7

The Symbolic-Consequences Argument in the Sex Robot Debate

John Danaher

7.1 Introduction

The television series *Humans* is a provocative and sometimes insightful drama about social robots.[1] It depicts a near-future in which realistic humanoid robots have become commonplace, acting as workers, home helpers, carers, and sexual playthings for their human creators. The majority of the robots are less-than-human in their intelligence and ability, and apparently lack sentience, but the main plotline concerns a particular group of these robots that has achieved human-level consciousness and intelligence. They struggle for freedom and

respect in a world in which their robot brethren are treated with either condescension or contempt.

In one episode, a group of (human) teenagers are having a house party. At the house party there is a robot serving drinks and catering to the attendees' needs. The robot looks like a human female. Some of the young men hurl abuse at her. One of them switches her off and then tells his mates that he is going to drag her upstairs to have sex with her. He is goaded on. At this point one of the main (human) female characters intervenes, telling her male peers to stop. When asked why, she responds by asking them whether it would be okay for them to knock out a real human female and have sex with her in similar conditions? They renege on their plan.

The writers of the show do not pause at this point and have the female protagonist expand on her objections. Like all good fiction writers they have learned to 'show not tell.' But I'm interested in the telling. Presumably the objection to the young men having sex with the switched-off robot had nothing to do with the potential harm to the robot. The robots within the show are—apart from the core group—deemed to be devoid of moral status, lacking the requisite consciousness and intelligence. They are—to use a phrase repeated in other chapters of this book—not *moral victims*. So why is it wrong for the young men to have sex with them in the suggested manner? The answer must lie elsewhere:

in the symbolic meaning of the act, and the consequences that might ensue from its permission.

As it happens, this combined concern for symbolism and its consequences is a common feature of several objections to the development and use of sex robots. Indeed, it is possibly the leading style of objection to sex robots in the current, admittedly small, literature. My goal in this chapter is to provide a detailed analysis of it, outlining its abstract structure, giving specific examples of its use, and evaluating its merits.

I will defend three main claims. First, I will agree with proponents of the symbolic-consequences argument that there are plausible grounds for thinking that sex robots will be symbolically problematic, both in how they represent human beings and in how they encourage a particular style of sexual interaction with those representations. Nevertheless, I will temper this conclusion by suggesting that this problematic symbolism is not essential, or incorrigible, or decisive. It can be *removed* and *reformed* under the right circumstances. Second, I will argue that this means that the consequences of the symbolism becomes all important. Will it cause people to act out in other problematic ways? Will it result in harm to the individual user or to the society in which they live? I'll argue that it is exceptionally difficult to answer those questions prior to the development of the technology, and this leaves us in an uncertain position regarding the strength of

the symbolic-consequences argument. Third, I'll suggest that the best way to address this uncertainty is to approach the development of sex robots as a social experiment, i.e., as something that should be subject to similar logistical and ethical standards as medical or psychological experiments.

Although I am interested in the symbolic meaning of love and sex with robots in general, I present all three of these arguments with a particular category of symbolic meaning and consequence in mind, namely: what does having sex with robots say about our understanding of consent to sex and the ethics of interpersonal sexual relationships? And what might the consequences of having sex with robots be for our attitudes and practices with regard to sexual consent and interpersonal sexual relationships? Given this focus, it behooves me to start with a brief primer on sexual consent and its relevance to the sex robot debate.

7.2　The Importance of Consent Norms to the Sex Robot Debate

I'll start by outlining the importance of consent in human-to-human sexual relationships. I'll then explain how it is relevant to the sex robots debate. I will work from first principles, beginning with some platitudes about the value of sexual experience in human life and the role that consent plays in ensuring its value.

It is relatively uncontroversial to say that sexual activity is an important and highly valued part of the human experience. In addition to being a source of pleasure, sexual activity is, for many people, a mark of intimacy and maturity. It provides the basis for a unique, mutual, and intersubjective bond. But it also has a dark side. Unwanted, coerced, or forced sexual activity can be physically and emotionally traumatizing, sometimes leading to lifelong personal and interpersonal difficulties.[2] Consequently, it is important to develop a system of sexual norms that distinguishes permissible sexual activity from impermissible sexual activity—preventing and punishing the latter, while, if not encouraging, at least facilitating, the former. In short, society needs to create a set of norms that protects *negative sexual autonomy* and facilitates *positive sexual autonomy*.[3] For most people, and most legal systems, *consent* is now deemed to be the 'moral magic' that performs this crucial function.[4] Consent is what ensures that the partners to the sexual act are willing (and hopefully enthusiastic) co-conspirators.

But what is consent and how do we ensure that it is present? In the human-to-human context, the answers to these questions are complex and controversial.[5] Westen (2004), for instance, argues that there are at least four distinct consent concepts that operate in moral and legal discourse. His framework for thinking about sexual consent distinguishes between consent as *subjective attitude* (i.e., willingness to accept or go along with something, not necessarily equivalent to a

desire) and *objective performance* (i.e., the communication of signals of willingness to another party). It also distinguishes between *factual* consent (i.e., what a person actually communicated and felt about an act) and *prescriptive* consent (i.e., the normative standards of communication we as a society demand or prescribe). The normative goal in human-to-human relationships is to ensure that the objective prescriptive performance matches the subjective factual attitude: i.e., that we communicate and act upon signals that are representative of our subjective willingness to engage in sexual activity. But it is often hard to craft workable guidelines to ensure that this happens. This is because it is difficult to figure out what a person's subjective attitude actually is, apart from the objective signals representing that attitude. So, when setting normative standards, we tend to focus on objective performance—and then run into the problem that there are many conflicting and ethically dubious views about when and whether an objective performance can be taken to signal consent. Some people think that certain styles of clothing and flirtatious behavior signal consent. Others think that a clearly communicated 'no' can mean yes, or that lack of resistance is a sign of encouragement. These views are morally flawed, but historically common.

Questions surrounding the appropriate norms of consent have become particularly notorious in recent years. There has been a highly publicized "crisis" of sexual assault and rape on university campuses. According to some US studies,

between one-in-five[6] and one-in-four women[7] are likely to experience unwanted sexual contact[8] during their time at university. Similar figures are reported in other countries. In Ireland (where I am located), a 2015 study of leading universities revealed that between one-in-seven and one-in-four women were victims of unwanted sexual contact.[9] The studies also suggest (as is true in non-university cases) that these incidents go underreported and under-prosecuted.

Studies of this sort have been the subject of criticism.[10] Some critics argue that such studies give inflated figures due to the language used in the surveys.[11] Some argue there are discrepancies in the figures that go unexplained. But even if these criticisms are correct, the likely 'true' number is still too high,[12] and most would agree that something ought to be done to address the problem. One of the more interesting solutions to the problem is to insist upon *affirmative consent standards* in sexual ethics. This is something that is now mandated in certain US states.[13] Affirmative consent standards stipulate that sexual contact is only permissible if there are clear and unambiguous affirmative signals of consent. No longer will people be able to infer consent from lack of resistance, clothing, and flirtatious behavior. More is needed.

How is any of this relevant to the sex robot debate? I, along with most other contributors to this volume,[14] believe that sex robots are unlikely to be moral persons. In other words, I believe they will (for the foreseeable future

anyway) lack the inner subjective life that makes consent so important in the human context. It might consequently seem that consent is completely irrelevant to the evaluation of sex robots. To talk about consent in the human-to-robot context is to commit a category mistake: to apply a concept that ought not to be applied.

Yet, this view seems to me to be in error. The framework distinguishing between actual subjective attitudes and prescribed objective performances allows us to see why. It is true that if robots are not moral persons, then they cannot be *victims* of unwanted sexual contact. But the robots themselves will presumably engage in objective performances in response to their users. Thus, they might respond approvingly, or disapprovingly, to their users' sexual advances.[15] These objective performances will either symbolically mimic or differ from the normatively accepted consent standards in society at large. This means that both the robot itself (in its appearance and behavior) and the act of having sex with the robot will have important symbolic properties when it comes to norms of sexual consent and interpersonal sexual ethics. The presence of these symbolic properties is what opens up the door to the symbolic-consequences argument.

