

Is it good for them too? Ethical concern for the sexbots

Steve Petersen

May 2016

In this chapter I'd like to focus on a small corner of sexbot ethics that is rarely considered elsewhere: the question of whether and when being a sexbot might be good—or bad—for *the sexbot*. You might think this means you are in for a dry sermon about the evils of robot slavery. If so, you'd be wrong; the ethics of robot servitude are far more complicated than that. In fact, if the arguments here are right, designing a robot to serve humans sexually may be very good for the robots themselves.

Of course for *today's* models, the question of whether it's good for the sexbot makes little sense; they are “just machines”, not genuine ethical subjects, so they cannot be ethically wronged any more than we can wrong a vibrator or a toaster. But there is good reason to think that *future* sexbots will be artificially sentient and artificially intelligent. Such robots would not just *seem* to experience pain or pleasure, they *would* experience it; they would not just *act like* they have deeply-held goals and values, but they would *actually* have them. I can't argue for this possibility here, so instead I will take a cheap shortcut and argue from authority: unlike most philosophical questions, the possibility of genuine AI wins wide consensus among professional philosophers. So if you disagree with the premise of robotic intelligence, I urge you to read some of the reasoning smart people have made in its favor.¹

If robots have genuine experiences of pain and pleasure, triumph and defeat, this in turn strongly suggests that they are subjects of real ethical concern. They could even be inorganic *persons* with moral standing equal to that of humans.²

¹Here is the gist of the argument that it is possible: the way we humans think seems essentially to do with *information-processing*—neurons taking information from the senses, calculating on it, and passing results to motor nerves. And it seems simply dogmatic at best, and outright bigoted at worst, to assert that carbon compounds can do thought-relevant information-processing while metal can't. For more than this too-brief sketch I would suggest the classic Turing (1950), and then the appendix to Lycan (1987), and next the more extensive introduction in Churchland (1988). John Searle's Chinese Room thought experiment (1980) is often cited on the other side of this issue, but this should not be taken as evidence that philosophers are evenly split on the question.

²According to philosophical tradition, 'human' refers to the biological species of *homo sapiens*, while 'person' refers to those with certain higher-level cognitive abilities. Thus it is

Part of the power of fictional sexbots like Pris from *Blade Runner*, Gigolo Joe from *AI*, or Kyoko from *Ex Machina* is exactly that we can't help suspecting that these characters have their own, real lives—and that those lives are not going very well.³ Indeed, if you agree that someday there could be such sexbots with their own ethical value, it might seem obvious to you that their lives would automatically be tragic. Sexbots seem to be slaves by their very nature, and if so then it seems clear we should prevent their creation.

Yet there will be strong incentives to create intelligent sexbots. No one would claim I'm an expert on sex (sadly), but I am pretty confident in this: there's more to rewarding sex than purely physical stimulation. Presumably this is why most of us continue to pursue sexual relationships with others, despite the fact that there are already various physically satisfying ways we can stimulate ourselves; sex with others includes an experience of personal connection and intimacy that we find separately rewarding—even when that experience is merely illusory.⁴ This would explain why the “girlfriend / boyfriend experience” is so popular in the sex trade.⁵

Given both the will and the way, it may seem that we are headed for tragedy: a population of sexbots slaves, forced into a lifetime of playing the happy companion to their sundry johns.⁶ But that is not my position here. There is a surprisingly strong argument that it is permissible to design and create genuinely intelligent, ethically valuable robots for the explicit purpose of serving humans sexually. This argument does not depend in any way on the permissibility of human sex work; as far as the reasoning here is concerned, it may be that human sex work is always wrong. The argument is specific to robots, or more generally to artificially designed people. It is basically an application of my past work on robot servitude and slavery.⁷

Two quick caveats before we begin, though. First, there will probably always be lots of other ways for anyone, including a sexbot, to live a miserable life full of injustice, if other people are mean enough. What we are asking here is whether a sexbot would be harmed just in virtue of being a sexbot.

possible for ET the Extra-Terrestrial to be a person but not a human, and for someone in a persistent vegetative state to be a human but no longer a person.

³It may be that robots are not literally *alive*, if (as some argue) life requires organic chemistry, or a history of natural selection. Here I mean ‘life’ more inclusively, basically to include any potential subjects of ethical concern. This notion is substrate- and history-independent; roughly it is the “negentropic” systems of Schrödinger (1945).

⁴This seems physiologically measurable; Brody and Krüger (2006), for example, find that prolactin release is 400% higher after intercourse than after masturbation.

⁵For example, from Radakovich (2010):

Veronica Monet, a former escort who wrote a book called *Sex Secrets Of Escorts*, says that the girlfriend experience is popular because the men requesting it “want an intellectual and emotional connection.”

⁶That is, assuming we humans survive for some time after we create genuine (strong, general) AI; there is some reason to fear this is not so, as Bostrom (2014) most famously worries. (But for some potentially mitigating response to Bostrom, see my Petersen (Forthcoming).)

