It Loves Me, It Loves Me Not:
Is It Morally Problematic to Design Sex Robots that Appear to Love Their Owners?

Sven Nyholm and Lily Eva Frank

Abstract: Drawing on insights from robotics, psychology, and human-computer interaction, developers of sex robots are currently aiming to create emotional bonds of attachment and even love between human users and their products. This is done by creating robots that can exhibit a range of facial expressions, that are made with human-like artificial skin, and that possess a rich vocabulary with many conversational possibilities. In light of the human tendency to anthropomorphize artefacts, we can expect that designers will have some success and that this will lead to the attribution of mental states to the robot that the robot does not actually have, as well as the inducement of significant emotional responses in the user. This raises the question of whether it might be ethically problematic to try to develop robots that appear to love their users. We discuss three possible ethical concerns about this aim: first, that designers may be taking advantage of users' emotional vulnerability; second, that users may be deceived; and, third, that relationships with robots may block off the possibility of more meaningful relationships with other humans. We argue that developers should attend to the ethical constraints suggested by these concerns in their development of increasingly humanoid sex robots. We discuss two different ways in which they might do so.

Key words: love, sex robot, deception, social robotics, emotional attachment

The company "Truecompanion.com" claims to sell "the world's first sex robot." The robot is called "Roxxxy," and the website advertising the robot makes some striking claims about the capabilities of this robot. Supposedly, Roxxxy

Sven Nyholm, Department of Philosophy and Religious Studies, Utrecht University, Janskerkhof 13, 3512 BL Utrecht, The Netherlands; s.r.nyholm@uu.nl.

Lily Eva Frank, Philosophy & Ethics, School of Innovation Sciences, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands; l.e.frank@tue.nl.

> knows your name, your likes and dislikes, can carry on a discussion and
> express her love to you and be a loving friend. She can talk to you, listen to
> you and feel your touch. She can even have an orgasm![1]

In an interview about Roxxxy, the robot's inventor Douglas Hines remarked that
a relationship with Roxxxy is about more than just sex. Sex, he stated, "is only a
small part of it."[2] Roxxxy can be a "true companion."

The company developing the sex robot "Samantha" makes similar claims
about their invention and has similar ambitions for their robot. In a 2017 television
appearance on the UK show "This Morning," inventor Arran Lee Wright described
the extent of Samantha's artificial intelligence. He claims that it enables Samantha
to respond to different social scenarios: she can talk to you ("she can talk about
animals, she can talk about philosophy") and even tell jokes. Much like Roxxxy,
Samantha is described as a true companion; as Wright explains, "you can tell her
'I love you' and she can respond."[3]

A third example is "Harmony," a sex robot developed by Matt McMullen.
Speaking about this robot and its capabilities, McMullen states in a video that "we
wanted to create a very open-ended kind of discourse that you can have with this
AI." That, he adds, "is what makes it fun." The head of the robot has "twelve points
of articulation in the face," enabling it to display a range of different facial expres-
sions. All these things "work together to make it come across as if she's alive." The
aim is to make people "feel a connection" when they are speaking with the robot.[4]

These three sex robots, then, are developed to be "more than a sex partner."
They are designed to make people "feel a connection," to make people want to say
"I love you" to the robot, so that users can see their robot as a "true companion."
In this article, we focus on this aim that some developers of sex robots have. We
discuss the ethics of pursuing the aim of creating a sex robot that will make users
respond emotionally, with attachment, affection, and even love. Our main ques-
tion is whether—and if so why—it might be ethically problematic to develop sex
robots designed to make their users feel an emotional connection and attachment
to the robots of the sort we typically associate with a loving human romantic rela-
tionship. Is this ethically problematic? And if so, for what reasons?

Our aim is not to argue that this is ethically problematic in all possible cir-
cumstances. Rather, our aim is to describe three conditions or circumstances under
which this might be considered ethically problematic. We will then discuss what,
if anything, could be done to counteract these ethical worries. In short, our three
areas of discussion are: firstly, conditions under which an individual's emotional

vulnerabilities are taken advantage of for the sake of someone else's gain, financial or otherwise; secondly, conditions under which users are deceived about what the robot's actual capabilities are; and thirdly, conditions under which the presence of the robot blocks off the option of having a more valuable and meaningful type of relationship with other human beings. Our suggestion in this article is that if one or more of these conditions obtain, the aim of creating a sex robot that people will get emotionally attached to becomes ethically problematic.[5] Accordingly, the ethically conscientious sex robot developer would need to seek ways of designing safeguards that forestall the mentioned ethical worries. We will offer some suggestions.

      We begin by situating this article in the growing literature on the ethics of sex robots. We then discuss the three ethical considerations just mentioned, one by one. We end by making two suggestions about what developers of emotionally engaging sex robots might do to offset the worries we describe below. Our aim, to restate, is not to argue that creating a sex robot that stimulates emotions and attachments will always and necessarily be ethically problematic. In keeping with our own previous contributions to the ethics of sex robots—as well as in keeping with contributions to this discussion by various other authors—we proceed on the assumption that the ethical discussion about sex robots should also explore ways of avoiding ethical problems sex robots might give rise to, as opposed to only identifying potential ethical problems (Nyholm and Frank 2017; Frank and Nyholm 2017; McArthur 2017).

