

Keeping *Homo Sapiens* 1.0

Anne Foerst,

St Bonaventure University, St Bonaventure, New York, USA

ABSTRACT

From a theological perspective, the human animal is estranged. The feat of self-awareness leads also to the rejection of other human beings as non-persons. In this article, I will first outline a theological concept of sin that is coherent with modern scientific findings. I will then discuss human bonding mechanisms and their evolutionary evolved limits. Finally I will argue that we ought not to improve *homo sapiens* but should still attempt to build an intelligence partner species (e.g. humanoid robots) that can teach us to be more tolerant and inclusive.

Keywords: sin, humans as God's images, personhood, humanoid robots, human-computer interaction

Many people agree that our species has quite a few design flaws. We become sick and have to die, we have a really bad memory, our children are born too early and depend on their parents too long, our upright walk gives us knee and hip problems. The most vulnerable part, our brain, is quite exposed, our throat is too easily slit, and, despite being omnivores, we can easily be poisoned. We know that most of these flaws are a result of our evolutionary development. But with gene-therapy and computer enhancement we might be able to improve ourselves. And what should we improve?

When reading Science Fiction literature, the theme often is our tendency to violence and hatred towards others. Most famously, Isaac Asimov invented a peaceful breed of robots guided by the three laws of robotics,¹ which made them incapable of doing harm to humans. In his early work, these robots are often depicted as "better people" but in his late work he discovered inconsistencies: because the robots had to obey any human command where the harm to another human being was not immediately obvious, they would ultimately end up harming humanity. So he has the mind-reading robot Giscard develop a "Zeroth Law" (A robot may not injure humanity or, through inaction,

allow humanity to come to harm) and modifies the First Law (A robot may not injure a human being or, through inaction, allow a human being to come to harm, except where it would conflict with the Zeroth Law.) With this additional stipulation, robots enter the same moral ambiguities as we humans live in because the question "What harms humanity?" cannot have an objective answer as the term "humanity" itself is ambiguous.

So could we improve humans so that they always know what is best for the greater good even though we don't know what it is? Is there a way to enhance *homo sapiens* so that we maintain all our wonderful capabilities but cease to fight and destroy each other? I will argue that this is not possible. Our flaws also lead to our greatest deeds, and our capability to hate is deeply connected to our capability to love and to bond. In other words, I find the human species quite wonderful as is! At the same time, though, I am convinced that we still could profit greatly from robots as our interaction with the children of our technological future can teach us something about who we are. In order to make this argument, I will first outline an anthropology that is influenced by both the Jewish and Christian traditions and scientific insights. Then I will introduce some insights from human-robot interaction and draw some conclusions.

A RELIGIOUS ANTHROPOLOGY

The first statement about humans made in the Bible is that they are created in the image of God. The Hebrew term *sælæm* for "image" means literally a "clay statue" and is usually used for a statue of a Divinity. So the first thing we learn about humans in the Bible is that each and every one of us, independent of looks, capabilities, creed, and gender, is a divine statue. However, like all other animals, humans don't have an inbuilt sense of species recognition. While the overall evolutionary process does influence whole species, it is maintained by individuals who care only about

their own survival and that of genetically close relatives. Primatologists have observed a strong sense of solidarity with insiders and the rejection of outsiders (including common occurrences of infanticide) among our closest relatives. Anthropologists have found that in many indigenous cultures the species definition (the equivalent of our term “human”) is equal to the name of the respective tribe. And we know from infant development that six-month-old babies are capable of distinguishing between chimp faces and can mimic all human sounds. At nine months, however, babies can only distinguish between those faces they are surrounded by and their sounds are already limited to those of their “mother-tongue”. This means that the Biblical *Imago Dei* is not descriptive but an ethical guideline.

But there are other statements in the Bible that are surprisingly coherent with modern scientific insights. Humans, like all other animals, depend on God-given food, and they have to die. Contrary to popular belief about the Bible, humans have always been mortal – what distinguishes us from other animals is that we are aware of our mortality. And this is the key to the human condition: estrangement for which the Bible gives a beautiful narrative. In Genesis 1, the text states after each act of creation *ki tow*, “and it was good”. Only after the creation of humans is this affirmation of goodness missing, but the whole of creation, including humanity, is very good. Humans, as God’s statues, are affirmed and accepted and part of something very good but they themselves are not good. Genesis 3 can be seen as explanation. When Adam and Eve in the poetic language of the text eat from the tree of knowledge of good and evil, they start to judge. Looking at each other after the so-called “Fall”, they recognize that they are naked and cover themselves up. Or, they start to judge nakedness as evil and therefore cover up.

