orang-outang", he wrote, "…is only a brute, but a brute of a kind so singular, that man cannot behold it without contemplating himself…" (1785/1766: 41). For Buffon and many others the inference one had to draw, he went on to write, was clear: that "[man's] body is not the most essential part of his nature." Can one feel some nervousness here? Certainly in oneself, if not in him.

Earlier that century, but with the greater license afforded a writer of fiction, Jonathan Swift depicted a very different result: the devastating impact on human self-understanding that the *homo sylvestris* could have. In the last part of *Gulliver's Travels* (1726) Lemuel Gulliver, bathing in a stream, is literally forced to embrace his lustful, brutish nature by a female Yahoo, who "enflamed by desire" leaps into the stream and embraces him. **[SLIDE 8]** His position in Houyhnhm society, among the creatures of perfect reason by whom he has striven to find acceptance, is suddenly made untenable. He is ejected by them and so is forced to return home, where he finds himself repelled by the smell of his wife and prefers to bed down with his horses. He is, you might say, driven insane by self-knowledge. It is worth asking why fulfilling the Delphic commandment could be so lethal.

When 150 years later Charles Darwin surveyed the biological continuities joining humans to the other great apes in *The Descent of Man*, he noted the many attempts to argue that "man is separated through his mental faculties by an impassable barrier from all the lower animals"; "I formerly made a

collection of above a score of such aphorisms," he wrote, "but they are not worth giving, as their wide difference and number prove the difficulty, if not impossibility, of the attempt" (1871: 47). Although his conclusion has widespread support today, the problem remains hotly debated.[9]

There are, then, two quite different reactions to the same body of evidence concerning both animals and machines: terror, however muffled, and wonder, however sentimentalized. I am going to assert without proof (because I do not know how to prove any such thing) that the uncanny lies at the root of both – in consequence of which both terror and wonder can be mined for the dangerous self-knowledge they yield.

I have argued along with Crombie and others that science is a moral undertaking, a corrective therapy, and I have argued with Giddens that this therapy works by confronting us with that which we are not. The peril here is the inability to see beyond the negation. Nietzsche's warning comes to mind: "He who fights with monsters should be careful lest he thereby become a monster. And if you gaze long into an abyss, the abyss will also gaze into you."[10] Hence, I would suggest, the odd excesses of attack against the human that we find in the writings of several scientists and others whose zeal is to prove that we amount to nothing, or less than nothing.[11] Biological anthropologist Melvin Konner sums up this attack: "It would seem", he concludes, "that we are sorted to a pulp, caught in a vise made, on the one side, of the increasing power of evolutionary biology… and, on the other, of the relentless duplication of human mental faculties by increasingly subtle and complex machines." He asks, "So what is left of us?" (1991: 120).

Yet again Kant's question, but with an ironic twist: What *are* we? And let me suggest an ironic response arising from the abyss these scientists have gazed

into – physicist Steven Weinberg's pointless, hostile universe,[12] for example: their discovery is that *human*kind isn't there. Konner refers to "the residue of uniqueness" left to us by such zealous pursuits. But, literary critic Jerome McGann points out, "what is left of us" is not *residual*. It is, as he says in biblical language, "the hem of a quantum garment" (2004: 201), one touch of which opens up a new world of possibilities.[13] With a slight stretch this is what Mori said the uncanny valley was all about: discovering more of human nature, or more human natures.

Thus I bring to an end the part of my talk that for a humanist comes more or less naturally. Let me conclude, however, by moving on to the hard part: to the future. What are we going to do about all this? The future to make happen, I suggest, is not one in which robotics and the humanities swap knowledge and devices across a two-cultured chasm. It is one in which they collaborate.

I assume that within robotics the knowledge which really matters is, as aeronautics engineer Walter Vincenti and historian Thomas Smith have separately argued, *indigenous* knowledge, i.e. in Smith's words, "the *fusion* of science with technology… [in] the distinctively social enterprise called R & D".[14] In contrast the humanities are rather asocial in their research practices, which are normally solitary rather than collaborative, communicated from the mind of the researcher to the mind of the reader via writing. But the introduction of computing into the humanities has meant that research is sometimes worked out in the design and construction of devices, such as software systems. Hence collaboration in the true sense has begun to develop, in the first instance out of necessity, because most scholars are technically incapable. But as a result some of them have likewise discovered a knowledge indigenous to the social work of collaboration centred on a common object (cf.

