

Please share your stories about how Open Access to this article benefits you.

Beyond the skin bag: on the moral
responsibility of extended agencies

by F. Allan Hanson

2009

This is the author's accepted manuscript version of the article, made available with the permission of the publisher. The original published version can be found at the link below.

Hanson, F. Allan "Beyond the Skin Bag: On the Moral Responsibility of Extended Agencies," *Ethics and Information Technology* 11:91-99, 2009

Published version: <http://dx.doi.org/10.1007/s10676-009-9184-z>

Terms of Use: <http://www2.ku.edu/~scholar/docs/license.shtml>

Beyond the Skin Bag: On the Moral Responsibility of Extended Agencies

Acknowledgements footnote: I am grateful to Richard Cole, Richard DeGeorge, Anthony Genova, Louise Hanson, Deborah Johnson, Rex Martin, and Evan Selinger for stimulating criticisms and suggestions as I prepared this essay.

F. Allan Hanson
University of Kansas
November 29, 2008

The growing prominence of computers in contemporary life, often seemingly with minds of their own, invites rethinking the question of moral responsibility. Eurotransplant is a highly computerized system that generates priority lists of recipients for organs on the basis of compatibility, age, waiting time, distance between donor and recipient, and balance among the several participating countries (Tufts 1996:1326). Ethical considerations are clearly involved, for Eurotransplant seeks to achieve “an optimal proportion between justice and efficiency—the medical ethical criteria,” and it is generally thought that it realizes these objectives better than previous procedures that relied entirely on human evaluations (De Meester, Persijn, Claas and Frei 2000:333). Where does the moral responsibility for the priorities generated by the computerized system lie: with the human programmers and users alone, or also with the databases and computer hardware and software?

It does not seem controversial to say that the moral responsibility for an act lies with the subject that carried it out. Thus our question about Eurotransplant raises the further question, what is the subject? This question has no absolute, definitive answer because the subject is a social construct that varies cross-culturally and historically. For some time the subject has been understood in our society to be the human individual. From that perspective, moral responsibility involved in Eurotransplant, or anything else, is limited to the humans involved. “After all,” it might be said, “if something goes wrong, we don’t punish the computers.” True

enough. But we very well may *blame* them, and blaming, no less than punishing, is generally taken to be a marker of moral responsibility. This gives a glimpse of the possibility that moral responsibility might include things beyond human individuals. But if there is to be such a thing, it must be grounded in a different concept of the subject. Recent social theory has proposed a way of thinking about the subject that is indeed conducive to a broader concept of moral responsibility. The purpose of this essay is to review that alternative view of the subject, to imagine what a concept of responsibility derived from it might look like, and to consider what advantages it might bring and the objections that are sure to be raised against it.¹

Individualism and extended agency

The view of the subject as only the human individual is known as methodological individualism. This theory holds that subjects are human beings entirely contained in their “skin bags” (Clark 2003), that maintain their identity and integrity over time and in their various dealings with other individuals and things of all sorts. Computers, other machines, tools, and animals on this view are understood as no more than objects that people encounter and manipulate in the course of their actions (Flew 1995, Jones 2000). An alternative view that has recently come on the theoretical scene goes under names such as extended agency, actor-network theory, distributed cognition, and cyborg. It maintains that most actions are undertaken by subjects that extend beyond the human individual to include other human beings and any number or kind of nonhuman entities (Clark 2003; Hanson 2004, 2007; Hutchins 1995, Haraway 1991, Law 1999; Selinger and Engström 2007).

A student looks up a book in a library’s online catalog. The methodological individualist insists that the subject in this case—the agent that carries out the task—is the student alone. Extended agency theory holds that the deed is accomplished by the combined entity consisting of

the student, the database of library holdings, and the automated hardware and software that put the two in contact. The basic reasoning behind this extension of agency beyond the individual is that if an action can be accomplished only with the collusion of a variety of human and nonhuman participants, then the subject or agency that carries out the action cannot be limited to the human component but must consist of all of them. The essential participation of nonhuman elements in high-tech applications such as Eurotransplant, robotic assembly lines or MRI procedures is especially clear. However, it is obvious that many activities—herding sheep with dogs, mowing a lawn, even driving a nail—require the participation of nonhuman beings or tools as well as people. Hence, although the growing importance of computers has been the most important factor in the recent development of extended agency theory, that theory applies to actions of all sorts.