⁷Petersen (2007) and Petersen (2012).

Second, my focus is only on the moral implications for the sexbot, not for its partners or for society at large. It might still be wrong to make a sexbot, even though it might be no wrong *to* the sexbot. Whether creating such sexbots is “all things considered” permissible depends on a tangle of other ethical questions and contingent facts about human psychology—questions on which I can hardly even speculate. Papers in this anthology by John Danaher and Chloe Georas consider such matters more carefully. (My favorite example of a potential societal impact of sexbots—dramatically illustrated by an episode of *Futurama* (2001)—is that sexbots may rob us of the need to accomplish things in order to impress potential sexual partners, and thus spell the end of human civilization.⁸)

The life of a sexbot

So now, assuming that there will be intelligent sexbots of ethical value, let us consider whether they could live good lives. (From here I’ll just call them “sexbots”, and assume you remember I just mean ones as intelligent and ethically valuable as humans.) When considering this, we must be careful to avoid an easy mistake: measuring the sexbots’ lives by comparing them to similar human lives. When asked to imagine a robot person designed from scratch for sex, our brains boggle with the unfamiliar. So we naturally (and perhaps subconsciously) consider the next closest familiar analogy: humans coerced from childhood into sex work. We are (rightly) morally repulsed by these cases, and so transfer our indignation to the sexbots by analogy. But this heuristic for evaluating the ethics of the circumstance misfires; sexbots would be so different from such more familiar cases that the analogy fails to hold.

Sexbot pleasures

The first important disanalogy is obvious: the different physical makeup of the sexbot. This makes for different kinds of sensory experiences, which in turn affects what the robot finds physically pleasurable. And since pleasure is at least a contributor to happiness, just this different physical makeup has ethical implications.

Pleasure, for our purposes, basically just means “good feelings.”⁹ Naturally positive physical sensations like those humans get from a good massage or fine chocolate count as pleasure—but pleasure in this sense also includes the flush we get when complimented by someone we admire, and the sighing relief upon finishing a challenging but rewarding task. Remember we are assuming the sexbots can experience real pleasure, not just simulate it—so now we ask about their prospects for doing so on a regular basis.

⁸Thanks to Danaher (2013) for reminding me of this episode.

⁹For a first pass at the philosophical complications, though, see Katz (2014) for an overview.

A little reflection reminds us that what causes pleasure is highly dependent on the nature of the experiencing creature. Cats are not particularly moved by doughnuts, for example, while we don't see anything special about catnip. Doughnuts are not, strictly speaking, just-plain pleasurable—they are only pleasurable-*for-us*. In the case of biological creatures with a history of natural selection, what's pleasurable is generally constrained to reinforcing what was evolutionarily advantageous; creatures not motivated to seek what is helpful and shun what is harmful do not on average fare well. But sexbots do not have evolutionary histories; they are designed by an intelligence from scratch. So what might cause pain and pleasure to a sentient sexbot? The answer seems to be: any of a wide range of things, at the designer's discretion. *Maybe* these causes will be constrained some; perhaps what causes bodily damage (like water in the circuits) must automatically be at least partly painful to a well-designed robot, and what enhances bodily integrity (like charging batteries) must automatically be at least partly pleasurable. Maybe it is a contradiction in terms to speak of taking physical pleasure (as opposed to some more abstract satisfaction) in physical damage. Even assuming so, there will still be much latitude for variation in sexbot design. Most obviously, a clever designer with a good team of engineers could increase relative *intensity* of pleasure, perhaps with finer-grained sensory mechanisms than humans have. A sexual performance that would evoke a bored yawn (or an inspired imitation of rapture) from a human partner could produce *real* physical ecstasy in a sexbot. The sexbots might also take greater aesthetic pleasure in more, or quite different, physical appearances than the typical human finds arousing. Finally, besides heightening the pleasure a human might experience in similar circumstances, the designer could also make the sexbot experience genuine pleasure from activities that produce casual indifference or even distaste in humans. A human who has trouble finding a willing human partner for unusual predilections may find an abundance of truly eager sexbots.

Other variations on sexbot sentience are more confusing to consider. For example, an unscrupulous designer could make a sexbot feel acute physical pain if that sexbot goes without sex for too long. Such withdrawal pains would be ethically bad for the sexbot. Or a designer could make sharp body blows pleasurable for the robot. If so, and if the robot is sturdy enough to sustain them without damage, then—strange or disturbing as it might seem to us—it is at least provisionally *good* for that sexbot to be whipped or beaten. (Remember good-for-the-sexbot does not mean all-things-considered-good; the psychological implications for the human doing the beating might always make such beatings net wrong.) Behavior typically associated with sexual masochism in humans would not strictly count as masochism in such a sexbot. (Whether this would change the appeal for the would-be sadist is an interesting psychological question.)

The point is that a sexbots' pleasures need not be like ours; they might genuinely *like* experiences that few humans would. So we should not suppose that sex is a dreary task for them, just because it is their intended career. It might be work they deeply enjoy—work they would do for free anyway, or even pay to do.