Deegan and McCarty 2012). Hence the stage is set for doing something together. But what?

The action, I've argued, is and will remain in the uncanny valley, at the point of contact between the best you are able to do and, for the humanities, the most adventurous ideas of text, image and sound. Our present construct to name this point of contact, *interface* – "a surface lying between two portions of matter or space, and forming their common boundary" – seems to me too much of a barrier; *skin* would be better if by that we meant living tissue with sensory receptors. To date we in the digital humanities have chiefly constructed mechanisms for delivery of codified knowledge across that interface, vending machines of knowledge, **[SLIDE 9]** leaving the core activity of interpretation, ideally at the speed of thought, untouched – and untouchable. We have to a large degree separated ourselves from the machine by relying on the ancient role of the servant or slave and thereby cast ourselves into the role of its master. For reasons I cannot amplify here, giving new life to the old master-servant relationship is a big mistake: it wastes the potential of both. The first question, in other words, is the social model we adopt for robotics in humanities research, or more accurately the different social models to suit the different, and mostly unstudied, circumstances in which scholarship is done.

Attempting to envision what sort of scholarly companion I might like to have, I am attracted by the primatologist Frans de Waal's description of the Bonobo ape. **[SLIDE 10]** Could we someday say of a hubot what de Waal says of the Bonobo, that

> When the lively, penetrating eyes lock with ours and challenge us to reveal who we are, we know right away that we are not looking at a "mere" animal, but a creature of considerable intellect with a secure sense of its place in the world. (1997: 1)

Certainly not in my lifetime, probably not in the lifetime of anyone here or soon to be at such a conference. But is that not what we are attempting? Would it not be exhilarating to try?

It's not as if such thoughts are new. The dream is ancient, the connection intimate. Research into "'crossing the lines' from one domain to another, and [seeing] what can then emerge" (as Chrystopher Nehaniv remarked), bringing together humans, animals and robots, has been going on for some time.[15] But what I think *is* new in what I have said lies in the humanities' view of the human not as "an ever-fixéd mark / That looks on tempests and is never shaken" but as an historical process of becoming. On the horizon of a humanistically educated imagination is not siege-warfare against the citadel of man by "monsters from the Id", or ourselves "all watched over by machines of loving grace",[16] or anything of the kind, but ourselves in contest with ourselves with the help of the arts and humanities, robotics, primatology and other sciences. The real danger won't come from hubots, rather from our disciplinary isolation from each other, e.g. from cultural theorists doing their own thing on the basis of what they think robotics aims to do while roboticists do their own thing on the basis of received knowledge or just untutored intuition. The dream I think we could share, if we do not share it already, is an extended version of Leibniz's *Theoreticos Empiricis felici connubio zu conjungiren*, as he said in a typical mixture of Latin and German, "to join theorists and empirics in a happy marriage" (Burke 2000: 16f). **[SLIDE 11]**

Works cited.

Agamben, Giorgio. 2004/2002. *The Open: Man and Animal*. Trans. Kevin Attell. Stanford CA: Stanford University Press.

Arnheim, Rudolf. 1957/1935. "A Forecast of Television". *Film as Art*. 188-98. Berkeley: University of California Press.

Baird, Davis. 2004. *Thing Knowledge: A Philosophy of Scientific Instruments*. Berkeley: University of California Press.

Bingham, Paul M. 1999. "Human Uniqueness: A General Theory". *The Quarterly Review of Biology* 74.2: 133-69.

Boesch, Christophe. 2012. "From Material to Symbolic Culture in Primates". *Oxford Handbook of Culture and Psychology*. Ed. Jaan Valsiner. 677-92. Oxford: Oxford University Press.

Brautigan, Richard. 1967. *All Watched Over by Machines of Loving Grace*. San Francisco: The Communication Company.

Bruner, Jerome S. 1956. "Freud and the Image of Man". *American Psychologist* 11.9: 463-6.

Cavell, Stanley. 1986. "The Uncanniness of the Ordinary". The Tanner Lectures on Human Values, Stanford University, 3 and 8 April. tannerlectures.utah.edu/lectures/documents/cavell88.pdf (3/4/13).