What does this have to do with moral responsibility? For methodological individualists the issue is simple. Every action is ultimately attributable to human individuals, and whatever role computers, robots, dogs, lawnmowers, or hammers may play is ancillary. Obviously, if only human individuals act, then only human individuals can be responsible for those acts, and that's the end of it. This view may be called *moral individualism*, the ethical twin of methodological individualism. When methodological individualist assumptions are deeply rooted, the mind is closed to any alternative to moral individualism.

If, however, one is open to extended agency theory's redefinition of the nature of acting subjects, then one can at least entertain the idea that responsibility may apply to nonhuman as well as human beings. This view may be called *joint responsibility*. Given the traditional dominance of individualism, both methodological and moral, extended agency as a theory of action requires something of a paradigm shift in conventional ways of thinking. And the notion

of joint responsibility demands even more, for many people find it counterintuitive to attribute moral responsibility to objects, animals, and other nonhuman things. To entertain the plausibility of such a paradigm shift requires closer consideration of the evidence and how it might be construed. Because moral individualism and joint responsibility are corollaries of the competing moral individualist and extended agency theories of action, first we will examine the argument for the subject as extended agency.

It is no ineluctable fact of nature that the subject is and must be the human individual. The subject, as I have said, is a social construct, and a great deal of evidence from various times and places indicates that the acting agent is frequently understood to be something other than the human individual. One common notion is that the things people do are often not of their own making but result from external forces working through them. Possession by demons and words from the deceased spoken by spirit mediums are obvious examples. The Sioux Indian Black Elk, who had powers of healing and clairvoyance, said: “Of course it was not I who cured. It was the power from the outer world, and the visions and ceremonies had only made me like a hole through which the power could come to the two-leggeds. If I thought that I was doing it myself, the hole would close up and no power could come through” (Neihardt 1988:205-205).

Explanations of action and responsibility for it are also often shifted from individuals as wholes to some of their constituent parts. Contact-period New Zealand Maori often explained how their “dark intestines” would not allow them to rest until they had taken revenge for the death of a kinsman (Hanson and Hanson 1983:128-29). Abraham Lincoln (1839) recounted the anecdote of “a witty Irish soldier, who was always boasting of his bravery when no danger was near, but who invariably retreated without orders at the first charge of an engagement, being asked by his captain why he did so, replied: ‘Captain, I have as brave a heart as Julius

Caesar ever had; but, somehow or other, whenever danger approaches, my cowardly legs will run away with it.” Closer to home, a Holy Grail in today’s biomedical science is to look within individuals to identify the genetic causes of behaviors and qualities such as attention deficit disorder, autism, intelligence, and many others.

Of greater interest here is consideration of agencies that include but extend beyond the individual. A widespread example is the corporate group. In many simpler societies this takes the form of a kinship group such as a lineage or clan. Moral responsibility lies with the group as a whole, making it appropriate to retaliate for grievances or injuries against any member of the offender’s group. The famous feud between the Hatfields and the McCoys is a further case in point. More common in our society is the corporation, a legal individual that is held responsible for what it does to or for other corporations, clients, and employees.

The idea of the individual as subject is only a few centuries old. Erich Fromm, following Jacob Burckhardt (1954 [1860]:100-101), held that the individual was born in the Renaissance. “Medieval society,” he wrote, “did not deprive the individual of his freedom, because the ‘individual’ did not yet exist....[Man] did not yet conceive of himself as an individual except through the medium of his social...role” (Fromm 1941:43). It is also a peculiarly Western idea (Ess 2006:222). Confucius held that “persons are not perceived as superordinated individuals—as agents who stand independent of their actions—but are rather ongoing ‘events’ defined functionally by constitutive roles and relationships as they are performed within the context of their specific families and communities” (Ames and Rosemont 1998:20). The same perspective is found in Zen Buddhism: “seeking after and grasping at a ‘coherent self’ that is non-existent from the outset only leads to a ‘suffering.’ The Buddhist idea of ‘codependent arising’ maintains that all things under the sun arise in a codependent relationship with each other. Nothing in the

world exists in complete independence and isolation from others. There is no such a thing as a solid basis that exists autonomously” (Nishigaki 2006:240). And finally, one prominent contemporary theorist has suggested that the conflation of the subject with the individual may be on the verge of disappearing, even in the West, “like a face, drawn in sand, at the edge of the sea” (Foucault 1970:387).