7.3   The Symbolic-Consequences Argument

As I mentioned in the introduction, the symbolic-consequences argument is popular in the contemporary debate about sex robots, particularly among those who object to the development and use of sex robots. But its popularity is implicit rather than explicit. Most proponents of the argument do not express it using the terminology of 'symbolic-consequences.' They make what they take to be unique and distinctive arguments. Thus, when I say that it is 'popular,' I am making a potentially controversial claim. I am saying that there is a common argumentative structure underlying many objections to sex robots. In this section, I want to identify that structure and illustrate it with examples from the literature.

The common argumentative structure is as follows:

(1) Sex robots do/will symbolically represent ethically problematic sexual norms. (*Symbolic Claim*.)

(2) If sex robots do/will symbolically represent ethically problematic sexual norms, then their development and/or use will have negative consequences. (*Consequential Claim*.)

(3) Therefore, the development and/or use of sex robots will have negative consequences *and* we should probably do something about this. *(Warning Call Conclusion)*

Some comments about this abstract formulation are in order.

First, the ethically problematic symbolism could take many forms. It could be linked to the robot's appearance and demeanor, or to the act of sexually engaging with the robot. For instance, in the consent case, it could be that the robot encourages the user to engage with it in a way that ignores or positively flouts the socially accepted norms of consent. It could also be that the physical representation of the robot embodies negative sexual stereotypes. Perhaps the robot represents a certain style of female appearance (maybe a "porn star"-esque style)? The behavior or movement of these sex robots may also be problematic, e.g., they may behave in an overly deferential, coquettish manner, representing women as submissive and subordinated creatures.

Second, the negative consequences of the symbolism could take many forms, some more immediate and direct than others. It could be that the user is directly and immediately harmed by the interaction with the robot. It could also be that the development and use of the robots sends a negative signal to the rest of society, thereby reinforcing a culture of sexism, misogyny, and/or sexual objectification. This "expressive" consequentialism is common in other symbolic objections to cultural practices.[16] The interaction with the robot could also have downstream effects on the user, changing his/her interactions with other human

beings and thereby having a harmful impact on those others as well. The negative

consequences need not be a dead certainty: they could have varying degrees of

probability attached to them. This is normal enough in a debate about a nascent,

emerging technology (it's normal enough in any debate about the consequences of

technological usage). But the uncertainties may make it difficult to draw firm

normative conclusions. I return to this problem later.

Third, the conclusion is something of a non sequitur in its current form.

The first part follows logically from the premises; the second part does not.

Nevertheless, I have tacked on this "warning call" because I think it is common in

the debate: most purveyors of these arguments think we ought to do something to

minimize the potential negative consequences. What this "something" should be

is another matter. Some people favor organized campaigns against the

development of sex robots;[17] others favor strong to weak forms of regulation.[18]

I have presented the abstract structure. Are there specific examples that

flesh out the premises in more detail? Indeed there are. I'll briefly describe three.

The first comes from the work of Sinziana Gutiu. She provides the most

extensive consent-based version of the argument so I will discuss her version at

the greatest length. Her starting presumption is that the majority of sex robots will

be targeted at heterosexual males and will depict a problematic, stereotypically

"ideal" woman. She defends this presumption by reference to literary precursors

to sex robots (e.g., the long-standing trope of male protagonists constructing ideal female partners, present for instance in the Adam and Eve myth) and current examples of robotic technology. Some of these current technologies do not involve actual sexbots (i.e., robots designed for sexual use) but do involve gynoid robots (robots designed to look and act like women) that are highly sexualized: "Aiko, Actroid DER and F, as well as Repliee Q2 are representations of young, thin, attractive oriental women, with high-pitched, feminine voices and movements. Actroid DER has been demoed wearing either a tight hello kitty shirt with a short jean skirt, and Repliee Q2 has been displayed wearing blue and white short leather dress and high-heeled boots."[19]

Current sex-robot prototypes (e.g., Roxxxy and the models from RealDoll) would seem to follow suit. For Gutiu, then, the physical structure of female robots alone serves to represent problematic norms of body shape, dress, and movement. The problematic symbolism is compounded when robots are designed for sexual use. As Gutiu puts it: "To the user, the sex robot looks and feels like a real woman who is programmed into submission and which functions as a tool for sexual purposes. The sex robot is an ever-consenting sexual partner and the user has full control of the robot and the sexual interaction. By circumventing any need for consent, sex robots eliminate the need for communication, mutual respect, and compromise in the sexual relationship. The use of sex robots results in the

dehumanization of sex and intimacy by allowing users to physically act out rape fantasies and confirm rape myths."[20]

It seems, then, that Gutiu fleshes out the first premise of the argument in the following manner:

(1*) Sex robots will symbolically represent ethically problematic sexual norms because (a) the majority will adopt gendered norms of body shape, dress, voice, and movement (e.g., they will be thin, large-breasted, provocatively clad, coquettish in behavior, and so on—this could vary from society to society); and (b) they will function as ever-consenting sexual tools, bypassing any need for mutual communication and mutual respect, and allowing users to act out rape fantasies and confirm rape myths.

She then turns to the negative social consequences of this symbolism. She distinguishes between two sets of harms. First, there are the obvious *social harms* arising from the symbolism. If the robots represent gendered norms of sexualized appearance and sexual compliance, they will contribute to and reinforce a patriarchal social order that is harmful to women. In particular, they will further distort our understanding of sexual consent. Campaigners have been fighting hard to make changes to the law surrounding rape and sexual assault. The changes made to date try to combat rape myths by clarifying the nature of sexual consent and assigning appropriate weight to the testimony of victims. Sex robots would

represent a step back in this fight because "they embed the idea that women are passive, ever-consenting sex objects, and teach users that when getting consent from a woman, 'only no means no.'"[21]

In other words, they would go against the recent reforms of consent standards and in particular the push for positive affirmative signals of sexual consent. This could obviously have an impact on women, who become victims of actual sexual assault and rape if users act out in the real world.

Second, in addition to the social harms and harms to others, there are the *harms to the users themselves*. For one thing, the users could internalize the problematic sexual norms through repeated use of the robots, which could alter their moral character and the nature of their interactions with other people. Also, and somewhat in tension with this idea, the robots could reinforce antisocial tendencies among users, encouraging them to withdraw from social interactions, and avoid the need for mutuality and compromise in their sexual lives. This latter notion was contradicted in the film *Lars and the Real Girl*. In that film, the use of a sex doll was therapeutic and enabled an introverted man to reintegrate with society. However, Gutiu dismisses this:

> Although it was an effective approach to a Hollywood film, sex
> robots are unlikely to help antisocial users better interact with
> women. It is doubtful that an individual who does not feel accepted

in society, and who finds an alternative way to meet their exact needs for companionship will, for some reason, want to integrate back into society, where they can risk rejection and face social discomfort.[22]

This suggests that Gutiu fleshes out the second premise of the argument in the following manner:

(2*) If sex robots adopt gendered norms of body shape, dress, behavior, etc., and function as ever-consenting sexual tools, their creation and use will: (a) reinforce patriarchal social norms and distort our understanding of sexual consent, which will ultimately harm women; and (b) will harm the users by encouraging them to internalize problematic sexual norms, and, for some, exacerbate their antisocial tendencies.

This, in turn, leads to the "warning call" conclusion. Gutiu thinks that something should be done to combat the problematic symbolism and likely negative consequences. She does not favor prohibition of sex robots. Instead, she favors various regulatory interventions. These could include, in particular, the demand that creators design robots in a certain way. They could also include the creative use of legal mechanisms to allow potential victims of harm arising from the use of sex robots to sue for damages. As an example, she suggests that a

person whose marriage dissolves after their partner starts using a sex robot be allowed to sue the manufacturer. This might seem unusual, but there are legal mechanisms (so-called "heart balm torts") that allow people to sue others for interfering with a legally protected relationship, so the idea is not without precedent.