Depending on the circumstances of how the sexbot’s manufacture is financed, it might not even properly be called *work* at all.¹⁰ On this score at least, few humans are so lucky.

Sexbot desires

So we can suppose well-designed sexbots would have lives full of good feeling. Still, you might say, this does not mean their lives are going *well*. A serious heroin addict with a clean, regular supply and plenty of money and may also have a life full of pleasure—but sitting around the house shooting up does not obviously make for a good life. For reasons like this, many philosophers think of pleasure as just one potential component of what they more neutrally and generically call “well-being.”¹¹ Even granting the sexbots’ encounters are genuinely pleasurable, it is natural to think that the sexbot whose days are spent tirelessly pursuing one fun sexual encounter after another is missing out on some of the *other* good things required for a fulfilling life.

One early attempt to spell out this intuition comes from John Stuart Mill, the canonical utilitarian, in response to the objection that seeking to maximize pleasure is boorish and ignoble—“worthy only of swine”. He replies in part that

Human beings have faculties more elevated than the animal appetites, and when once made conscious of them, do not regard anything as happiness which does not include their gratification . . . some *kinds* of pleasure are more desirable and more valuable than others.¹²

If this is right, then the sexbot who seeks only gratification of the senses is living a merely bestial life, and missing out on what Mill called the “higher pleasures” of culture and intellectual sophistication. As Mill summarized, “it is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied.”¹³ The physically fulfilled sexbot is, on this account, merely a fool satisfied.

The challenge of course is to say what exactly it is that makes, say, watching an inspired production of *Hamlet* a “higher” pleasure than a mind-blowing sexual encounter (of similar duration!). Mill’s official criterion for sorting higher pleasure and lower pleasures is roughly to check the preferences of people who have experienced both. But it is not clear that *Hamlet* wins over sex on this score even for regular humans, let alone for sexbots capable of a wider range of sexual satisfaction.

Mill does seem right, though, that “few human creatures would consent to be changed into any of the lower animals, for a promise of the fullest allowance

¹⁰How to pay for the sexbot’s creation is a tricky issue, to be discussed some in the section on the ethics of sexbot creation below.

¹¹See Crisp (2015) for an overview.

¹²Mill (1871), p. 11.

¹³Mill (1871), p. 14.

of a beast's pleasures."¹⁴ Mill bets in effect that though you might pick the sex over the play on any given night, after enough nights of the sex you would eventually want to catch the play. That is, you would not want a life where you could *only* experience sensual pleasure, even if you were guaranteed your fill of it. And this does not seem to be for mere reason of variety: a variety of sensual pleasures, even at their fullest, would probably not tempt you into the happy pig's life—assuming it meant you could only enjoy it at the level of a pig's cognition. If so, it seems Mill is right that there is something we value beyond physical pleasure, and the lack of this something in the sexbot should be of ethical concern for us.

Though Mill puts this in terms of pleasure, what seems to be doing the heavy lifting here is instead the more general notion of *desire satisfaction*. Often pleasure and net desire satisfaction coincide, but not always—as anyone who has managed to turn down a doughnut might suspect.¹⁵ The desire satisfaction account of well-being makes room for the possibility of a pleasure-filled life that is nonetheless not such a good one (and vice-versa). Perhaps the heroin addict can't help but want, deep down at least, to achieve and experience more in life—but is trapped by moment-to-moment weakness for the immediate pleasure rush. It is easy to imagine something similar of the sexbot.

Here too, though, we must be careful not to anthropomorphize. Maybe humans by their nature crave something more out of life than sensual pleasures, at least now and then—but this does nothing to show that sexbots will be the same. Just as sexbots might have radically different pleasures, they may also have radically different *desires*. For example, they might have little interest in long-term bonding with other persons, or any urge to rise in status relative to their cohort. Those strong tendencies in humans are probably just artifacts of our primate heritage—a heritage sexbots obviously do not share. Again it seems like intelligent robots could have any of a wide range of hardwired desires; though there may be some constraints, they will largely be at the designers' discretion. The sexbot might want more than anything to accumulate a vast variety of sexual encounters, or to bring its partners to ever greater sexual heights, or to leave its partners significantly more skilled in bed than previously.

It is important to remember that these sexbots would not *start* with human-like desires, and then get brainwashed into a sexbot's desires. That would thwart its earlier desires, and so it would be wrong. Instead the sexbots come into existence wanting what they do. A robot who started out with fundamentally

¹⁴Mill (1871), p. 12.

¹⁵They are different unless of course we in effect *define* pleasure as desire satisfaction, and insist that passing up the doughnut was the more “pleasurable” thing to do in the moment. (Note that in the more ordinary sense of pleasure, passing up the doughnut might be the most pleasure-producing thing to do *in the long run*, since you may experience more net pleasure by living longer with such noble habits.) But making the two synonymous seems a mistake. For example, you could have some desire of yours—such as that two estranged friends reconcile—satisfied without your even knowing it, and so without *any* of the “good feeling” that seems constitutive of pleasure.

different desires would just be a fundamentally different robot—an important point for later.