To think of the subject as something other than the human individual, then, is by no means unusual in human experience. They have widely been considered to be extended agencies consisting of pluralities of human and nonhuman beings and constituted by the contexts and events in which they participate. One can think of them as much as verbs as nouns, the doings of activities. It is a notion consistent with physicist David Bohm’s view of the world as informed by relativity and quantum theory, in which everything is an unbroken flow of movement and supposedly durable things such as observer and observed are only momentarily stabilized forms of movement that form wholes for a time and then flow apart to join in new configurations (Bohm 1980:xi, 47).

To bring this into the practical realm, consider the well known phrase “guns don’t kill people; people kill people.” That is an absurd thing to say because it falsely implies that the possession of a gun is not pertinent to killing, or to the subject that possesses it. Of course guns don’t go around shooting people all by themselves, but everybody knows that a person with a gun is a far greater threat to kill someone than a person without one. Selinger and Engström (2007:575-76), following Ihde (1990, 2002), explain that human beings are changed when they use certain technologies: a man-with-a-gun is a different being or subject than the same man without a gun. Similarly, it is commonly said that some people are transformed when they get behind the wheel of a car. Or consider the saying, “if all you have is a hammer, everything looks

like a nail.” That is, the possibilities for action depend not just on human beings, but on the available means of action as determined by the relationship of humans with technology and/or other components of extended agencies. The entity that acts—the subject—is the extended agency (person/gun, person/hammer), not just the human individual.

The varieties of responsibility

If moral responsibility for an act lies with the subject that undertakes it, and if the subject includes nonhuman as well as human beings, then so may moral responsibility. Two versions of this point of view are distinguishable. One holds that at least some nonhuman entities may have moral responsibility in their own right. The other is that moral responsibility belongs to extended agencies as wholes rather than specifically to any human or nonhuman parts of them.

As for the first position, several authors who address the moral implications of automation ask whether automated agents such as computers and robots can be morally responsible (Coleman 2004, Friedman and Kahn 1992, Schick 1997, Hall 2000, Sparrow 2004). It is widely acknowledged that nonhumans participate in activities with moral import (e.g., Eurotransplant’s computers) but the salient question here is whether they possess, all by themselves, the mental qualities generally accepted as necessary for moral responsibility. Prominent among these are intentionality, the capacity to act voluntarily, and awareness of the consequences of what they do. Dennet (1997) and those cognitive scientists represented in Dietrich (1994) are more willing than most to credit automated agents with mental qualities such as these, and are thus more open to the notion that they are, or eventually will be, morally responsible. But most thinkers are unwilling to go this far, and they resort to locutions about technological objects being quasi-responsible, conducive to moral behavior, implicated in it, and

the like (Johnson and Powers 2005, Johnson 2006, Floridi and Sanders 2004, Stahl 2006, Verbeek 2006).

Possibly future development of automated systems and new ways of thinking about responsibility will spawn plausible arguments for the moral responsibility of nonhuman agents. For the present, however, questions about the mental qualities of robots and computers make it unwise to go this far. Moreover, this perspective's focus on nonhuman agents in their own right actually shares the moral individualist tendency to separate them from humans. Extended agency and joint responsibility theory aims precisely to overcome that separation. Peter-Paul Verbeek anticipates that when it is overcome, "ethics can move beyond the fear that nonhuman objects will...suffocate human subjects and direct its attention to the moral quality of associations of subjects and objects" (Verbeek 2009:255). This enables recognition that "technologies play a fundamentally mediating role in human practices and experiences, and for this reason it can be argued that moral agency is distributed over both humans and technological artifacts" (Verbeek 2008:24, see also Verbeek 2009:257).

However, while it is relatively straightforward to attribute actions to extended agencies, other knotty issues come into play when it is claimed that they are also morally responsible. Most important among these are the notions that 1) moral responsibility is associated with deserts and 2) it requires awareness of consequences and freedom of action. These matters are conventionally restricted to human individuals. But a case for joint responsibility can dismiss the first of these as irrelevant, and, as for the second, can argue that the joint responsibility of extended agencies is not only plausible but that it accords even better with certain everyday notions than does moral individualism.

Deserts

A standard objection to joint responsibility by proponents of moral individualism is that, as animals, machines, and tools are neither punished nor rewarded for the activities in which they participate, they do not share moral responsibility for them. The first response to this objection that a proponent of joint responsibility might make is to point out that deserts are not a criterion for attributing responsibility because the question of deserts comes up only *after* a determination about responsibility has been made. Especially in cases potentially involving capital punishment, criminal trials are often divided into the determination of guilt phase and the penalty phase. The first phase is limited to the question of whether the defendant is guilty of (i.e., responsible for) the crime. Only after that question has been decided in the affirmative is the second phase of the trial convened to determine the punishment. The punishment (desert), that is to say, is a *consequence* of an already-made decision about responsibility, not a factor in making that decision.