A second variation on the symbolic-consequences argument can be found in the work of Kathleen Richardson and *The Campaign Against Sex Robots.*[23] This work is discussed and critiqued at length elsewhere in this volume,[24] so I will only offer a brief summary here. The major objection to sex robots in Richardson's work stems from what she perceives to be the analogy between human-sexbot interactions and human-prostitute[25] interactions. She believes that the current development of sex robots is being modeled on a particular understanding of the interactions between humans and sex workers. In other words, the goal of the designers and creators of sex robots is to create an interactive experience between the robot and the human user that is roughly equivalent to the interaction between a sex worker and their client. The robots consequently symbolically represent that style of interaction. She cites the work of David Levy in support of this view.

This is problematic for two reasons. First, human-sex worker interactions are themselves ethically problematic. They are based on asymmetries of power.

The client's will and interests dominate over those of the sex worker. There is no concern for the inner mental life, wants, or needs of the worker. The sex worker is thus objectified and instrumentalized. By symbolically mimicking such interactions, sex robots represent approval for this style of interaction. Second, in doing so, sex robots will encourage their users to perpetuate negative attitudes toward women. This will reinforce a misogynistic and patriarchal culture in which women are subordinated and oppressed. Richardson thinks we should respond to these problems by instituting an organized campaign against the development of sex robots. This argument fits very much within the symbolic-consequences model.

A final variation on the symbolic-consequences argument comes from some of my own work. In a paper published a couple of years back,[26] I suggested that there *might* (I was tentative) be reason to outlaw the manufacture and/or use of certain kinds of sex robot on essentially symbolic grounds. In particular, I singled out robots that were designed to cater to rape fantasies and pedophiliac tendencies. My argument was intended to be *purely symbolic* in nature. I suggested, following the work of Stephanie Patridge, that there was something intrinsically wrong with our reactions to certain symbolic representations.[27] In this sense, the person who enjoys having sex with a robot that mimics resistance to sexual advances or that is designed to look like a child is analogous to the person

who laughs at a racist joke or enjoys racist artworks. They express something about their moral character that is worthy of social condemnation. On some occasions, this may be sufficiently serious to warrant legal prohibition. In this manner, my argument didn't really appeal to consequences at all. Nevertheless, I did suggest (as I will suggest again below) that consequences are always relevant to the ethical evaluation of symbolic representations as they may serve to outweigh or reinforce the problems with the symbolism. The crucial question is: Do the negative/positive consequences outweigh or reinforce the problems with the symbolism? This is actually an exceptionally difficult question to answer and may warrant a whole new approach to the development of sex robots.

But this is to get ahead of the argument. For now I want to move away from outlining the structure of the symbolic-consequences argument to a critical evaluation of its two main premises. Is the symbolism of sex robots likely to be problematic? If so, how? And how can we evaluate the alleged consequences of this symbolism? I answer these questions over the next two sections.

7.4   Are Sex Robots Symbolically Problematic?

To determine whether sex robots are symbolically problematic, we first need a better understanding of symbolic value and its importance in human social life. Andrew Sneddon's paper on the topic is instructive in this regard.[28] It makes two

claims that are relevant to the present inquiry. The first is that there are two distinct ways in which symbols can be valuable; the second is that symbols are valuable because they govern the relational aspects of human life. I want to briefly explain both of these claims because I think they serve to highlight the strength of the symbolic claim in premise one of the argument.

Let's start with a general account of symbols. Following C. S. Peirce's work on representation, Sneddon argues that symbols exist when three things are present: (1) a *symbolic object or practice*, i.e., some object or practice that is taken to *stand for* or *represent* something else; (2) an *interpreter*, i.e., someone who decides that the object stands for or represents something else; and (3) a *ground for interpretation*, i.e., something that justifies or supports the interpreter's take on what the symbol stands for. A painting is a symbolic object: the lines of paint on the canvas are taken to represent and stand for something by the person viewing the painting (this could be some event in the real world, some commentary on religion or politics, or some reflection of the artist's inner turmoil). The viewer's interpretation can be justified on a number of grounds (e.g., the *similarity* between the lines of paint and some event or object in the real world, some conventional or proposed theory of art, or some *causal relationship* between what was going on in the painter's mind when they were creating this artwork and the artwork itself).[29] Even with this simple example we see that

symbolic representation is a complex thing. A symbolic object can be taken to represent many different things on many different grounds. Furthermore, we see that symbolism is distinct from communication: symbols can exist without some original communicator who is trying to convey a message. All that matters for symbolism is that you have the object, the interpreter, and the grounds for interpretation.

This account of symbols applies straightforwardly to the sex robot case. Take the arguments in the previous section. The proponents of these arguments are the *interpreters*. The sex robots (real or imagined) are the *symbolic objects* that are taken to stand for, or represent, something else by the interpreters. What they are taken to stand for or represent varies slightly between the interpreters. They all agree that the robots will tend to stand for or represent women (or, in my case, also possibly children). Furthermore, they all think that they stand for a particular understanding of women (or children) as sexual playthings. There are then some differences in terms of how the behavior and interaction with the robots stands for something else. Richardson, for instance, thinks that the interaction represents the relationship between a sex worker and a client; and Gutiu thinks it represents a problematic set of beliefs about norms of consent and the status of women. The interpreters then justify or support their interpretations on various grounds. The most obvious ground is the *resemblance* between the

robots and the real world human beings and actions they represent. The intentions

of the creators and users are also additional, supporting grounds for the

interpretations. To me, at any rate, this understanding of the symbolism of sex

robots makes sense. With relatively few exceptions,[30] these robots are created and

desired because they provide some kind of facsimile of a sexual encounter with a

real human being. They are not simply devices for sexual stimulation or release—

we already have those—they are something more, by virtue of what they

represent.

But why does it matter? Why should we care what they represent if the

robots themselves are not moral victims? This is where Sneddon's distinctions

between the different kinds of symbolic value, and the social importance of

symbols, is relevant. Sneddon claims that there are two distinct ways in which

symbols can be valuable (or disvaluable, as the case may be). The first is that they

can be valuable in *virtue of what they are taken to represent*. This is obvious

enough, but it has some important repercussions. If the symbol is valuable (or

disvaluable) in virtue of what it represents, then you must first understand the

value of what it is taken to represent before you can understand the value of the

symbol. Thus, the disvalue that attaches to sex robots that are taken to represent

women (or children) as passive, ever-consenting sexual playthings must be

understood in terms of the disvalue that attaches to the view that women (or

children) actually *are* passive, ever-consenting sexual playthings. The history of sexual violence and oppression, the disregard for individual autonomy and rights, the harm and trauma that results from unwanted sexual contact—these are all reasons to balk at the notion that women (or children) should be understood in these terms. These reasons carry over (in an attenuated form) to the symbolic representations (i.e., the sex robots). The symbols thus share in the disvalue of what they represent.

The second way in which symbols can be valuable (or disvaluable) is *in and of themselves*, i.e., apart from what they represent. Sneddon says that the 'N-word' (i.e., 'nigger') is a strong case of this. Although the disvalue attaching to the N-word originated in real world practices of abuse and oppression, the word itself has now taken on such an incendiary aura that to even *mention* it in discourse (as opposed to *use* it as a term of abuse) is taken to be problematic. To prove the point: I suspect many people reacted negatively when they saw the real word being mentioned by me a couple of sentences ago, even though I wasn't using the word to refer to anyone or any group. Something doesn't sit right with its mere presence on the page. This is why the euphemism "N-word" has become common. People want some way to refer to the symbol without actually using it. This second type of symbolic value is rare in its purest form. But oftentimes there is a hybrid form of symbolic value where the symbol is valuable (or disvaluable)

by virtue of what it represents; but this doesn't fully explain the value attaching to it—there is something intrinsic to the symbol as well. It is difficult to see how this could happen in the case of sex robots. But it might. It might be that even mentioning or referring to sex robots takes on a negative (or positive) aura regardless of what they are taken to represent. Indeed, there is a sense in which this is already true. When I say to colleagues that I am writing and editing a book about sex robots, they seem to get immediately uncomfortable and dismissive. This might be because their minds instantly conjure up images of potential sex robots, and they then think about what the robots might represent, but it might also be that the mere mention of the concept is doing all the work. It is a difficult thing to disentangle. Fortunately, it does not matter going forward. For the remainder of this section, I will simply assume that the disvalue attaching to sex robots arises by virtue of what they are taken to represent and not from anything intrinsic to the robots themselves.