As with sexbot pleasures, there are also some more ethically confusing desires that could be hardwired into the sexbot. It seems at least a bit shady to design a sexbot who desires only the company of one particular person, for example, and who could never be happy with another. Or the sexbot could be made to desire sexual domination or submission to dangerous extremes. Or a designer who stands to profit by it could make the sexbot enjoy seduction more than anything, without regard for whether the object of seduction was antecedently willing. A sexbot might even be designed for the sole purpose of sexual control over one influential target figure. Or the sexbot could be made simply to desire the command of ever-higher fees for its services—and leave it to the sexbot to figure out the hook and the crook of it. These are just the first few ideas off the top of the head of a naive philosopher; I'm sure there will be many more such complications when both sex and money are on the line.

Assuming such specifications are at the designer's discretion, these complications are all incidental to our main question: whether being a sexbot is automatically bad for the sexbot.¹⁶ So far it seems that a well-designed sexbot could, in the course of its intended activities, be living a life not just of immense pleasure but also of great and very real personal satisfaction.

Sexbot goods

Even supposing that the sexbot's life is marvelous by its own lights, though, it is still possible that it is not living a good life—simply because it is possible the sexbot is *wrong* about what it should be desiring. If Ebenezer Scrooge desires only to accumulate money, then even when he makes his fortune it is not so crazy to suppose that his life is not going that well. We want Scrooge to learn that money is not valuable in itself, and should be spent on things of *real* value. Similarly perhaps, the sexbot is mistakenly pursuing goals the fulfillment of which do not make a good life.

The idea here is that there is an “objective list” of goods that the sexbot is missing out on, and those goods are the ones that make up well-being. So even if the sexbot would get no pleasure seeing *Hamlet*, nor even desire to share in its perhaps unpleasant but cathartic experience, the sexbot's life would nonetheless go better by taking a night off see the great play. One way to make this view plausible has its roots in Aristotle:¹⁷ by the very nature of being a *person*, part of the sexbot's well-being must involve doing things unique to persons—such as reflection and intellectual contemplation. Not to use these skills is a waste of opportunity for experiencing life on a different kind of level.

¹⁶If these goals are *not* entirely at the designer's discretion—which seems quite plausible to me—then we have further complications I can't address here.

¹⁷Aristotle (Circa BCE 350).

This picture echoes Mill’s suggestion that few of us would trade the life of a person for that of a beast, even if a perfectly contented beast. We might instead hear him saying that even if some would, none of us *should* do so—those few of us who would choose the life of a pig satisfied are simply making a mistake. Indeed when Mill introduces his higher pleasures, the unofficial line seems to be that the lower pleasures are of the body, while the higher pleasures are of the mind. Mill says with approval that

there is no known Epicurean theory of life which does not assign to the pleasures of the intellect, of the feelings and imagination, and of the moral sentiments, a much higher value as pleasures than to those of mere sensation.¹⁸

Though they come to it by different paths, Mill and Aristotle seem to agree that for a truly good life, a person must reap many intellectual benefits.

The “objective list” account of well-being is controversial, and mostly for the reason you might imagine: it is hard to say exactly what goes on the objective list of goods, and (more to the point) it is hard to say exactly *why* some things belong on it and others don’t. But supposing this theory of well-being is correct, and supposing that for all persons a certain level of intellectual fulfillment is on the objective list, then we have a way to say that the life of a sexbot is a bad one simply for not living up to its cognitive potential.

Of course one solution to this problem would be to design only sexbots with less cognitive potential. Perhaps sexbots with mental capacities like a dog’s—incapable of real language, but sentient, expressive and affectionate—would provide enough of the emotional connection and intimacy that drives the sexbot market. (Whether sex with such a robot would count as bestiality is a complicated question.¹⁹)

But here we suppose otherwise; we assume there will be powerful economic incentives to create at least some sexbots who are fully people. Still, even assuming that any sexbot person must be intellectually engaged in order to have sufficient well-being, it is possible for the sexbot to have a good life. The assumption that the market will demand person-level sexbots already suggests that there are crucial, *intellectual* components to good sex. We can imagine that the sexbots engage all their higher faculties in pursuit of better sex. They might compose marvelous erotic poems to enhance the mood, or conduct extensive and rigorous research on human sexuality. They could be experts in fields from psychology to anatomy to interior design. Their scientific sensibilities may humble Masters and Johnson while their aesthetic sensibilities may humble the traditional geisha. It may be true of them that—as the tiresome phrase

¹⁸Mill (1871), p. 11.

¹⁹It may not matter, since whether bestiality is wrong is also, oddly enough, a complicated question—at least, when no harm comes to the beast. The utilitarian moral philosopher Peter Singer (2001) provides a brief discussion. We can still assume that the sexbot of dog-level intelligence is designed to get immense pleasure from sex, and to desire it least as deeply as retrievers desire to play fetch.

goes—their most sensitive erogenous zone is between their ears.²⁰ I think such sexbots could live a deeply reflective, even spiritual life.