The same pattern characterizes virtually all applications of deserts, be it in formal situations or in the give and take of daily life. We consider punishment or rewards only after we have decided that what was done was bad or good. In behavioral psychology, rewards and punishment are known as positive and negative reinforcements. They are applied after a person or experimental animal has behaved in a certain way, positive reinforcement if the aim is to encourage future repetition of that behavior and negative if the aim is to extinguish it. In all these cases, the decision about whether a certain form of behavior is good or bad, desirable or undesirable always comes before a decision to apply deserts of one sort or another.

It might still be maintained that even if deserts are after-the-fact responses to decisions about moral responsibility rather than criteria for making them, deserts are still unique markers of responsibility. That is, all applications of deserts occur when moral responsibility is involved,

and only then. The rejoinder to that would be, in the first place, that it is simply not true, and in the second, it is not evidence for moral individualism because deserts are applied not only to human individuals, but to other entities as well.

Deserts are often used in response to behaviors that have no moral implications at all, as when psychologists give treats or electric shocks to rats that push buttons they are being trained to push or avoid. They are also used in cases where the behavior itself may be morally pertinent but the subject performing it is not considered to be morally responsible. Children are routinely rewarded and punished for desirable and undesirable behavior before they are old enough to be responsible. Owners constantly praise or chastise their pets (“Good dog!” “Bad dog!”), give them treats to reward appropriate behavior, and sometimes subject them to corporal punishment such as spanking. In the Middle Ages the animal as well as the human participant could be hanged if found guilty of bestiality, and bells that were rung to summon crowds to an uprising were flogged or destroyed (Ihde 2006:273-74). People criticize tools, machines and other inanimate objects that are badly designed or made, blame them for poor performance, and praise things of all sorts that function well. Yet, with the possible exception of medieval notions about beasts and bells, none of these—be it a small child, an animal, or an inanimate object—is thought to be morally responsible for what it does.²

Moreover, the notion that the application of deserts supports moral individualism cannot stand because they are often applied to subjects other than human individuals. It is possible to view punishments such as cutting off the hand of a thief or castrating a rapist as directed specifically against a part of the body rather than the persons as a whole. In a similar vein, sharing much with Lincoln’s “cowardly legs” anecdote recounted above, Jesus said “if your eye causes you to sin, pluck it out; it is better for you to enter the kingdom of God with one eye than

with two eyes to be thrown into hell.” The sentiment is important enough to be repeated almost verbatim in two gospels: Mark 9:47, quoted here, and Matthew 18:9. In both of these places Jesus also expresses similar attitudes about one’s hand or foot.

As for subjects that extend beyond the individual, it has already been noted that corporations are held to be morally responsible, legally as well as in public opinion, for their acts. They are praised and given citations, sued and fined or even (the corporate equivalent of capital punishment) dissolved for their misdeeds. It is the same with those informal extended agencies that are ubiquitous in everyday life, which form spontaneously to undertake particular activities and then dissolve when the task is finished and their components reform in other combinations for other activities. This is especially clear with certain kinds of negative sanctions that have the goal of breaking up an offending extended agency. Conviction of reckless driving or DUI often includes suspension of one’s driver’s license. When a child misbehaves with a toy by using it destructively or refusing to share it, a typical punishment is to take the toy away. The conventional view is that the child or criminal is being punished. But the problem emerges from the *conjunction of* child and toy, driver and vehicle. If either component were lacking—the child *or* the toy, the driver *or* the car—the offense would not have occurred and there would be nothing to punish. Thus it is reasonable to understand the punishment as directed against an offending relationship rather than a particular part of it. While we tend to think that the toy is being taken away from the child, for large toys with fixed positions such as a sandbox or a jungle gym, the child is removed from the toy. But what is really going on is that the offending toy-child relationship is being terminated.

The same principles apply, in mirror image, for positive sanctions. What are appreciated, praised, and encouraged are, at bottom, beneficial deeds and their results. These are often the

products of extended agencies rather than individuals acting alone. Even when individuals are honored they often acknowledge that the achievements were possible only because of the contribution of other people, education, and other circumstances. Frequently rewards are conferred explicitly upon extended agencies as wholes. This is obvious in the case of championship trophies and prizes awarded to debate and athletic teams, but there are many other examples. If a company decides to make a special investment in a given department or project, or a foundation makes a grant for an innovative educational program or for research, the award typically includes bonuses or salaries for the leading individuals, financial support for lower level personnel, and enhanced resources for office, classroom or laboratory space and equipment. The recipient of the reward is, in such a case, the department, school, or project team, i.e., the extended agency itself.