This still doesn't quite tell us why symbols are valuable or disvaluable. We know that they can be valuable (or disvaluable) in two distinct ways, but we don't know why they acquire this value (or disvalue) in the first place. The answer to that question lies in the importance of human sociality and the role of symbols in mediating and facilitating human social life. Human beings are a social species. Key moments in our technological and social history are typically

marked by increases in social cooperation and coordination.[31] Symbols are

essential to this progress. Anthropologists and historians have often commented

upon this.[32] For example, in his surprise best-selling book *Sapiens* (2011), Yuval

Noah Harari argues that human social evolution has been marked largely by our

ability to create fictional, abstract structures that we overlay onto our physical

reality. These fictional structures get reinforced and communicated through

symbolic representations. The most obvious and important of these, of course, are

the languages we use to encode and communicate our beliefs, laws, customs, and

norms. But other symbolic representations play a part too, from national flags and

sculptures, to scientific theories, to works of architectural beauty and wonder. All

of these things serve to create a heavily symbolic social environment in which we

live out our lives. These symbols dictate social roles and social beliefs. They tell

us how we should relate to, and understand, one another.

This provides support for Sneddon's claim that symbols are valuable

precisely because they govern the purely *relational* aspects of human life. In this

regard, they are distinct from other sources of value, such as harms/benefits and

rights/duties. Harms and benefits are, in Sneddon's vocabulary, *constitutively* and

evaluatively individualistic. In other words, harms/benefits are things that happen

to, or accrue to, individual human beings, and we care about them because of

what they do to individuals. Symbols, on the other hand, are both constitutively

and evaluatively relational. They are constitutively relational because they are made up of objects, signs, practices, etc., that represent or stand for something else. Thus, they always stand in relations to both human interpreters and that which is being represented in symbolic form. Furthermore, they are evaluatively relational because they are important by virtue of how they mediate the relationship we have with others and the world around us. Thus, a racial slur is (negatively) value-laden because of what it says about the relationship between the user of the slur and the person or race in question. The same goes for the use of a sex robot with symbolically disturbing properties. Its use says something about the relationship between the user (and the society that facilitates the user) and the people or group represented in the robotic form.

Where does this leave us with respect to premise one of the symbolic-consequences argument? It seems to leave us with much to be said in its favor. Given the centrality of symbols in human social life, proponents of these arguments have reason to be concerned. They seem to be justified in suggesting that, at least some (and maybe many), sex robots will be taken to stand for and represent our attitudes toward real people (specifically, women and children) due to both their resemblance to real people and the intentions of the creators and users. Furthermore, it seems plausible to suggest that they will tend to represent those real people in a particular way: as ever-consenting sexual playthings. It is

hard to escape this interpretation of the symbolism. If sex robots are designed and marketed for sexual use, the user will want them to be available and ready for use whenever they are switched on. They are unlikely to have an appetite for the mutual conversation and objective performances demanded by our consent norms. Since this could be taken to symbolically encode a disregard for preferred norms of sexual consent, it seems plausible to say that there is something symbolically disvaluable about sex robots. The same logic applies to other aspects of the symbolism (e.g., the gendered beauty norms, the asymmetry of power, the lack of mutual respect).

But the argument cannot end there. The problematic symbolism of sex robots is contingent in two important ways: it is *removable* and *reformable*. It is possible to embrace the symbolic critique without rejecting the permissibility of sex robots. With regards to *removability*, it is important to remember that the appearance and behavior of sex robots is not some Platonic essence that is fixed and irrevocable. Sex robots need not be large-breasted, thin-waisted, porn star-esque waifs.[33] No doubt there will be significant pressures in favor of this representation.[34] But it is conceivable that one could create and design a sex robot to look and act more like a 'real' woman; to represent a more progressive set of norms around sexual consent and beauty, and interpersonal relations. For instance, the robot could be programmed so as not to be an "ever-consenting"

sexual tool. The robot might sometimes randomly refuse its user, and always provide positive affirmative signals of consent when it is willing to proceed. Enforcing and ensuring a more positive set of representations might be a good target for regulation in this area. Furthermore, to the extent that robots are designed to cater to rape fantasies or pedophiliac tendencies, this is something that could be outlawed or banned. In short, it is conceivable to imagine a world in which sex robots do not share the problematic symbolism highlighted by the arguments discussed in the preceding section. Whether it is possible to realize that world is another matter.

This brings us to the second important way in which the symbolism of sex robots is contingent. Some people might resist the suggestions in the previous paragraph on the grounds that it is very difficult to avoid the problematic symbolism involved in the creation of a robot that looks like a real woman and is to be used solely for sexual purposes.[35] This argument, however, ignores the fact that symbolic interpretations are, in virtually all cases, *polysemous* and *highly contested*. It is often only because we live in a particular cultural environment—with its own set of socially accepted symbolic interpretations—that we fail to see this contingency. Brennan and Jaworski provide a fascinating insight into this phenomenon in their discussion of symbolic objections to markets.[36] To many moral philosophers, the idea of paying for certain goods and services (mourners at

funerals, sex, best man speeches, kidneys) necessarily leads to the moral tainting of those goods and services.[37] To pay your spouse for sex, they say, would necessarily corrupt the intimacy and mutuality of the marital relationship, reducing it to a cold and emotionless commercial transaction. But not all cultures share this belief. In the Merina tribe of Madagascar, it is expected that husbands pay their wives after sex as a sign of respect. To the Merina tribe, money does not symbolize distance or a lack of affection. Quite the contrary, in fact.[38] And it is not just money whose social meaning is contingent either. Brennan and Jaworski discuss several other examples illustrating the social and cultural contingency of the meaning that attaches to symbolic practices. The most famous example is the social meaning that attaches to our treatment of the dead. According to Herodotus, the Persian King Darius once noted the discrepancies between Greeks and Callation cultural norms on this score. The Greeks thought that you expressed respect for the dead by burning their bodies on a funeral pyre; the Callations thought that this was to treat the dead like a piece of trash. They preferred to eat them instead. Needless to say the Greeks were abhorred by this notion.

This contingency of symbolic meaning has important implications for how we think about the symbolic meaning of sex robots. At the moment, we may well live in a culture that attaches negative meanings to the representation of women (and potentially children) as sexual playthings. But this could be capable of

radical change. There could be (distant) future cultures where having sex with a robot does not carry the same negative connotations. It may actually signal safety and respect. Don't misunderstand this claim. To say that the social meaning that attaches to sex robots can be radically altered in this manner is not to say that we should radically alter it. It is simply to say—as Brennan and Jaworski point out—that the meaning of a symbolic practice cannot be treated as a given in our ethical analysis. The meaning of the practice is itself up for ethical scrutiny, and, under the right circumstances, there might be strong moral grounds for thinking that we should reform the meaning that attaches to the practice. What circumstances might these be? The cultural meaning of dead bodies is, again, instructive:

> [C]onsider that some cultures developed the idea that the best way to respect the dead was to eat their bodies. In those cultures, it really was a socially constructed fact, regardless of one's intentions, that failing to eat the dead expressed disrespect, while eating rotting flesh expressed respect. But now consider that the Fore tribe of Papua New Guinea suffered from prion infections as a result of eating the rotten brains of their dead relatives prior to that practice being banned in the 1950s. The interpretative practice of equating the eating of rotting flesh with showing respect is a destructive, bad practice. The people in that culture have strong moral grounds to change what expresses respect.[39]

The point is clear. In some cases, the *consequences* of sticking with a particular set of social meanings can be destructive. The Fore tribe's belief that they should respect the dead by eating their brains has such destructive consequences that it needed to be changed. The same conclusion can be drawn in relation to the symbolism of sex robots. Thus, while I might be inclined to agree with Gutiu and Richardson (and myself-of-three-years-ago) that the current social meaning of sex robots is problematic, that is not the end of the story. The consequences of having robots with that problematic symbolism turns out to be the critical factor. If the consequences are positive, then we may need to reform the symbolic meaning.