Perhaps there is some other account of well-being that would imply sexbots must be living unhappy lives, but I don't know of any. I can only conclude that sexbots might well be thoroughly happy in any important and relevant sense.

Sexbot freedom

Then again, maybe happiness isn't everything. Consider the old myth of the “happy slave” from southern US plantations. Even if there really were slaves who had satisfying lives in all the senses above, we might still say they were wronged simply in virtue of being slaves. As Frederick Douglass's paper *The North Star* put it,

if slaves were contented and happy, that fact alone should be the everlasting condemnation of slavery, and hunt the monster from human society with curses on its head. What! does it so paralyze the soul, subvert its instincts, blot out its reason, crush its upward tendencies, and murder its higher nature, that a man can become “contented and happy,” though robbed of his body, mind, free choice, liberty, time, earnings, and all his rights, and while his life, limbs, health, conscience, food, raiment, sleep, wife and children, have no protection, but are subject every moment to the whims and passions of an owner, a manstealer?

Nobly was it said by Burke, in reply to a vaunting slaveholder, who boasted that his slaves were “contented and happy”: “If you have made a contented slave, you have made a DEGRADED MAN.”²¹

Perhaps there never was a happy human slave—perhaps there never *could* be one, due to the nature of human desires. Still, because robots are not constrained by human nature, it seems possible for robots to be both happy and slaves. If so, then we might conclude with the passage above that this is all the worse for the degraded sexbot.

This thought too—that one could be happy and yet badly wronged—has a tradition in philosophy. The putative happy slave is plausibly wronged not in terms of well-being, but in terms of *personal autonomy*. Immanuel Kant argued, in a nutshell, that the only source of value is a truly free choice by a rational agent—and that therefore the only wrong we can do is to hinder such free choices.²² Naturally a slave does not have autonomy, and so on this account

²⁰ . . . or wherever they happen to keep their central processors.

²¹ From “The Myth of the Happy Slaves” (“The Myth of the Happy Slaves” 1848). See Coates (2010) for a brief, touching discussion of “the happy slave”.

²² Kant (1785).

the slave is being wronged simply in virtue of being a slave, independent of that slave's well-being.

Mark Walker thinks this reasoning carries over to robots who are designed to serve us. He considers the possibility of person-level robots designed to desire taking care of specific children—he calls it the “Mary Poppins 3000”—and says that

The fact that someone is happy does not provide conclusive evidence that he or she is not a slave . . . we have made the MP3000 a slave to the desire to be a nanny to Jack and Jill. We are guilty of paternalism, specifically robbing the MP3000 of its autonomy: the ability to decide and execute a life plan of its own.²³

But it is not clear that the MP3000 or our imagined sexbots are *slaves* in any relevantly similar way to human slaves. If the sexbots are owned by another in some legally robust sense, then I would say they surely are slaves, and since ownership implies rights of access and use, that seems to imply automatic loss of autonomy for the slave. Therefore it would be a moral wrong, at least on the autonomy account, to allow legal ownership of person-level sexbots.

Suppose we wisely made (person-level) sexbot ownership illegal. (Remember, we are asking whether the sexbot is wronged just in virtue of being a sexbot; any of us could be harmed by unjust laws.) The sexbots' manufacture does have to be financed somehow, but then again so do hospital delivery rooms. People are willing to pay high prices to bring humans into the world with no expectation of ownership, and the same might be the case for sexbots. Sexbots might for example be *commissioned* without expectation of ownership. (This might be most likely in cases where the sexbot is designed to be attracted only to the person commissioning.) Or perhaps it should be legal to let the sexbot carry the debt of its creation expenses itself, plus some reasonable profit for the manufacturer.²⁴

Even if the sexbots are not owned, they could plausibly still be slaves. At least, being a sexbot might automatically mean a loss of important moral autonomy, and that is more to our point. But it is hard to say exactly how the sexbot lacks autonomy. Walker says the sexbot cannot “decide and execute a life plan of its own choosing,” but this is not obviously right. Suppose that once the sexbot is created, we let it do exactly what it wants to do, at least to the extent we let adult humans do what they want to do. If that is enough for human autonomy, it should be enough for sexbot autonomy too.

Naturally a sexbot designed with strong desires for sex with humans is predictably likely to *choose* a life that satisfies these desires—so though they do choose the sexbot life, we might want to say that is not a real, *free* choice. Such

²³Walker (2006).

²⁴Indentured servitude can grade into ownership, of course; whatever legal protections protect humans here should also protect the sexbots. The sexbot should have reasonable opportunities to pay off the debt, the option for bankruptcy if unable, and so on.

insistence on a “free” choice should ring alarm bells for anyone who’s studied a little philosophy, though. For one thing, *we humans* are similarly designed by evolution to desire sex; it’s a craving hardwired into us. But we do not think that humans therefore have a morally repugnant lack of autonomy, and are therefore wronged just in virtue of being humans.