To summarize, 1) the issue of deserts comes up only after decisions about moral responsibility have already been made, 2) recipients of deserts may be considered to be either morally responsible or not morally responsible, and 3) those recipients may be animals, inanimate objects, and extended agencies as well as human individuals. For these reasons, joint responsibility theory holds that deserts are not decisive factors in assigning moral responsibility and certainly do not constitute evidence for moral individualism.

Freedom of action and awareness of consequences

Two people collude in a murder. One provokes the victim to chase him into the street and the other runs the victim down with a car. Neither conspirator could have accomplished this deed alone. Everyone would agree that they are both responsible for the crime. But it is also true that it could not have been done without the car. Does it share in the moral responsibility? Moral individualists would say no, because the car has no choice of action and no awareness of

the consequences of what it did. The reason for this is the presumption that only human beings can act of their own volition and know (or, to include negligence, should know) the consequences of what they are doing. The criteria of awareness and free choice are so strong that when they are not met, even adult humans are released from responsibility. People are not held morally responsible for acts they are forced to do, for consequences of their deeds that were completely unforeseeable, or if they are unable to evaluate the consequences of their acts (as with the McNaghten Rule that a defendant who was unable to distinguish between right and wrong is not guilty by reason of insanity).

Joint responsibility theory is in basic agreement with these propositions, and makes no claim that the car has moral responsibility by itself. It diverges from moral individualism in that it attributes responsibility to the extended agency as a whole, including both the human perpetrators and the car.³

Awareness of consequences and freedom of choice, which indeed are limited to the human conspirators, are necessary but not sufficient conditions for moral responsibility. It is also necessary for the deed actually to have been done. If the victim in our murder example had not been run down, there could be no responsibility for running him down. This was done by the extended agency that included the two humans and the car. Given that moral responsibility cannot exist but for the action of the extended agency, it lies with the extended agency as a whole and should not be limited to any part of it.

This argument can be run in a positive as well as a negative sense. Two people cooperate to rescue a toddler who had fallen into a well. The one descends into the well, a rope tied around her waist, to seize the toddler. The other, at the upper end of the rope, pulls his partner and the toddler out of the well. Everyone would agree that both human rescuers are responsible for the

good deed. But it is also true that the rescue could not have occurred without the rope. Absent the rope there would have been no rescue and therefore nothing to be responsible for. Hence the extended agency of rescuers-and-rope is responsible for the deed.

Obviously, the term “responsibility” takes on a different meaning here from its conventional usage. It is not, however, that the conventional sense of the word is being stretched to include extended agencies. The conventional sense of “responsibility” is defined by moral individualism, which is an artifact of the underlying theory of methodological individualism. Extended agency is an alternative theory to methodological individualism, and the concept of joint responsibility is equally an artifact of extended agency theory. This is the paradigm shift mentioned earlier. To entertain the possibility of joint responsibility entails recognizing extended agency as a distinct theory of action.

Joint responsibility can also be anticipatory or forward-looking (see Johnson and Powers 2005:100). A father hands the car keys to his sixteen year old daughter who has just gotten her driver’s license. He admonishes her that, as the driver of a car, she is assuming new and greater responsibilities. In this usage, “responsibility” refers not to what has happened but to what *might* happen: she might hit a pedestrian or another car, cause injury to people riding in her car, get a speeding ticket, and so on. She is responsible to take proper precautions to avoid such events. But she can assume those particular responsibilities only when she gets behind the wheel of a car. Here again, if the action of an extended agency is necessary for the potential occurrence of an event or circumstance with moral import, then the anticipatory moral responsibility should be attributed to the extended agency as a whole rather than limited to the human part of it.

Intentions

Extended agency theory generates an argument for intentions that parallels the one for joint responsibility. I will not develop a detailed analysis here, but simply indicate its main lineaments. Intending, like forward-looking responsibility, is a before-the-fact concept that concerns what may happen rather than what has already happened. To intend involves willing, but it is limited to activities that have a reasonable probability of actually being attempted, as well as a reasonable chance of being achieved. It makes sense, for example, to say that one would like to jump across the Grand Canyon, or wishes one could do so, but no sense to say that one intends to do so. In Salman Rushdie's novel *Shalimar the Clown*, Shalimar is a highly accomplished Kashmiri tightrope walker. His dream is to become so adept that he can dispense with the rope entirely. He would like to do this, he wants to do it, and at one point in the novel he even does it, escaping from prison by running along the top of a wall and continuing to run through the air after the wall comes to an end. But the reader can countenance that episode only by suspending disbelief, for in the real world it is not possible for a human being to run through the air supported by nothing. Therefore, in the real world, Shalimar could not intend to do it because it cannot be done.