Crombie, A. C. 1994. *Styles of Scientific Thinking in the European Tradition. The History of Argument and Explanation especially in the Mathematical and Biomedical Sciences and Arts*. 3 vols. London: Duckworth.

Cetina, Karin Knorr. 1991. "Epistemic Cultures: Forms of Reason in Science". *History of Political Economy* 23.1: 105-22.

Darwin, Charles. 1871. *The Descent of Man, and Selection in Relation to Sex*. Vol. I. New York: D. Appleton and Company.

Daston, Lorraine, ed. 2004. *Things that Talk: Object Lessons from Art and Science*. New York: Zone Books.

--- and Gregg Mitman, eds. 2005. *Thinking with Animals: New Perspectives on Anthropomorphism*. New York: Columbia University Press.

Dautenhahn, Kerstin, Alan H. Bond, Lola Cañamero and Bruce Edmonds, eds. 2002. *Socially Intelligent Agents: Creating Relationships with Computers and Robots*. Multiagent Systems, Artificial Societies, and Simulated Organizations. Boston MA: Kluwer Academic Publishers.

--- and Chrystopher L. Nehaniv, eds. 2002. *Imitation in Animals and Artifacts*. Complex Adaptive Systems. Cambridge MA: MIT Press.

de Buffon, Comte. 1785/1766. *Natural History, General and Particular*. Trans. William Smellie. 2nd edn. Vol. VIII. London: W. Strahan and T. Cadell.

de Waal, Frans and Frans Lansing. 1997. *Bonobo: The forgotten Ape*. Berkeley CA: University of California Press.

de Waal, Frans, Robert Wright, Christine M. Korsgaard, Philip Kitcher and Peter Singer. 2006. *Primates and Philosophers: How Morality Evolved*. Ed. Stephen Macedo and Josiah Ober. Princeton: Princeton University Press.

Deegan, Marilyn and Willard McCarty, eds. 2012. *Collaborative Research in Digital Humanities*. Farnham, Surrey: Ashgate.

Dening, Greg. 2002. "Performing on the Beaches of the Mind: An Essay". *History and Theory* 41: 1-24.

Descartes, René. 1998/1637. *Discourse on Method and Meditations on First Philosophy*. Trans. Donald A. Cress. Indianapolis IN: Hackett Publishing Company.

Freud, Sigmund. 1919. "Das Unheimliche". *Imago* 5/6: 297-324.

---. 1920/1917a. "One of the Difficulties of Psychoanalysis". Trans. Joan Riviere. *International Journal of Psychoanalysis* 1: 17-23. Originally published as "Eine Schwierigkeit der Psychoanalyse", *Imago* 5.1: 1-7.

---. 1920/1917b. *A General Introduction to Psychoanalysis*. Trans. G. Stanley Hall. New York: Boni & Liveright. Originally published as *Vorlesungen zur Einführung in die Psychoanalyse*, Berlin: Internationaler Psychoanalytischer Verlag, 1917.

Galison, Peter. 1994. "The Ontology of the Enemy: Norbert Wiener and the Cybernetic Vision". *Critical Inquiry* 21.1: 228-66.

---. 1997. *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press.

Giddens, Anthony. 1991. *Modernity and Self-Identity: Self and Society in the Late Modern Age*. London: Polity.

Gorman, Michael E., ed. 2011. *Trading Zones and Interactional Expertise: Creating New Kinds of Collaboration*. Cambridge MA: MIT Press.

Gray, Jeffrey A. 1991. "Fear, Panic, and Anxiety: What's in a Name?" *Psychological Inquiry* 2.1: 77-8.

Hacking, Ian. 1995. *Rewriting the Soul: Multiple Personality and the Sciences of Memory*. Princeton: Princeton University Press.

Haraway, Donna. 2008. *When Species Meet*. Posthumanities, Vol. 3. Minneapolis MN: University of Minnesota Press.

Henderson, David. 1939. *Psychopathic States*. Thomas William Salmon Memorial Lecture. London: Chapman and Hall.

Horvitz, Eric and Bart Selman. 2011. "Interim Report from the Panel Chairs, AAAI Presidential Panel on Long-Term AI Futures, August 2009". www.aaai.org/Organization/Panel/panel-note.pdf (4/4/13)

Humphreys, Paul. 2004. *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford: Oxford University Press.