7.5   Are Sex Robots Consequentially Harmful?

We have reached a critical point. If the argument in the preceding section is correct, then there may well be problems with the symbolism of (at least some) sex robots, but that problematic symbolism is likely to be contingent in two important ways: (1) the particular features of the robots that warrant the problematic interpretation might be removed and changed; and (2) the social meaning of any symbolic representation, no matter how strongly negative it seems to be, is capable of being reformed. This contingency means that the consequences of the symbolism become all important. We may then logically ask: Are the design, manufacture, sale, and use of sex robots likely to reinforce and

exacerbate the problematic symbolism? Or could these factors have positive

consequences that are capable of outweighing (and thus warranting changes in)

the symbolic interpretation?

Proponents of the symbolic-consequences argument claim that the

consequences will exacerbate and reinforce the problems with the symbolism.

Recall Gutiu's claims about how gendered sex robots will cause their users to

withdraw from society and/or interact with real women in problematic ways.

Richardson echoes these claims in her 'campaign' against sex robots. Both have

plausible-sounding arguments for believing that these negative consequences will

follow. Someone who has sex with a robot on a regular basis may grow

accustomed to the belief that their sexual partners should always be 'ready to go.'

They may grow frustrated with the need for mutual agreement and meaningful

consent in human-to-human relationships. This may cause them to withdraw from

such relationships, or to be more aggressive in those sexual encounters. Either

way, the consequences would seem to be bad for our collective attempts to

improve the norms of sexual consent and interpersonal sexual relationships. On

top of this, there may be other, more subtle and difficult-to-assess effects. The

mere presence and cultural acceptance of symbolically problematic sex robots

might have negative consequences for the experience of women living in the

societies that accept their existence. The women might feel less welcome and less

respected. They might acquire a 'false consciousness' about their position and place in society.

The problem with these plausible-sounding arguments is that they need to be weighed against other, often equally plausible-sounding, arguments suggesting that the consequences might not be as bad as we just supposed. There are a few possibilities to consider. For one thing, the design and use of robots that cater to, say, rape fantasies or pedophiliac tendencies might have a *cathartic* as opposed to *emboldening* effect on their users. In other words, the robots might create a "safe space" in which these problematic sexual desires can be expressed without harming others. This "cathartic" view of human desire is contentious, but if utilized in the right therapeutic setting—perhaps with complementary psychotherapy—it is possible that these robots could be used to wean people away from their problematic desires and dispositions. More generally, sex robots that are designed to symbolically represent more progressive attitudes toward women and sexual consent could be used to educate young people as to the socially accepted sexual norms. Thus, far from reinforcing patriarchal and misogynistic attitudes, the robots might help to undermine them. On top of this, there are, as other contributions to this volume suggest (e.g., McArthur, Di Nucci), positive consequences that may ensue from the development of sex robots, including the improvement of the access to, and satisfaction of, positive

sexual rights. These consequences would need to be weighed against competing negative consequences.

What do we do with these contradictory, plausible-sounding arguments? I will suggest a modest, skeptical response—similar to the response that I have given elsewhere.[40] I will suggest that plausible-sounding arguments are not going to be enough. To decide who has the better of it, we need good empirical evidence. And we simply do not have that at the moment because we do not have many sex robots in existence, and so we do not have any empirical studies of their uses and effects. All we have are analogies with other, potentially similar phenomena, like hardcore pornography. And those analogies are not encouraging.

People have long worried about the negative effects of pornography on users and the societies in which they live. People worry that regular exposure to, and use of, pornography will have addictive effects, causing the user to constantly seek out new 'highs' in their pornographic viewing, and alter the users' attitudes toward sexual behavior and (in particular) women. Over the years, thousands of experimental and epidemiological studies have been published supporting different views on this question. Many studies do indeed find that users of pornography are (slightly, but significantly) more likely to embrace promiscuity,[41] engage in risky sexual behavior,[42] have worse relationships,[43] have disturbing attitudes toward women, and be more likely to engage in acts of sexual

aggression.[44] But other studies dispute these claims, suggesting that users of pornography are more likely to have progressive attitudes toward women,[45] that pornography can be associated with positive relationship outcomes,[46] and that correlations between pornography use and sexual aggression fail to disentangle cause and effect (i.e., higher pornography consumption may be an effect of negative attitudes and aggression, not a cause).[47] Claims regarding the addictive effects of pornography are also hotly disputed.[48] And virtually every researcher in this field laments the low quality and biased nature of the available evidence.[49]

This is not encouraging in two respects. It suggests that finding out the consequences of sexual symbolism is exceptionally difficult. And it suggests that the evidence we end up with may be ambiguous and disputed—which would be of little help to proponents or opponents of the symbolic-consequences argument.

Of course, the analogy between pornography and sex robots is imperfect. I have argued elsewhere that the embodied nature of the interaction between the robot and the user may have stronger causal effects than the consumption of pornography.[50] In viewing pornography, there is some psychological distance between the user and the symbolic object; in the case of sex robots there is a direct and immediate interaction with the symbolic object. But this argument is still speculative and it's unclear in which direction the stronger causal effects

might flow. Will it have an emboldening or cathartic effect? At this point in time, we just don't know.

7.6   An Experimental Approach to Sex Robots

This leaves us in a tricky position. We have grounds for thinking that at least some of the symbolic properties of sex robots are ethically problematic, but that these properties are contingent in two respects (*removability* and *reformability*). We also have grounds for thinking that the consequences will be the decisive factor, but that if analogous case studies are any guide, these consequences are going to be exceptionally difficult to work out. I want to conclude by arguing that this state of affairs should encourage us to take an explicitly experimental approach to the development of sex robots.[51]

In this respect, I am influenced by the work of Ibo van de Poel, and his colleagues, on new technologies as social experiments.[52] To understand their thinking, take the case of the iPhone (or smartphones, more generally) and ask yourself a simple question: What was Apple thinking when they introduced this product back in 2007? It was an impressive bit of technology, poised to revolutionize the smartphone industry, and set to become nearly ubiquitous within a decade. The social consequences were to be dramatic. Looking back, some of those consequences have been positive: increased connectivity, increased

knowledge, and increased day-to-day convenience. But a considerable number of the consequences have been quite negative: the assault on privacy, increased distractability, endless social noise. Were any of these possible consequences weighing on the mind of Steve Jobs when he stepped onstage to deliver his keynote on January 9, 2007? Some possibly were, but more than likely they leaned toward the positive end of the spectrum. Jobs was famous for his "reality distortion field"; it's unlikely he allowed the negative to hold him back for more than a few milliseconds. It was a cool product and it was bound to be a big seller. That's all that mattered. But when you think about it, this attitude is pretty odd. The success of the iPhone and subsequent smartphones has given rise to one of the biggest social experiments in human history. The consequences of near-ubiquitous smartphone use were uncertain at the time. Why didn't we insist on Jobs giving it a good deal more thought and scrutiny? Imagine if instead of an iPhone he was launching a revolutionary new cancer drug? In that case, we would have insisted upon a decade of trials and experiments, with animal and human subjects, before it could be brought to market. Why are we so blasé about information technology as compared to medication?

As Van de Poel notes, technologies like the iPhone have two key properties at their time of launch: (1) they have significant impact potential (i.e., they could change society in dramatic ways); and (2) they have unknown and

uncertain effects. Sex robots would seem to share these two properties. If the arguments in this chapter (and throughout this book) are correct, sexbots have significant impact potential. And, as I just pointed out above, they definitely have unknown and uncertain effects. This does not mean we should ban or prevent their creation, assuming this is practical (see Danaher, Earp, and Sandberg in this volume), but it should give us some pause. There is a well-known 'control dilemma' associated with the launch of any new technology with significant impact potential.[53] During the early phases of development, the technology will be easy to control and change in response to feedback, but its social effects will be poorly understood. But during later phases, as the technology becomes more ubiquitous and its social effects (possibly) better understood, it will be effectively impossible to control and change.

This presents policymakers and innovators with a difficult choice. Either they choose to encourage the technological development, and thereby run the risk of profound and uncontrollable social consequences, or they stifle the development in the effort to avoid unnecessary risks. Both choices seem far from optimal. This conundrum inherent in innovation has led to a number of controversial and (arguably) unhelpful approaches to the assessment of new technologies. Developers are encouraged to conduct cost-benefit analyses of any new technologies with a view to bringing some quantificational precision into the

early phase. This is then usually overlaid with some biasing-principle such as the *precautionary principle*—which leans against permitting technologies with significant impact potential—or the *procautionary principle*—which does the opposite. We can imagine such principles being applied to the development and use of sex robots. People who emphasize the potentially negative consequences are likely to favor the precautionary approach; people who emphasize the potentially positive consequences are likely to favor the procautionary one.