It is a uniqueness of the human condition that the “I–Thou” relationship implies that the “Thou” is another, and different from me. This means that in recognizing an “I” I also recognize many “non-I’s”, people I am separated from, or estranged from.

Infants develop the sense of I–Thou at approximately 36 months of age. Before this age, they do not have a sense of self and other. They cannot put themselves into someone else’s shoes and don’t recognize even their parents as people different from them. In itself, the development of a self–other recognition is an enormous achievement. While several other animals (including chimps and

dolphins) recognize themselves in a mirror, we seem to be the only species with a developed sense of self-awareness. So why is this development in the Biblical language tied to the “Fall”? Because, through this self-awareness, we become aware of our mortality and we become aware of differences among human animals. Because we judge, we use otherness to create categories of Good and Evil to justify rejection of other beings who are not like us. Hence, theologically speaking, the key to the human condition is estrangement – from oneself, from others, from nature, and from God.

When God according to Christian doctrine becomes human, it is to understand the human condition and to understand why it is impossible for us to “love God with all our hearts and minds” and to “love our neighbor like ourselves”. Thus, Jesus, as true God and true human, bridges the gap between humans and God.

And it is no coincidence that in the main ethical framework from Jesus, as laid out in the Sermon on the Mount (Matthew 5–7) the judging of others is the most criticized action.

It is fascinating how the human capability to judge is so deeply rooted in who we are and the cause for such greatness and yet, so many flaws. In paradise, before the so-called “Fall”, Adam and Eve were innocent but also not fully human. The I–Thou recognition does imply a loss of innocence (you are not “I”) but is also the key to understanding humans.

Christian tradition calls our senses “the portal to the soul”. We share our senses with all other mammals but our judging makes the human use of senses unique. Only because we judge our sensory impressions do we go out of our way to create positive sensory input: we create music for positive auditory input, we create art for positive visual input. We cook and create perfumes for positive taste and smell input, and we create wonderful materials that feel good on our skin. And the judgment of each human is unique and reveals who each of us is. In other words, humans become artists because we judge. And this artistic capability sets us apart even from tool-using and symbol-processing chimps.

But our sensory judgments are also deeply influenced and even manipulated by our society. We often deem our cultural sensory judgments superior to those of other cultures. Just look at how many people look askance at foods from other cultures

or clothes or music. It is hard after a certain age to value unfamiliar sensory inputs as highly as those we grew up with. So the judgment of our senses does not only serve as a tool for uniting people in a society; it also serves as a tool for estrangement. While it makes *homo sapiens* uniquely creative, it also creates estrangement from those who don't share one's taste or who do not appreciate one's sensual judgment. And we use quite arbitrary criteria to reject others: the way they look, the way they smell, the way they dress – all these and more features can be used by us to judge the other not just as “an other”, but as inferior to us.

This means that the very act of judging gives us humans our greatness but also causes the rejection of otherness. The Christian message of redemption does not call us back to paradise as we aren't truly human there. Instead, it teaches us to accept this estrangement and its wonderful sides but to try to limit its bad sides. In this theology, “sin” and “guilt” are two different concepts with “sin” describing human ambiguity and estrangement a neutral state of being.

The insight to our estrangement and the need for redemption is theological but the reasons for this estrangement can be found in scientific theories. We have mentioned that sounds and facial recognition are already socialized by the age of nine months. After that age, babies will be very good at identifying and recognizing faces and sounds they are familiar with but increasingly lose the capability to distinguish between faces and sounds from other cultures or species.

In addition to these limits to human bonding, analysis of military and church history has shown that a group which has more than approximately 150 members becomes unstable and often falls apart. There seems to be a natural limit to the amount of people with whom we can bond. The mirror cells that enable us to have empathy with others only work in shared physical space and it seems that true bonding can only happen in physical proximity.