Just as moral responsibility requires awareness of consequences and freedom of action, intending requires will. Inanimate objects and animals can no more intend to do something than they can be aware of the consequences of what they do or are free to decide whether or not to do it. Most important for our purposes, however, is that intention also shares the requirement with forward-looking responsibility that the act in question could actually take place. If it is not possible, it cannot be intended. If it is an act that can be accomplished only by an extended agency that includes nonhuman as well as human components, such as walking on a tightrope, the intention to do it lies with the extended agency as a whole. It must include a human

component because that is the locus of the will, which is necessary to intentions. But it must also include the nonhuman components, because without them the act could not be done, and therefore it could not be intended.

Moral individualism and causal responsibility

If pressed, a proponent of moral individualism might accept the proposition that responsibility extends beyond the human individual, but still insist that human and nonhuman entities have different *kinds* of responsibility. The responsibility of the car in our murder example or the rope in our rescue example is purely causal (as, for example, high wind might be responsible for blowing down a tree) while that of the human perpetrators, who alone are aware of what they are doing and have the freedom to act otherwise, is both causal *and* moral.

It is possible to frame parallel arguments for intending, but I will limit the response to responsibility. First, assigning different kinds of responsibility to different components of an extended agency depends on the ability to distinguish unequivocally between what the various components do. Sometimes this is not difficult. In our rescue example, the exact roles played by each of the two human rescuers and the rope are easy to identify. But with extended agencies that include automated expert systems featuring decisions made by computers, differentiating tasks is much more difficult. As Johnson and Powers point out, “the distribution of tasks to computer systems integrates computer system behavior and human behavior in a way that makes it impossible to disaggregate in ascribing moral responsibility” (2005:106).

It is especially difficult to call upon initial human programming to explain the behavior of artificial intelligent systems that are the result of CAD-CAM (computer-assisted design, computer-assisted manufacture) and that can learn. After a time such systems may develop into something quite different from their beginnings, independently of any further human input.

Consider Tierra, a computer program designed by biologist Thomas Rey to simulate evolution by natural selection. He began with a self-replicating digital creature consisting of 80 instructions. Its progeny were designed to replicate themselves with fewer instructions, being selected for by using less CPU time in an environment where that is a scarce resource. Over the generations and with no further human intervention, the system produced new, more efficient creatures. Small parasitic creatures emerged that co-opted the features of larger ones. In response, some of the host creatures developed immunity to the first generation of parasites, following which new parasites capable of penetrating the hosts' defenses were replicated. Rey, the original programmer, could not predict the developments that were taking place (Turkle 1998:321). If any system with clear moral significance, such as Eurotransplant, incorporates machine learning, it would eventually become impossible to sort out how the various parts of the extended agency influence each other and what part of an action each of them does. At that point the only recourse is to assign responsibility to the extended agency as a whole.

It might be argued, however, that the difficulty in distinguishing what humans do from what computers do in cases such as this is purely a practical problem. As such, it has no bearing on the theoretical or in-principle issue of where moral responsibility lies. But practical issues often determine how various situations are regarded, and this is as true of individualism as it is of extended agency theory. It is commonly recognized that an individual's behavior is determined by multiple factors, including genetic make-up, childhood experience, formal training, and momentary impulse. It should be possible in principle to sort out the particular contribution of each factor in a given situation, but this is seldom done because the task would be extremely complex, time consuming, and highly controversial. Instead, moral individualists adopt the practical option of attributing responsibility to the person as an undifferentiated whole. Cases

where the contributions of humans and computers to tasks such as predicting the weather or allocating donor organs are difficult to disentangle are no different.

Second, joint responsibility actually accords better in certain ways with general understandings of responsibility than does moral individualism. Recall the father admonishing his teenage daughter about her new responsibility, now that she will be driving a car. It is possible that he articulated similar warnings several years earlier, when she learned to ride a bicycle. “Be careful not to run into people or things, don’t crash your bike or hurt yourself, and especially don’t ride into the street without looking.” Her responsibility with the bicycle is, however, considerably less momentous than that with the car. Imagine now that she later becomes President of the United States and assumes the heaviest responsibility that any human being can have: to order a nuclear attack. Here her responsibility is infinitely greater than that associated with riding a bike or driving a car.