This is not a satisfactory state of affairs. These solutions focus on the first horn of the control dilemma: they try to con us into thinking that the social effects are more knowable at the early phases than they actually are. Van de Poel suggests that we might be better off focusing on the second horn. In other words, we should try to make new technologies more controllable in their later phases by taking a deliberately experimental and incremental approach to their development. Approaching new technologies as social experiments will require both a perspectival and practical shift. It will require us to think about the technology in a new way and put in place practical mechanisms for ensuring effective social experimentation. These practical mechanisms will have epistemic and ethical dimensions.

On the epistemic side of things, we need to ensure that we can gather useful information about the impact of technology and feed this into ongoing and

future experimentation. This means that as the technology is developed and made available to users, logistical frameworks need to be put in place to ensure that we can gather data on the social and personal effects of the technology. This is important in the case of sex robots, because, if analogous cases are anything to go by, it may be difficult to gather data after the technology has been released. If we want to avoid the endless and empirically unsatisfactory avalanche of studies that have become common in the pornography debate, we need to do something now, while we still have control. Tracking and surveillance of users may be the most plausible course of action (since tracking and surveillance is often built-in to new technologies)—but this leads to ethical problems (discussed below).

On the ethical side of things, we need to ensure that our ongoing and incremental experiments with the technology will respect certain ethical principles. One of Van de Poel's major contributions to the social experiment debate is his attempt to develop a comprehensive framework of principles for ethical technological experimentation. He does this by explicitly appealing to the medical analogy. Medical experimentation has been subject to increasing levels of ethical scrutiny since World War II. Detailed theoretical frameworks and practical guidelines have been developed to enable biomedical researchers to comply with appropriate ethical standards. The leading theoretical framework is probably

Beauchamp and Childress's *Principlism*. This framework is based on four key

ethical principles:

Non-maleficence: Human subjects should not be harmed.

Beneficence: Human subjects should be benefited.

Autonomy: Human autonomy and agency should be respected.

Justice: The benefits and risks of experimentation ought to be fairly distributed.

These four principles are general and vague. The idea is that they

represent widely shared ethical commitments that can be developed into more

detailed practical guidelines for researchers. Again, one of the major strengths of

Van de Poel's work is his review of existing medical ethics guidelines (such as

the Helsinki Declaration and the Common Rule) and his attempt to code each of

those guidelines in terms of Beauchamp and Childress's four ethical principles.

He shows how it is possible to fit the vast majority of the specific guidelines into

those four main categories. The only real problem is that some of the guidelines

focus on who has responsibility for ensuring that the experimentation follows the

guidelines, not on the four principles used by Beauchamp and Childress. This is

something that is important in relation to the development of non-medical

technologies too. Concern about responsibility and liability gaps are rife in the

literature about social robotics (see Di Nucci, and Danaher, Earp, and Sandbergin this volume).

These Helsinki and Common Rule guidelines were developed with the vagaries of medical experimentation in mind. We need something that can apply to a technology like sex robots. This requires some adaptation and creativity. Van de Poel has come up with a list of sixteen conditions for ethical technological experimentation. They are illustrated in the table below, which also shows how they map onto Beauchamp and Childress's principles.

[Insert Table 7.1 Here]

These guidelines are relatively self-explanatory, but I will briefly run through the main categories and discuss how they might apply to the experimental development of sex robots.

As you can see, the first seven are all concerned with the principle of non-maleficence. The first condition states that other means of acquiring knowledge about a technology must be exhausted before it is introduced into society. So manufacturers of sex robots should acquire knowledge about their effects on users and the reactions of others before releasing them more generally to consumers. The second and third conditions demand ongoing monitoring of the social effects of technology and efforts to halt the experiment if serious risks become apparent.

This will require some ongoing tracking and monitoring of initial beta users to ascertain social effects. This, of course, brings with it certain privacy and autonomy risks, which will need to be addressed through appropriate data protection laws and informed consent provisions (see below). The fourth condition focuses on the containment of harm. It accepts that it is impossible to live in a risk-free world and to eliminate all the risks associated with technology. Nevertheless, harm should be contained as best it can be. So if we learn early on that particular forms of human-sex robot interaction have harmful effects, we should act to mitigate and contain those harms as soon as possible. The fifth, sixth, and seventh conditions all encourage an attitude of incrementalism toward social experimentation. Instead of trying to anticipate all the possible risks and benefits of technology, we should try to learn from experience and build up resilience in society so that any unanticipated risks of technology are not too unsettling.

The next two conditions focus on beneficence and responsibility. Condition eight stipulates that whenever a new technology is introduced there must be some reasonable prospect of benefit to the user and to society at large. This is quite a shift from current attitudes. At the moment, the decision to release a technology is largely governed by economic principles: what matters is whether it will be profitable, not whether it will benefit people. I think the condition of

benefit can probably be met in the case of sex robots (other contributions to this book outline some of the reasonable prospects of benefit), but there must also be clear acknowledgment of and respect for the potential harms. Condition nine is about who has responsibility for ensuring compliance with ethical standards. This is an important condition for those who are interested in the legal side of this debate. Those who develop and release these technologies should do so in a responsible and socially conscientious fashion. They should be made to reflect on the potentially negative consequences of releasing a sex robot that clearly represents some problematic symbolism and be forced to take legal responsibility for their decision to do so. Furthermore, their decision to do so should be scrutinized in light of the other principles in this framework.

Conditions ten to thirteen are all about autonomy and consent at both an individual and societal level. Condition ten requires that those who use and may be affected by the technology are properly informed as to the risks and benefits. This will require that appropriate educational and informational materials be provided to people who purchase and are affected by the use of sex robots. Condition eleven says that majority approval is needed for launching a social technological experiment. This might be the most controversial element of the framework. It suggests that decisions about when and whether to launch a new technology with high impact potential should not be a left solely in private,

corporate hands. It should be opened up to public scrutiny and debate. I agree that public deliberation about the merits of developing sex robot technology would be a good thing. And, in some sense, I hope that this book and the contributions it contains can play a part in that public debate. But I am not sure that a "majority approval" condition is either practical or desirable. Van de Poel himself notes that this could lead to the tyranny of the majority—with majority groups imposing technological experiments on the minorities who are most affected by them. Conditions twelve and thirteen try to mitigate for this potential tyranny by insisting on meaningful participation for those who are affected by the technology, including a right to withdraw from the experiment. This would seem to be most important in the case of sex robots, particularly if the symbolism is most likely to implicate minority groups or those who lack political power, but how one could ensure a right to 'withdraw' from the experiment is unclear.

The final set of conditions all relate to justice. They too should help to mitigate the potential for a tyranny of the majority. They insist that the benefits and burdens of any technological experiment be appropriately distributed, and that special measures be taken to protect vulnerable populations. Condition sixteen also insists on reversibility or compensation for any harm done. This is where something like Gutiu's earlier proposal about the use of civil liability laws could become appropriate. If the great "sex robot experiment" backfires, and

adversely affects women or children or other more specific groups of people, then facilities should be put in place to ensure that these adverse effects can be compensated for, and, where possible, reversed. Explicit consideration for ways in which to distribute the benefits and burdens should also help to determine which side of the symbolic-consequences debate should be allowed to win out.

This experimental approach is certainly not a panacea. But it does encourage a more thoughtful, less knee-jerk, approach to technological developments like sex robots.

7.   Conclusion

Let me conclude by returning to the opening example: the attempt by the teenagers to have sex with a switched-off robot in the TV show *Humans*. Recall how one of the female protagonists objected to this on the grounds that they would not do this to a real woman. What I have argued in this chapter is that her objection can be spelled out in terms of the symbolic-consequences argument. The problem with switching off a robot and having sex with it lies not in the harm it does to the robot, but rather in what it symbolizes—a general disregard and/or contempt for norms of consent in interpersonal sexual relationships—and the potential negative effects of that symbolism—harm to real women and/or harm to

the user of the robot. Several contributors to the current sex robot debate have voiced similar objections.