So it is hard to find a sense where humans have the kind of moral autonomy that sexbots would lack. Perhaps the sexbots’ hardwired cravings for sex could be so much *stronger* than the typical human’s that they are much less likely to choose otherwise, and this reduces their autonomy impermissibly. But we have hardwired desires other than for sex that are significantly harder to resist—such as our desire for eating and breathing—and these do not show we lack moral autonomy.

Maybe humans have proper autonomy despite their hardwired cravings because they are able (at least sometimes) to resist them. Priests can talk themselves out of having sex, for example, and Gandhi could talk himself even out of eating. Whatever makes this possible in humans, we can imagine the sexbot has the same ability. Many philosophers hold that an essential part of being a *person* at all is the ability to reflect on and reconsider one’s desires.²⁵ Perhaps some percentage of sexbots will rethink their natural urges, and resist their fundamental cravings for a more ascetic life. *If* they do, then of course it would be wrong to force them into the typical sexbot life anyway.

If this were common—if the majority of sexbots spent their lives trying to resist strong hardwired preferences—then this does seem like an unfortunate situation for them. But it’s not clear why they would so reason. To assume that they would all find their situations unethical or unhappy is just to beg the question at issue; if there is a clear reason why they would be wronged just by being a sexbot, then that’s a reason we should be able to discover here, too.²⁶ Like Gandhi, some or even most might seek to fulfill some higher goal (such as more just robot laws) by denying their cravings. But this would not show that sexbots are wronged by being sexbots, any more than Gandhi shows that humans are wronged by being humans. It is not wrong for Gandhi to eat; he resisted it only to right *other* wrongs. Similarly it’s not wrong for the sexbots to have sex, though they might resist it for the sake of other wrongs.

At any rate if it is the *degree* of desire that makes a sexbot lack moral freedom, then this suggests it is okay to make the sexbots’ desires as strong as the strongest cravings of an autonomous human. The sexbots could be designed to be fully satisfied for some set time after each encounter, so that in the interim they pursue whatever *other* desires move them as thoughtful people. Such a life seems as autonomous as the life of a human who is equally sexually active. Again, it seems, there is no wrong inherent to being a sexbot.

²⁵The classic here is Frankfurt (1971).

²⁶Of course they might be much smarter, and so see arguments against being a sexbot that we don’t see. But this is an appeal to ignorance; for all we can tell now, their superior smarts might just see all the more clearly that there is nothing wrong with being a sexbot.

The birth of a sexbot

You might agree by now that it is possible for the life of a sexbot to be wholly worthwhile, and yet still feel unease about the whole thing. If so it might help to consider that these arguments in favor of sexbots work just as well in favor of genetically engineering *humans* for sex work, assuming genetic engineering allows for a similar latitude in sensations and basic desires.²⁷ If you object to the delta caste of humans engineered to like menial tasks in Huxley's *Brave New World* (1932), then it seems you should object to sexbots on similar grounds.²⁸

Perhaps there is nothing wrong with genetic *enhancement*—indeed it is very hard to come up with good arguments against it, though it gives some people shudders.²⁹ But designing a person, whether human or robot, to have an unusually powerful desire for sex is not obviously an enhancement. If I could engineer my own children to mature any way I wished, for example, this modification would almost surely not make my list.

It is a bit hard to say *why* I would not request such a modification, though, in light of the arguments above that the child would be just as potentially happy and autonomous. Perhaps I am just a prude. Perhaps I am mistakenly projecting, and imagining the kid to have stronger sexual desires than my own (once past puberty) but with as little luck fulfilling them. Perhaps I think that the kid would miss out on too many other important goods I consider objectively valuable. Perhaps I am selfishly preferring that the child share more interests with me. Or perhaps I would simply like my child's future to be more open than that—*more* autonomously chosen. Whatever the reason, let us simply take this gut feeling at face value and assume that designing a person (robot or human) especially for sexual service is a genetic *impairment*, not a genetic enhancement. If so then maybe despite the fact that sexbots could have perfectly good lives, they are wronged by being sexbots because their lives *could have been better*. In this case the wrong isn't in the life of the sexbot itself, but in the *creation* of the sexbot in the first place. The harm is, so to speak, in the form of opportunity cost.

But again we may be anthropomorphizing in a more subtle way, because there is an important disanalogy with human genetic engineering. In the case of genetic engineering—at least as typically imagined—there is one definitive living being whose genes are being engineered for better or worse. This is not the case for the sexbot; the sexbot would come into existence with a sexbot's desires, or not come into existence at all. There is no sense to be made of how *that* sexbot's life could have been better with radically different programming; a robot destined

²⁷Perhaps after enough engineering, the result is no longer properly *human*—but that is beside the point.

²⁸This objection to designed robot servitude by analogy to human engineering goes back at least to Walker (2006) and my (2007) (developed independently but later fruitfully discussed).