From birth onward, humans are trained in social interaction. The two main learning strategies for the infant are turn-taking and mimicking, and both require social interaction. By the time babies become self-aware, most of the social cues and responses are so deeply embedded in the social structure of their environment that they are not even aware of them. This means that even if there are limits to our physical capability of bonding, we usually don't notice them as physical limits

but attempt to find other explanations. So humans create myths and legends that justify their physical limitations. And, usually, these narratives use concepts of superiority. One just has to look at the Bible. The chosen people are the Jews; Jesus speaks only to Jews and refers to pagans as mangy dogs under a table. While Paul starts to include Pagans in the Christian framework, all those who do not accept Jesus as the Christ or Messiah are now condemned. We find similar exclusivity in most other meaning-giving narratives and creation myths.

HUMAN–ROBOT INTERACTION

While humans have quite extensive limits to their capability of bonding with other humans and tend to judge humans with features different from them as inferior, studies in human–computer interaction have also revealed our amazing capability to bond with beings totally different from us. Cliff Nass, sociologist at Stanford, has done some fascinating experiments. In one, he asked several people to test a computer-learning program that was intended to be introduced to elementary schools – the program was very bad. Some of the testers were computer specialists and some were laypeople. After they had tested the program for a while, the computer on which they worked asked them to evaluate its performance. For the most part, people responded positively. Afterwards, these same testers were led into another room with other computer terminals and were asked to evaluate the learning program again. Here, on these different computers, their answers were less positive about the quality of the tested software but they still sounded somewhat satisfied. Finally, a human with pen and paper asked the testers for their opinion on the software and the testers were very negative about it. Such a program should never be used in school, they said.

Interestingly enough, the testers had not voiced these criticisms to either the computers they had tested the program on, or the computers in the other room on which they had done a second evaluation. These same people, when asked if they would ever be polite to a computer or think they could hurt its feelings, rejected such a notion vehemently.

This experiment suggests that somehow we seem to apply our rules of politeness to non-human entities such as computers. Obviously, the participants in the experiment did not want to hurt the computer's feelings. They even assumed a level of kinship between different computers and, therefore, applied similar rules of politeness on the computer on which they did a second evaluation.

They didn't give these machines their true, very critical opinion either out of a desire not to hurt the feelings of the second computer by criticizing one of its "fellow computers" or because they assumed some "contact" between the two so that the second would tell the first what had been said.

In another experiment, Nass placed people and computers inside one room. Half of the computers had green monitors while the other half had blue monitors. Half of the people wore green arm badges; the other half wore blue ones. All played interactive games together and it turned out that for those people with blue arm badges it was much more successful to use computers with blue screens to reach their goal than to use "green" machines. The same, of course, was valid for the other side. So, slowly, the people with green arm badges bonded with the green-monitor machines and the "blue" people with the "blue" machines.

And here is the surprising result. After approximately half an hour, the people wearing the blue arm badge expressed more solidarity with the computers with the blue screens than with the humans with the green arm badges; the same was true for the humans with the green arm badges. It seems that through the interactive games and the experienced benefit of interacting with the machines with one's color code, the color code took over as a definition for "my" group. The entities with the other color code, no matter whether humans or machines, tended to be rejected.

This seems to imply that humans bond with the entities of their own group no matter whether they are human or not. Human beings are social mammals. And, as Nass's experiments shows, we seem to be able to accept anyone or anything into our group with whom we can sufficiently interact. As soon as such a stranger is accepted into a group, he or she is seen as an equal part of the group; that group defines itself by the members that both belong and do not belong to it. As we all know and perhaps do ourselves, people treat their cars and stereos as people as well. In a way, it saves a lot of time and energy to do so. After all, humans are educated from birth on how to interact with their fellow human beings. It is necessary for a baby to be able to do so as its survival depends on it. Throughout our lives, we learn patterns of behavior – such as being polite and not openly criticizing someone. It is very easy to apply these ingrained rules to every entity we interact with. It is very hard not to do so as it demands a conscious effort of us.

ARE ROBOTS PERSONS?