But this style of objection faces several hurdles. While there are grounds for thinking that sex robots could symbolically represent a troubling attitude toward women (and maybe children) and the norms of interpersonal sexual relationships, the troubling symbolism is likely to be contingent in two ways. It is likely to be *removable* in many instances and *reformable* in others. What will ultimately matter are the consequences of the symbolism. These consequences are going to be difficult to work out. There are plausible-sounding arguments in favor of positive consequences and plausible-sounding argument in favor of negative consequences. What we lack is data. To address these problems, I suggest that we adopt an explicitly experimental approach to the development of sex robots. This approach should be guided by ethical principles and should build in logistical frameworks that allow for experimental data to be gathered and fed back into the process of incremental development.

Adopting this experimental approach will be a difficult thing to do. It will require significant changes in our perspective and attitude toward technological development. But it may be our best bet if we are to avoid the risks associated with developing this potentially high impact technology.

Table 7.1. Van De Poel's Principles for Ethical Technological Experiment.

Non-maleficence: Do no harm by ensuring …

1. Absence of other reasonable means for gaining knowledge about risks and benefits.

2. Monitoring of data and risks while addressing privacy concerns.

3. Possibility and willingness to adapt or stop the experiment.

4. Containment of risks as far as reasonably possible.

5. Consciously scaling up to avoid large-scale harm and to improve learning.

6. Flexible setup of the experiment and avoidance of lock-in of the technology.

7. Avoiding experiments that undermine resilience.

Beneficence: Do good by ensuring that it is …

8. Reasonable to expect social benefits from the experiment.

Responsibility: **Be sure that there is a …**

9. Clear distribution of responsibilities for setting up, carrying out, monitoring, evaluating, adapting, and stopping the experiment.

Autonomy: Respect autonomy by ensuring that …

10. Experimental subjects are informed.

11. The experiment is approved by democratically legitimized bodies.

12. Experimental subjects can influence the setting up, carrying out, monitoring, evaluating, adapting, and stopping of the experiment.

13. Experimental subjects can withdraw from the experiment.

Justice: Ensure that there is a fair distribution of the benefits and burdens of the technology by ensuring that…

14. Vulnerable experimental subjects are either not subject to the experiment or are additionally protected or particularly profit from the experimental technology (or a combination).

15. A fair distribution of potential hazards and benefits.

16. Reversibility of harm, or, if impossible, compensation for harm.

Notes

1. *Humans* originally aired in June 2015. For details see,

http://www.channel4.com/programmes/humans.

2. To be clear, in identifying these two classes of sexual activity I am not

ignoring the possibility that much sexual activity is neutral or not particularly

valuable. I ignore this "middle" category since consent would still be essential to

its permissibility.

3. See Di Nucci in this volume.

4. Heidi Hurd, "The Moral Magic of Consent," *Legal Theory* 2, no. 2 (1996):

121–46.

5. For useful overviews see: Alan Wertheimer, *Consent to Sexual Relations*

(Cambridge: Cambridge University Press, 2003); Peter Westen, "Some Confusion

about Consent in Rape Cases," *Ohio State Journal of Law* 2 (2004): 333–59; Peter

Westen *The Logic of Consent.* (London: Ashgate, 2004); Douglas Husak, "The

Complete Guide to Consent to Sex: Alan Wertheimer's *Consent to Sexual*

Relations," *Law and Philosophy* 25, no. 2 (2006): 267–87; Kimberly Ferzan,

"Clarifying Consent: Peter Westen's '*The Logic of Consent,*'" *Law and*

Philosophy 25 (2006): 193–217; and Heidi Hurd, "Was the Frog Prince Sexually

Molested? A Review of Peter Westen's *The Logic of Consent*," *Michigan Law Review* 103, no. 6 (May 2005): 1329–46.

6. Christopher Krebs, Christine Lindquist, Tara Warner, Bonnie Fisher, and Sandra Martin, "The Campus Sexual Assault Survey," National Institute of Justice, Washington, DC (2007),

https://www.ncjrs.gov/pdffiles1/nij/grants/221153.pdf.

7. David Cantor, Bonnie Fisher, Susan Chibnall, and Carol Bruce, "Report on the AAU Campus Climate Survey on Sexual Assault and Sexual Misconduct," The Association of American Universities (2015),

http://www.aau.edu/uploadedFiles/AAU_Publications/AAU_Reports/Sexual_Assault_Campus_Survey/Report%20on%20the%20AAU%20Campus%20Climate%20Survey%20on%20Sexual%20Assault%20and%20Sexual%20Misconduct.pdf.

8. I use this term, rather than "rape or sexual assault" because of the controversies alluded to in the text about the language used in the surveys supporting these figures.

9. J. Humphreys, "11% of Women Students Believe They Were Victims of Sexual Assault," *Irish Times*, June 4, 2015. The 11% figure comes from focusing

on the past year; the one-in-four figure comes from focusing on how many people were victims of sexual assault and/or attempted assault.

10. For instance, compare the figures in the two US-based surveys with the figures in the National Victimization of Crime survey. This survey suggests lower overall incidence and higher incidence for non-students: Sinozich, S. and Langton, L *Rape and Sexual Assault Victimization Among College-Age Females, 1995–2013* (Department of Justice, Bureau of Statistics, November 2015) available at http://www.bjs.gov/content/pub/pdf/rsavcaf9513.pdf

11. Libby Nelson, "Why Some Studies Make Campus Rape Look Like an Epidemic While Others Make it Look Rare," *Vox,* December 11, 2014, http://www.vox.com/2014/12/11/7378271/why-some-studies-make-campus-rape-look-like-an-epidemic-while-others.

12. There is an interesting question here as to whether any figure higher than zero is acceptable. In principle, the answer is no: we obviously do not want anyone to be the victim of rape or sexual assault. In practice, it may be impossible to reduce the incidence of a particular crime to zero.

13. California Senate Bill 967 Student Safety: Sexual Assault makes it a condition of funding for the state's universities to adopt affirmative consent

standards; New York adopted a similar piece of legislation in Senate Bill S5965. Both pieces of legislation affect proceedings within universities and do not change the definition of consent for rape and sexual assault within the criminal law.

14. Cf. chapter by Petersen in this volume.

15. This is apparent from existing prototypes. Roxxxy—who is billed as the world's first sex robot—comes with preprogrammed personality types. One of them ('Frigid Farah') is said to respond to the user's advances with disapproval. See http://www.truecompanion.com/shop/faq.

16. Jason Brennan and Martin Jaworski, "Markets Without Symbolic Limits," *Ethics* 125, no. 4 (2015): 1053–77.

17. Danaher, Earp, and Sandberg in this volume.

18. Sinziana Gutiu, "The Roboticization of Consent," in *Robot Law*, eds. Calo, Froomkin, and Kerr (Cheltenham, UK: Edward Elgar Publishing, 2016); Sinziana Gutiu, "Sex Robots and the Roboticization of Consent," *We Robot Law Conference Miami* (2012), robots.law.miami.edu/wp-content/uploads/2012/01/GutiuRoboticization_of_Consent.pdf; John Danaher, "Robotic Rape and Robotic Child Sexual Abuse: Should They be Criminalized?"

Criminal Law and Philosophy, forthcoming (2015), doi: 10.1007/s11572-014-9362-

19. Gutiu, "Sex Robots," 5.

20. Ibid., 2.

21. Ibid., 15.

22. Ibid., 17.

23. Kathleen Richardson, "The 'Asymmetrical' Relationship: Parallels Between Prostitution and the Development of Sex Robots," *SIGCAS Computers & Society* 4, no. 3 (September 2015): 290–293, https://campaignagainstsexrobots.org/the-asymmetrical-relationship-parallels-between-prostitution-and-the-development-of-sex-robots.

24. Danaher, Earp, and Sandberg in this volume.

25. Richardson prefers the term "prostitute" to "sex worker." She argues that the latter tends to legitimize the practice. I tend to prefer "sex worker" and revert to that for the remainder of this chapter.