Kant, Immanuel. 1992/1800. *Lectures on Logic.* Ed. and trans. J. Michael Young. Cambridge: Cambridge University Press.

Keller, Evelyn Fox. "Language and Ideology in Evolutionary Theory: Reading Cultural Norms into Natural Law". In Sheehan and Sosna 1991: 85-102.

Konner, Melvin. 1991. "Human Nature and Culture: Biology and the Residue of Uniqueness". In Sheehan and Sosna 1991: 103-24.

Lakatos, Imre. 1970. "Falsification and the Methodology of Scientific Research Programmes". In *Criticism and the Growth of Knowledge*. Ed. Imre Lakatos and Alan Musgrave. Cambridge: Cambridge University Press.

Lloyd, G. E. R. 2007. *Cognitive Variations: Reflections on the Unity and Diversity of the Human Mind*. Oxford: Clarendon Press.

---. 2012. *Being, Humanity & Understanding*. Oxford: Clarendon Press.

Mason, Matthew T. 2012. "Creation Myths: The beginnings of robotics research". *IEEE Robotics and Automation Magazine* 19.2 (June): 72-7.

Markoff, John. 2009. "Scientists Worry Machines May Outsmart Man". *New York Times*, 26 July. www.nytimes.com/2009/07/26/science/26robot.html (4/4/13).

Mazis, Glen A. 2008. *Humans, Animals, Machines: Blurring Boundaries*. Albany NY: State University of New York Press.

Mazlish, Bruce. 1967. "The Fourth Discontinuity". *Technology and Culture* 8.1: 1-15.

McCarty, Willard. 2012. "The Residue of Uniqueness". *Historical Social Research / Historische Sozialforschung* 37.3: 24-45.

McGann, Jerome. 2004. "Marking Texts of Many Dimensions". In *A Companion to Digital Humanities*. Ed. Susan Schreibman, Ray Siemens and John Unsworth. Oxford: Blackwell.

Midgley, Mary. 1985. *Evolution as a Religion: Strange Hopes and Stranger Fears*. London: Routledge.

Monod, Jacques. 1972/1970. *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology.* Trans. Austryn Wainhouse. London: Collins.

Mori, Masahiro. 2012/1970. "The Uncanny Valley". Trans. Karl F. McDorman and Norri Kageki. *IEEE Robotics and Automation Magazine* 19.2 (June): 98-100.

Nehaniv, Chrystopher L., ed. 1998. *Computation for Metaphors, Analogy, and Agents*. Berlin: Springer-Verlag.

--- and Kerstin Dautenhahn, eds. 2007. *Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communication Dimensions*. Cambridge: Cambridge University Press.

Nietzsche, Friedrich. 1921/1886. *Beyond Good and Evil*. Trans. Helen Zimmern. New York: Modern Library. Originally published as *Jeneseits von Gut und Böse*, in *Nietzsche's Werke*, Band VII, Stuttgart: Alfred Kröner.

Payne, Junaidi and J. Cede Prudente. 2008. *Orangutans: Behavior, Ecology, and Conservation*. Cambridge MA: MIT Press.

Penn, Derek C., Keith J. Holyoak and Daniel J. Povinelli. 2008. "Darwin's mistake: Explaining the discontinuity between human and nonhuman minds". *Behavioral and Brain Sciences* 31: 109-178.

Rorty, Richard. 1979. *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.

Schulz, Bruno. 1998/1935. "An Essay for S. I. Witkiewicz". *The Collected Works of Bruno Schulz*. Ed. Jerzy Ficowski. London: Picador.

Sheehan, James J. and Morton Sosna, eds. 1991. *The Boundaries of Humanity: Humans, Animals, Machines*. Berkeley CA: University of California Press.

Shklovsky, Viktor. 1965/1917. "Art as Technique". In *Russian Formalist Criticism: Four Essays*. Trans. Lee T. Lemon and Marion J. Reis. 3-24. Loncoln NB: University of Nebraska Press.

Smith, N. Kemp. 1957. "Fear: Its nature and diverse uses". *Philosophy* 32.120: 3-20.

Smith, Roger. 2007. *Being Human: Historical Knowledge and the Creation of Human Nature*. New York: Columbia University Press.

Smith, Thomas M. 1976. "Project Whirlwind: An Unorthodox Development Project". *Technology and Culture* 17.3: 447-64.