We appreciate the different degrees of responsibility in these examples immediately and intuitively, but they are more readily explained by joint responsibility than by moral individualism. If moral responsibility is restricted to human beings, then differences in responsibility must be explained in terms of human differences. But in this case the woman, who continues to ride a bicycle and drive a car after she has become President, is a constant. She is expected to exercise the most thoughtful prudence in all of these contexts. Hence moral individualism gives no satisfactory explanation for the obvious differences in her responsibilities. Using the concept of joint responsibility, the difference is explained by recognizing that there is not one subject in these scenarios, but three. More damage can be done by the extended agency woman-driving-a-car than woman-riding-a-bike, and a great deal more damage still can be done by woman-with-her-finger-on-the-nuclear-trigger. When we recognize the responsible parties in

these three cases as three quite different extended agencies, we can readily explain the differences in the moral responsibilities they carry.

Finally, the joint responsibility perspective encourages constructive, moral behavior in all contexts. Under moral individualism people are isolated in their skin bags, independent of other things. They of course have obligations to others, but the others remain, precisely, Other, ultimately alien from the Self. In contrast, extended agency theory emphasizes the multiple connections between humans and nonhumans of all descriptions in systems of action ranging in scope from the immediate all the way to the global. This is more consistent with recent emphases on ecological thinking. When the subject is perceived more as a verb than a noun—as a way of combining different entities in different ways to engage in various activities—the distinction between Self and Other loses both clarity and significance. When human individuals realize that they do not act alone but together with other people and things in extended agencies, they are more likely to appreciate the mutual dependency of all the participants for their common wellbeing. The notion of joint responsibility associated with this frame of mind is more conducive than moral individualism to constructive engagement with other people, with technology, and with the environment in general.

REFERENCES CITED

Ames, R. and H. Rosemont

1998 *The Analects of Confucius: A Philosophical Translation*. Ballantine Books, New York.

Bohm, David

1980 *Wholeness and the Implicit Order*. London: Routledge & Kegan Paul.

Burckhardt, Jacob

1954 [1860] *The Civilization of the Renaissance in Italy*. New York: Modern Library.

Clark, Andy

2003 *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*.
New York: Oxford University Press.

Coleman, Kari Gwen

2004 *Computing and Moral Responsibility*. In *Stanford Encyclopedia of Philosophy*, Fall
edition. E. N. Zalta, ed *Stanford Encyclopedia of Philosophy*.

De Meester, Johan, Guido G. Persijn, Frans H. J. Claas and Ulrich Frei

2000 "In the Queue for a Cadaver Donor Kidney Transplant: New Rules and Concepts in the
Eurotransplant International Foundation." *Nephrology Dialysis Transplantation* 15: 333-
338.

Dennett, D. C.

1997 *When HAL Kills, Who's to Blame? Computer Ethics*. In *HAL's Legacy: 2001's Computer
As Dream and Reality*. D. G. Stork, ed . Cambridge, MA: MIT Press.

Dietrich, Eric, ed.

1994 *Thinking Computers and Virtual Persons: Essays on the Intentionality of Machines*. San
Diego: Academic Press.

Ess, Charles

2006 *Ethical Pluralism and Global Information Ethics*. *Ethics and Information Technology*
8:215-226.

Flew, Anthony

1995 *Thinking About Social Thinking*. 2nd edition. Amherst, NY: Prometheus Books.

Floridi, Luciano, and J.W. Sanders

2004 *On the Morality of Artificial Agents*. *Minds and Machines* 14:349-379.

Foucault, Michel

1970 The Order of Things: An Archaeology of the Human Sciences. New York: Vintage.

Fromm, Erich

1941 Escape From Freedom. New York: Rinehart & Company.

Giere, Ronald N.

2008 Human Moral Responsibility Is Moral Responsibility Enough: Comment on F. Allan Hanson's "The Anachronism of Moral Individualism and the Responsibility of Extended Agency". *Phenomenology and the Cognitive Sciences* 7:425-427.

Hall, J. Stors

2000 Ethics for machines, Discuss.Foresight.Org/~Josh/Ethics.Html, August 16, 2004.

Hanson, F. Allan

2004 The New Superorganic. *Current Anthropology* 45:467-482.

2007 The Trouble with Culture: How Computers are Calming the Culture Wars. Albany: State University of New York Press.

2008 The Anachronism of Moral Individualism and the Responsibility of Extended Agency. *Phenomenology and the Cognitive Sciences* 7:415-424.