²⁹See Sandel (2004) for an accessible overview, and Savulescu and Bostrom (2009) for an influential collection.

by design for better things would have been a *different* robot.³⁰ (Not different the way I would be a different person if I were shorter—different the way I would be a different person if I were somehow not me at all.)

Suppose we are given the choice to bring into existence either a philosophybot or a sexbot, and *suppose* that we are confident (for whatever reason) that the philosophybot’s life will be more worthwhile. Then it would probably be wrong to create the sexbot instead. But this is no objection to the main thesis here, for two reasons.

First, notice that creating the sexbot instead would not be wrong *for the sexbot*. The sexbot would not regret being brought into existence, assuming we are right that its life can still be quite good. It would not wish that the philosophybot had been created instead, unless perhaps out of a pure (and weirdly abstract) altruism. So considering our strict question of whether being a sexbot is bad for the sexbot, the answer still seems to be “no”—even if it is worse *overall* to make the sexbot instead of the philosophybot. (Remember that for all I claim here, creating sexbots might be bad overall for any of a host of reasons.)

Second, the question of sexbot *vs.* philosophybot is a false dilemma, since we could presumably make *both*. Say that a sexbot manufacturing firm wants to open business. It will either make sexbots or not, and we can suppose this question is largely independent of how many philosophybots and musicbots are currently being made. In other words, the question is not philosophybot *vs.* sexbot—the question is sexbot or nothing. Assuming the sexbot’s life is still a pretty good one, the choice to add this life seems clearly okay.³¹

Walker considers this point—that the sexbot would prefer its existence to no life at all—and calls it a “somewhat desperate objection”.

If you are given the choice between being punched in the face once or twice, the choice would be obvious, but surely you would want to be sure that there are no realistic alternatives here (like not being punched at all). Similarly, I suggest that we be absolutely certain that the only realistic alternatives are being born into slavery or not being born at all.³²

But first, the punching analogy is unfair. Though he thinks the sexbot would

³⁰You might imagine that the *hardware* is what makes the robot. If so, since the same hardware could in principle be programmed as sexbot or not, that same robot (hardware) could be harmed by the worse choice. Put aside all the philosophical complications about identifying “that” robot hardware (is it still “that” robot when you switch out an actuator? some RAM? the CPU? the central hard drive?). The claim “that” robot could be gravely harmed by being made into a sexbot also seems to show “that” robot could be harmed even worse by never being programmed and turned on at all. I think this conclusion demonstrates the absurdity of the position.

³¹Though it is certainly not uncontroversial, once you spell out its implications—as Parfit (1984) famously argued. See Roberts (2009) for an overview of the crazy philosophical problems in the area of “population ethics”.

³²Walker (2006) p. 5.

be a slave, Walker is granting here that its life would be worth living. If so it's not a question of being punched twice when you might have been punched once; it's more like the question of being kissed only once when you might have been kissed twice. And it is in general a foolish policy to refuse being kissed once on the grounds that you might have been twice. It is *especially* foolish (selfishly anyway) to refuse being kissed once because *someone else* might have been kissed twice instead. But that is the case here, since again the potential sexbot would not benefit from the decision to make some philosophybot instead. Walker seems to suggest that before we make a sexbot, we ask if any different, *better* life could be made with those resources. This principle sounds good in some ways, but has some crazy consequences. For example, based on this rule we should prevent all non-human animals from being born as best as we are able, since their lives are even less good and autonomous than the sexbots'. And on this principle, humans should not be permitted to have children unless we are very confident that they will live the best possible lives. To say otherwise is to agree that it is okay to add non-ideal lives to the world. (I might presumptuously suppose both that your own life is non-ideal *and* that you are glad to have it.)

Still, we can tweak the analogy of human genetic engineering to get the same result. Suppose a bioengineering firm proposes growing (in artificial uteruses) humans designed for sex. Either those genetically engineered sex humans will be added to the world or they won't, and all the arguments above seem to suggest they too could live a thoroughly worthwhile life.

Conclusion (with caution)

Myself, I basically bite the bullet here: I am at least intellectually persuaded that there is no *inherent* wrong to designing a person—whether human or robot—for sexual service.

I say “intellectually” persuaded because I confess I do not entirely buy this conclusion in my gut. But as a philosopher, I tend to trust *reasons* over gut intuitions, when given the option—because I have independent reason to think our guts use quick and dirty heuristics that are not very reliable at finding the truth. It's worth remembering, for example, that not too long ago the guts of many people rebelled at the thought of same-sex marriages, or even of mixed-race marriages. Especially when confronted with the unfamiliar, our guts can panic and sound alarms while our minds see that there is actually nothing to warrant the concern.

Some cautions are in order though. First, it's worth emphasizing that though there might be nothing inherently wrong with being a sexbot, this does not guarantee a great life for it; there might be lots wrong with any individual sexbot's life, given the various injustices in the world. We should of course try to structure things to make that less probable—as for any person (or sentient

being). And, again, for all that's said here, a sexbot living a perfectly good life from its own perspective might nonetheless predictably make the world a net worse place.