Stibitz, George R. 1993. *The Zeroth Generation. A Scientist's Recollections (1937-1955) from the early binary relay digital computers at Bell Telephone Laboratory and OSRD to a fledgling minicomputer at the Barber-Colman Company.* Privately printed.

Tulp, Nicolaes. 1641. *Observationum Medicarum. Libri Tres. Cum aeneis figuris*. Amsterdam: Ludovicus Elzevirium.

Vaesen, Krist. 2012. "The cognitive bases of human tool use". *Behavioral and Brain Sciences* 35: 203-62.

Vincenti, Walter G. 1990. *What Engineers Know and How They Know It: Analytical Studies from Aeronautical History*. Baltimore: Johns Hopkins University Press.

Viveiros de Castro, Eduardo. 1992. *From the Enemy's Point of View: Humanity and Divinity in an Amazonian Society*. Trans. Catherine V. Howard. Chicago: University of Chicago Press.

Weinberg, Steven. 1983/1977. *The First Three Minutes: A modern view of the origin of the universe.* London: Flamingo.

Williams, Raymond. 1983/1976. *Keywords: A vocabulary of culture and society.* Rev. edn. New York: Oxford University Press.

Notes.

[1]不気味 (bukimi), "uncanny", consisting of 不 (bu), "un-", 気 (ki), "energy, life-force, feeling", 味 (mi), "flavour, taste", suggesting a weird feeling in reaction to something not named; の (no), "of"; 谷 (tani), "valley"; 現象 (gensho), "phenomenon". My thanks to Reiko Takeda for the explanation and to Anton Bogdanovych for alerting me to Mori's hypothesis initially.

[2] Note, for example, the meeting convened by the American Association of Artificial Intelligence, 21-22 February 2009 at Asilomar, California, to discuss "potential long-term societal influences of AI research and development", e.g. "how… AI successes in multiple realms and venues [might] lead to significant or perhaps even disruptive social changes", research.microsoft.com/en-us/um/people/horvitz/AAAI_Presidential_Panel_2008-2009.htm (4/4/13); see Horvitz and Selman 2011; Markoff 2009. My point is effectively made by the fact that the AAAI felt the need to commission a report and convene such a meeting.

[3] For *On the Human* see onthehuman.org and the Summer 2009 issue of *Daedalus* that resulted from it; for the Stanford conference, Sheehan and Sosna 1991.

[4] See esp. Viveiros de Castro 1992; Galison 1994.

[5] Freud 1920/1917a and 1920/1917b; cf. Mazlish 1967: 2 and Bruner 1956. Note, however, that I am arguing for the value of the disturbing human-machine conjunction.

[6] id quod generat ad quod vult scientias, in *Novum Organum*, I.xlix.

[7] Crombie 1994: 8; for Bacon also see 1208-9 and 1572-86.

[8] "cum homine similitudinem: ut vix ovum ovo videris similis", Tulp 1641: 3.56, p. 274; cf. de Waal 1997: 7.

[9] See e.g. Penn, Holyoak and Povinelli 2008 with the numerous responses that follow; Bingham 1999; Daston and Mitman 2005; de Waal et al 2006; Haraway 2008; Vaesen 2012.

10 "Wer mit Ungeheuern kämpft, mag zusehn, dass er nicht dabei zum Ungeheuer wird. Und wenn du lange in einen Abgrund blickst, blickt der Abgrund auch in dich hinein." Nietzsche 1921/1886: 146.

11 Examples are Humphreys 2004: 8 and 156; Monod 1972/1960: 160; Weinberg 1993/1977: 155. For a revealing examination of the language used in such attacks see Keller 1991: 85-102 and on Weinberg and Monod, Midgley 2002/1985: 86-94.

12 Quoted by Midgley 2002/1985: 86-7.

[13] McGann 2004: 201; see Matthew 9:20.

[14] Vincenti 1990; Smith 1976: 463; cf. Baird 2004; Daston 2004.

15 Nehaniv 1998; see also Nehaniv and Dautenhahn 2007; Dautenhahn, Bond, Cañamero and Edmonds 2002; Dautenhahn and Nehaniv 2002; cf Mazis 2008. These edited collections came to my attention at the last minute and so only represent belated recognition of an active area of research.

16 Forbidden Planet (1956); Brautigan 1967.