Hanson, F. Allan and Louise Hanson

1983 Counterpoint in Maori Culture. London : Routledge & Kegan Paul.

Haraway, Donna J.

1991 Simians, Cyborgs, and Women: The Reinvention of Nature. London: Free Association Books.

Hutchins, Edwin

1995 Cognition in the Wild. Cambridge, MA: MIT Press.

Ihde, Don

- 1990 Technology and the Lifeworld. Bloomington: Indiana University Press.
- 2002 Bodies in Technology. Minneapolis: University of Minnesota Press.
- 2006 Forty Years in the Wilderness. *In Postphenomenology: A Critical Companion to Ihde*. E. Selinger, ed. Pp. 267-290. Albany: SUNY Press.

Johnson, Deborah G.

- 2006 Computer Systems: Moral Entities But Not Moral Agents. *Ethics and Information Technology* 8:195-204.

Johnson, Deborah G., and Thomas M. Powers

- 2005 Computer Systems and Responsibility: A Normative Look At Technological Complexity. *Ethics and Information Technology* 7:99-107.

Jones, Richard H.

- 2000 Reductionism: Analysis and the Fullness of Reality. Lewisburg, PA: Bucknell University Press.

Law, John

- 1999 After ANT: Complexity, Naming and Topology. *In Actor Network Theory and After*. J. Law, and J. Hassard, eds. Pp. 1-14. Oxford: Blackwell.

Lincoln, Abraham

- 1839 Speech on National Bank delivered to the Illinois House of Representatives, December 20, 1839. *The Writings of Abraham Lincoln*, vol. 1.
<http://www.classicreader.com/book/3237/>

Neihardt, John

- 1988 Black Elk Speaks. Lincoln: University of Nebraska Press.

Nishigaki, Toru

2006 The Ethics in Japanese Information Society: Consideration on Francisco Varela's *The Embodied Mind* From the Perspective of Fundamental Informatics. *Ethics and Information Technology* 8:237-242.

Schick, Theodore, Jr.

Winter 1997 Can a Robot Have Moral Rights? *Free Inquiry* 18:42-44.

Selinger, Evan, and Timothy Engström

2007 On Naturally Embodied Cyborgs: Identities, Metaphors, and Models. *Janus Head* 9:553-584.

Sparrow, Robert

2004 The Turing Triage Test. *Ethics and Information Technology* 6:203-213.

Stahl, Bernd Carsten

2006 Responsible Computers? A Case for Ascribing Quasi-Responsibility to Computers Independent of Personhood or Agency. *Ethics and Information Technology* 8:205-213.

Tufts Anette

1996 "Eurotransplant to Allocate Kidneys by Computer." *Lancet* 347(9011):1326

Turkle, Sherry

1998 "Cyborg Babies and Cy-Dough-Plasm: Ideas about Self and Life in the Culture of Simulation." *In* *Cyborg Babies: From Techno-Sex to Techno-Tots*. R. Davis-Floyd and J. Dumit, eds. Pp. 317-329. New York: Routledge.

Verbeek, Peter-Paul

2006 The Morality of Things: A Postphenomenological Inquiry. *In* *Postphenomenology: A Critical Companion to Ihde*. E. Selinger, ed. Pp. 117-128. Albany: SUNY Press.

- 2008 Obstetric Ultrasound and Technological Mediation of Morality: a Postphenomenological Analysis. *Human Studies* 31:11-26.
- 2009 Cultivating Humanity: Toward a Non-Humanist Ethics of Technology. In *New Waves in Philosophy of Technology*. J. K. Berg Olsen, E. Selinger, S. Riis, eds. Pp. 241-263. Hampshire: Palgrave Macmillan.

¹ This essay represents my second attempt to deal with these issues. For the first, see Hanson (2008) and, for a rejoinder, Giere (2008).

² It is important to bear in mind that the concept of joint responsibility under discussion here does not advocate the moral responsibility of nonhuman things in their own rights, but only the responsibility of extended agencies consisting of humans and nonhumans taken together.

³ Strictly speaking the agencies involved in these events include many components in addition to those specified here, such as gasoline to fuel the car, the weather, the conditions of the road, and so on. There comes a point of diminishing returns, where the participation of some elements is sufficiently inconsequential (the make and year of the car) or so constant (the effect of gravity, the presence of oxygen) that they can be disregarded. As this is a discussion of general concepts rather than exact allocation of responsibility for particular events, it considers only the most obvious variables of any extended agency.