When I was part of the Kismet project at MIT, I often presented videos of the social and emotional robot Kismet to non-technical audiences. And whenever I asked the question as to whether Kismet deserves to be treated as a person, I got very different reactions. While some people were willing to concede such a possibility, many others rejected such a notion violently. When asked why they wouldn't assign Kismet personhood, they would usually list features that humans have and Kismet lacked that were, in their opinion, crucial for the assignment of personhood.

Let's have a look at the most commonly mentioned of these features. Kismet does not have language and is far away from anything like speech. It babbles and mimics but there is no understandable language. Does this argue against Kismet's personhood? I would think not because if language were a criterion, newborn babies wouldn't be persons either and if we were to treat our babies as non-persons they would not develop properly.

Kismet also has no body but consists of just a head and a neck. But it is nonetheless embodied as all parts of the head are in constant interaction with the world.

Kismet's building blocks are metal and plastic; there is no organic material. In a way, it cannot die. But is death a condition for personhood? After all, when we deny our fellow humans personhood, we don't accept this criterion as they are certainly mortal.

Kismet has no humor. But we shared many laughs while we interacted with Kismet, as members of a family will have when interacting with a baby. And a baby does not have explicit humor either. Kismet has no sense of smell. But perhaps that will be developed at some point.

All these arguments can be rejected for two reasons. One, it might very well be that through technical progress all these elements will be developed at some time in the future. Even the lack of organic material might be overcome as some researchers today work on organic chips based on viruses. Second, and much more importantly, every empirical criterion that might exclude Kismet from the community of persons will also exclude human beings from it. A counter-argument that was often presented at this point is that we talk about normal grown-ups and not borderline cases. But I cannot accept this. If we want to draw a clear

line between humans as persons and Kismet, the community of humans must include everyone, not just the “normal” cases. After all, who defines what is normal and what is not?

So all arguments based on empirical properties against Kismet’s personhood are invalid. This, however, means that a discussion about the personhood of future AI has powerful ramifications for humans.

The Bible assigns personhood to all human beings as a divine gift. Humans are chosen partners of God and as such are special. But on an empirical level we do distinguish between the concept of “being human” and the concept of “being a person”. We are capable of denying our fellow humans personhood (we do so in every war and genocide) and yet we might assign non-human beings personhood, as seen in Nass’s experiments. No animal has “inbuilt” species recognition and neither do we.

So how can robots help to make us “better people”? For one, if we look at arguments against Kismet’s personhood that are based on empirical properties we will realize that it is the same kind of judgment that leads us to reject human beings who are different from us. If people are excluded from the community of persons due to their skin color, their gender, or their religion, this exclusion is based on empirical properties – even though usually justified by superiority myths (African-Americans are more criminal, women have smaller brains, Muslims are radical etc.). If people react with fear towards robots and, thus, reject them and justify this rejection with empirical properties, then we can learn from this action about our rejection of other people insofar as it might be rooted in fear. This means that the creation of robots can teach us to be more tolerant and inclusive.

But this is not the only way robots can turn us into “better people”. According to Jewish mysticism (the Kabbalah) the creation of humanoids called “Golems” is a prayer. We are created in the image of God and, hence, participate in the divine creativity. Whenever we are creative we celebrate God. But when we re-create ourselves, the most complex being we know of, we celebrate God the most. And, indeed, the building of robots makes us modest because it is so very hard. Any insect is more complex and capable of richer behaviors than the most complicated robots ever build.

I think that *homo sapiens 1.0* is a pretty interesting and fascinating animal and that any change would most likely destroy valuable features. However, we ought to attempt to re-build ourselves, not to better our own species but to create a partner species that teaches us to be more humble and less anthropocentric and egotistical. We should work towards the goal of technical counterparts of ourselves that might teach us tolerance and the appreciation of otherness, capabilities that have been curtailed by our evolutionary development.

NOTES

1. 1. A robot may not injure a human being or, through inaction, allow a human being to come to harm. 2. A robot must obey orders given to it by human beings, except where such orders would conflict with the First Law. 3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

CORRESPONDENCE

Anne Foerst, St Bonaventure University, Depts of Theology and Computer Science, St Bonaventure, New York, NY 14778, USA.
E-mail: afoerst@cs.sbu.edu

Published online 19 October, 2011
ISSN 1749-3463 print/ ISSN 1749-3471
DOI: 10.1080/17493460802480446
© 2008 Artifact