26. John Danaher, "Robotic Rape and Robotic Child Sexual Abuse: Should They be Criminalized?" *Criminal Law and Philosophy*, forthcoming (2015), doi: 10.1007/s11572-014-9362-.

27. Stephanie Patridge, "The Incorrigible Social Meaning of Video Game Imagery," *Ethics and Information Technology* 13, no. 4 (2010): 303–12; Stephanie Patridge, "Pornography, Ethics and Video Games," *Ethics and Information Technology* 15, no. 1 (2013): 25–34.

28. Andrew Sneddon, "Symbolic Value," *Journal of Value Inquiry* (September 2015), doi: 10.1007/S10790-015-9519-4.

29. The most prominent grounds of representation mentioned by Sneddon include *resemblance*, *convention*, *causal connection*, and *stipulation*.

30. What do I mean by this? I think, perhaps, some people might like robotic sexual partners *qua* robots. They might like, and prefer, them to real human beings precisely because they do not view them as being equivalent to or symbolically similar.

31. As an illustration, see Ian Morris, *Foragers, Farmers, and Fossil Fuels* (Princeton: Princeton University Press, 2015).

32. Terrence Deacon, *The Symbolic Species: The Co-Evolution of Language and the Brain* (New York: WW Norton, 1998).

33. A comment is apposite here: I am uncomfortable about shaming any particular style of bodily appearance. I write about the "'porn star'-esque" body shape in a somewhat pejorative fashion because I take it that proponents of the symbolic consequences argument have that pejorative view.

34. Paul Treanor (2015) argues that this is the dominant male preference in his commentaries on sex robots. See Paul Treanor, "Bimboid Utopia: What do Men Want?" (2015), https://politicalaspects.wordpress.com/2015/02/06/bimboid-utopia-what-do-men-want.

35. Cf. Frank and Nyholm in this volume.

36. Brennan and Jaworski, "Markets," (2015).

37. Michael Sandel, *What Money Can't Buy* (London: Penguin, 2012).

38. The example of the Merina tribe is taken from Viviana A. Zelizer, *The Social Meaning of Money* (Princeton: Princeton University Press, 1996) and is discussed in Brennan and Jaworski (2015).

39. Brennan and Jaworski "Markets without Limits," 1067.

40. John Danaher, "New Technologies as Social Experiments: An Ethical

Framework," in *Philosophical Disquisitions* (2016),

http://philosophicaldisquisitions.blogspot.ie/2016/03/new-technologies-as-social-

experiments.html.

41. Dolf Zillman and Jennings Bryant, "Effects of Prolonged Consumption of

Pornography on Family Values," *Journal of Family Issues* 9, no. 4 (1988a): 518;

Dolf Zillman and Jennings Bryant, "Pornography's Impact on Sexual

Satisfaction," *Journal of Applied Social Psychology* 18, no. 5 (1988b): 438–453;

Dolf Zillman and Jennings Bryant, "Effects of Massive Exposure to

Pornography," in *Pornography and Sexual Aggression*, eds. Neil Malamuth and

Edward Donnerstein, (Amsterdam: Elsevier, 1984); and Jochen Peter and Patti M.

Valkenburg, "Adolescents and Pornography: A Review of 20 Years of Research,"

Journal of Sex Research 53, nos. 4–5 (March 2016): 509–31. doi:

10.1080/00224499.2016.1143441.

42. Matija Sinković, Aleksandar Stulhofer, and Joan Božić, "Revisiting the

Association Between Pornography Use and Risky Sexual Behaviors: The Role of

Early Exposure to Pornography and Sexual Sensation Seeking," *Journal of Sex

Research* 50, no. 7 (2013): 633–41. doi: 10.1080/00224499.2012.681403; Emily

Harkness, Barbara Mullan, and Alex Blaszczynski, "Association Between

Pornography Use and Sexual Risk Behaviors in Adult Consumers: A Systematic Review," *Cyberpsychology Behavior and Social Networking* 18, no. 2 (January 2015): 59–71, doi: 10.1089/cyber.2014.0343.

43. Samuel L. Perry, "Does Viewing Pornography Reduce Marital Quality Over Time? Evidence from Longitudinal Data," *Archive of Sexual Behavior*, doi: 10.1007/s10508-016-0770-y.

44. I treat sexual aggression and disturbing attitudes toward women as a pair since most studies focus on heterosexual pornography. See Gert M. Hald, Neil M. Malamuth, and Carlin Yuen, "Pornography and Attitudes Supporting Violence Against Women: Revisiting the Relationship in Nonexperimental Studies," *Aggressive Behavior* 36, no. 1 (January 2010): 14–20, doi: 10.1002/ab.20328; Drew A. Kingston, Neil M. Malamuth, Paul Fedoroff, and William L. Marshall, "The Importance of Individual Differences in Pornography Use: Theoretical Perspectives and Implications for Treating Sexual Offenders," *Journal of Sex Research* 46, nos. 2–3 (April 2009): 216–32, doi: 10.1080/00224490902747701; and Paul J. Wright, Robert S. Tokunaga, and Ashley Kraus, "A Meta-Analysis of Pornography Consumption and Actual Acts of Sexual Aggression in General Population Studies," *Journal of Communication* 66, no. 1 (2015): 183–205, doi: 10.1111/jcom.12201.

45. Taylor Kohut, Jodie L. Baer, and Brendan Watts, "Is Pornography Really about 'Making Hate to Women'? Pornography Users Hold More Gender Egalitarian Attitudes Than Nonusers in a Representative American Sample," *Journal of Sex Research* 53, no. 1 (2016): 1–11, doi: 10.1080/00224499.2015.1023427.

46. The evidence here is suggestive only. In a self-selecting sample it was found that some couples were unaffected or positively affected by pornography use: William A. Fisher and Lorne Campbell, "Perceived Effects of Pornography on the Couple Relationship: Initial Findings of Open-Ended, Participant-Informed, "Bottom-Up" Research," *Archives Sexual Behavior* (July 2016), doi: 10.1007/s10508-016-0783-6. The following study finds that *female* use of pornography can be correlated with positive relationship outcomes: Samuel L. Perry, "Does Viewing Pornography Reduce Marital Quality Over Time? Evidence from Longitudinal Data," *Archive of Sexual Behavior*, doi: 10.1007/s10508-016-0770-y.

47. See studies from Malamuth et al. (1988a, 1988b, and 1984). Also, recent studies of Dutch adolescents—Hald et al. (2013)—find that pornography is one small factor among many that accounts for sexual aggression.

48. David Ley, Micole Prause, and Peter Finn, "The Emperor Has No Clothes: A Review of the 'Pornography Addiction' Model," *Current Sexual Health Reports* 6, no. 2 (2014): 94–105.

49. Many of the studies cited in the preceding endnotes lament this fact. A particularly telling illustration of the problem is the report on the effects of pornography exposure on children compiled by Miranda A. H. Horvath, Lilian Alys, Kristina Massey, Afroditi Pina, Mia Scally, and Joanna R. Adler, "Basically ... Porn is Everywhere: A Rapid Evidence Assessment on the Effects that Access and Exposure to Pornography has on Children and Young People," London, Middlesex University, Office of the Children's Commissioner (2013), http://www.mdx.ac.uk/__data/assets/pdf_file/0026/48545/BasicallyporniseverywhereReport.pdf. In this study, they found more than 40,000 results in their searches of the academic literature on the causal effects of pornography exposure, but only a few hundred studies met their inclusion criteria, and even then they noted the partial, low quality, and sometimes biased nature of the research.

50. Danaher, "Robotic Rape,".

51. I previously published some of the following paragraphs in Danaher (2016).

52. Ibo Van de Poel, "An Ethical Framework for Evaluating New Technologies," *Science and Engineering Ethics* 22, no. 3 (2016): 667–86; Zoe Robaey and Arno Simons, "Responsible Management of Social Experiments: Challenges for Policymaking," in *Responsible Innovation*, 2nd ed., eds. Bert-Jaap Koops, Ilse Oosterlaken, Henny Romijn, Tsjalling Swierstra, and Jeroen van den Hoven (Berlin: Springer International Publishing, 2015): 87–103.

53. David Collingridge, *The Social Control of Technology* (London: Frances Pinter, 1980).