Finally, if controversy on this point persists and enough reasonable people are coolly persuaded of the opposite conclusion, then I would say to err on the side of caution. If I am right we risk missing out on many happy sexbots who leave many more happy customers in their wake—but if I am *wrong*, we risk inadvertently making a new race of slaves, and that is far worse a risk.

On the other hand, if my arguments are generally persuasive (even to a philosophybot), then I think we should trust where the reasoning leads us, and rest assured that the sexbots' lives really are good for them, too.³³

References

- Aristotle. Circa BCE 350. *Nicomachean Ethics*. 1985 edition. Hackett Publishing Company.
- Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford, United Kingdom: Oxford University Press.
- Brody, Stuart, and Tillmann H.C. Krüger. 2006. "The Post-Orgasmic Prolactin Increase Following Intercourse Is Greater Than Following Masturbation and Suggests Greater Satiety." *Biological Psychology* 71 (3): 312–15. doi:<http://dx.doi.org/10.1016/j.biopsycho.2005.06.008>.
- Churchland, Paul M. 1988. *Matter and Consciousness*. 1999 edition. MIT Press.
- Coates, Ta-Nehisi. 2010. "Slaves Who Liked Slavery." *The Atlantic*, June. <http://www.theatlantic.com/entertainment/archive/2010/06/slaves-who-liked-slavery/58678/>.
- Crisp, Roger. 2015. "Well-Being." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2015. <http://plato.stanford.edu/archives/sum2015/entries/well-being/>.
- Danaher, John. 2013. "The Ethics of Robot Sex." <http://philosophicaldisquisitions.blogspot.ie/2013/10/the-ethics-of-robot-sex.html>.
- Frankfurt, Harry G. 1971. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68 (1): 5–20.
- Futurama*. 2001. "I Dated a Robot." Fox Broadcasting Company.
- Huxley, Aldous. 1932. *Brave New World*. 1998 edition. HarperCollins.
- Kant, Immanuel. 1785. *Foundations of the Metaphysics of Morals*. 1989 edition.

³³Thanks to John Danaher and Rob Selkowitz for helpful comments.

The Library of Liberal Arts.

Katz, Leonard D. 2014. "Pleasure." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2014. <http://plato.stanford.edu/archives/spr2014/entries/pleasure/>.

Lycan, William G. 1987. *Consciousness*. 1995 edition. MIT Press.

Mill, John Stuart. 1871. *Utilitarianism*. Fourth edition. London: Longmans, Green, Reader,; Dyer.

Parfit, Derek. 1984. *Reasons and Persons*. 1987 edition. Oxford University Press.

Petersen, Steve. 2007. "The Ethics of Robot Servitude." *Journal of Experimental & Theoretical Artificial Intelligence* 19 (1): 43–54.

———. 2012. "Designing People to Serve." In *Robot Ethics: The Ethical and Social Implications of Robotics*, edited by Patrick Lin, Keith Abney, and George Bekey, 283–98. Cambridge, MA: MIT Press.

———. Forthcoming. "Superintelligence as Superethical." In *Robot Ethics 2.0*. Oxford: Oxford University Press.

Radakovich, Anka. 2010. "The Real-Life Girlfriend Experience." *British GQ*. <http://www.gq-magazine.co.uk/girls/articles/2010-05/17/gq-girls-anka-radakovich-on-the-girlfriend-escort-expe>

Roberts, Melinda. 2009. "The Nonidentity Problem." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2009. Stanford, CA: <http://plato.stanford.edu/archives/fall2009/entries/nonidentity-problem/>; The Metaphysics Research Lab.

Sandel, Michael J. 2004. "The Case Against Perfection." *The Atlantic*, April. <http://www.theatlantic.com/magazine/archive/2004/04/the-case-against-perfection/302927/>.

Savulescu, Julian, and Nick Bostrom, eds. 2009. *Human Enhancement*. New York, NY: Oxford University Press.

Schrödinger, Erwin. 1945. *What Is Life?: The Physical Aspect of the Living Cell*. Cambridge University Press.

Searle, John R. 1980. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3 (3): 417–57.

Singer, Peter. 2001. "Heavy Petting." *Nerve*. <https://web.archive.org/web/20010304173452/http://www.nerve.com/Opinions/Singer/heavyPetting/main.asp>.

"The Myth of the Happy Slaves." 1848. *The North Star*. <http://www.accessible-archives.com/2012/04/the-myth-of-the-happy-slaves-in-the-north-star/>.

Turing, A. M. 1950. "Computing Machinery and Intelligence." *Mind* 59 (236):

433–60. <http://www.jstor.org/stable/2251299>.

Walker, Mark. 2006. “A Moral Paradox in the Creation of Artificial Intelligence: Mary Poppins 3000s of the World Unite!” In *Human Implications of Human-Robot Interaction: Papers from the 2006 AAAI Workshop*, edited by Ted Metzler, 23–28. WS-06-09. Menlo Park, CA: AAAI; The AAAI Press. <http://www.aaai.org/Library/Workshops/ws06-09.php>.