1. Background

1.1. What Is a Sex Robot?

We focus particularly on sex robots in this article, rather than social robots more generally, for three reasons. First, this gives the article a sharper focus. Second, as we will see below, there are some general points discussed in general terms in ethical debates about social robots that have not yet been discussed more specifically in relation to sex robots in particular. And third, the types of ethical concerns we discuss in this article appear particularly poignant in relation to sex robots designed to cause their users to become attached to them. The idea of robots designed for sex, but also designed to create an emotional attachment in their user by appearing to like or love the user, is a very striking one.[6]

      By a sex robot, we here understand a robot that is designed to look like a human, with a certain degree of artificial intelligence, designed specifically for

sexual purposes (Nyholm and Frank 2017). We understand such a robot as being able to move in ways made to facilitate sexual interactions with the robot. And we understand the robots in question to be designed to appear able to talk and in other ways communicate with their users. For example, Samantha—the second above-mentioned robot—has the look and is the size of a small adult woman. Moreover, Samantha has skin designed to feel like human skin, and is also equipped with various sensors spread out across her body that enable the robot to respond to different forms of touch and interaction. As was also mentioned above, Samantha can respond to human speech. Similarly, the robot Harmony is compared to the smart phone feature "Siri" in the above-mentioned video featuring Harmony's developer Matt McMullen.[7]

In the future, it might be that robots will become advanced enough in their artificial intelligence that it might make sense to attribute to them emotions and other types of internal states that will make them similar to human beings in their capacities and their agency. We do not wish to exclude that possibility here. We think that if one accepts a broadly functionalist understanding of human mental states, there is no principled reason why future robots might not be able to have a type fairly sophisticated minds and emotional lives (Frank and Nyholm 2017). However, at present—and certainly for the sake of this discussion—we are taking it that robots like the three instances described above lack such capabilities. Currently, they are not able to have emotions or any inner life resembling that of a human being.

Rather, the types of robots we will focus on in what follows are artifacts that are designed to outwardly appear human, mimicking human behaviors, and giving the impression that they possess emotions, internal states, etc. Thus, when a robot like Roxxxy is described as having an "orgasm," we understand such a robot as behaving in a way that makes it look, sound, or feel, like a human having an orgasm. Hence when we ask whether the aim of creating sex robots designed to create emotional attachments and feelings of commitment in human users, we are asking this question about sex robots that could not be understood as being able to reciprocate these types of emotions and attachments.

1.2. Love between Humans and Robots?

In a previous article, we investigated the possibility of mutual love between humans and robots (Nyholm and Frank 2017). We described three clusters of ideas commonly associated with romantic relationships in western cultural contexts, important presuppositions about mental capacities associated with these ideas, and

the question of whether it is conceivable that robots one day might be able to live up to these criteria. Specifically, we investigated the ideas of (i) love as involving being a "good match"; (ii) mutual love in the sense that two or more people value their relationship and each other in the distinct particularity as the unique individuals they are; and (iii) mutual love as involving the idea of a commitment to one another. We argued that these common ideas about love presuppose capacities such as that of being able to value a person, of having a will, and of being able to make free choices. We did not conclude that robots could never come to have such capabilities. But we did conclude that creating such robots is currently a great challenge. No presently existing robots would be able to participate in relations that can truly be described as living up to all the mentioned three ideas associated with the ideal of mutual love (cf. Danaher 2018; see also Hauskeller 2017).

What about the human side of the equation? How likely is it that humans would feel a strong attachment to a sex robot prompted by the ways in which the robot responds to them? According to robotics researcher Matthias Scheutz, in an article on "the inherent dangers of unidirectional emotional bonds between humans and social robots," many humans are prone to develop emotional bonds with robots of all sorts—from military robots to robot pets and even to robotic vacuum cleaners (Scheutz 2012). Scheutz thinks, as the title of his article implies, that this is an "inherent danger" related to all sorts of social robots. A social robot is here defined as any robot designed to interact with human beings; among the types of social robots Scheutz mentions in the article are sex robots designed to have sexual interactions with humans. According to Scheutz's assessment and the research he describes, then, it is highly likely that many human users would indeed develop strong emotional bonds with sex robots designed to stimulate such responses.[8]

That some humans are likely to develop a strong emotional bond with sex robots with which they interact is supported by extensive evidence about how easily humans anthropomorphize technology. Humans are even prone to develop sentimental attachments to robots not specifically designed to create such feelings (Turkle 2011; Breazeal 2002). This suggests that the same patterns will emerge in the case of technologies specifically designed to elicit these feelings.

Even very simple artefacts that lack any artificial intelligence or any type of functional autonomy can become the objects of strong emotional attachments. This is suggested by cases such as those of people who view themselves as having loving relationships with sex dolls. For example, a man living in Michigan who calls himself "Davecat" has received a lot of media attention for his claim that he has been "married" for over 15 years to a sex doll named Sidore (Beck

2013). Davecat describes himself as a proponent of "synthetic love." He has made several media appearances in which he describes his love life, e.g., in the BBC documentary *Guys and Dolls*. Similarly, in Japan, there is a trend whereby men have "virtual girlfriends." The Japanese government has even had worries about an apparent decrease in interest among young people in finding stable romantic partners with whom to form families (Rani 2013; Jozuka 2016).

If people like Davecat and the mentioned group of Japanese males form what they experience as romantic bonds with sex dolls or avatars ("virtual girlfriends"), it is quite likely that the same would happen with sex robots people might interact with. That is especially so if findings from robotics, psychology, and human-technology interaction research are taken into account by developers with the specific aim to create robots that induce emotional bonding. In the videos referred to above, the creators of Roxxxy, Samantha, and Harmony explicitly state that they are drawing on such research results. Accordingly, we take the type of scenario that we wish to discuss here—i.e., sex robots that fulfill their design goal, namely, to make people feel that they are emotionally bonding with the robot—is a likely one. It is a scenario that needs to be taken seriously from an ethical point of view.

1.3. The Ethics of Sex Robots thus Far

The ethics of sex robots is a quickly growing field. As mentioned above, one of our own contributions to this field is about whether there could ever be mutual love between a human and a sex robot (Nyholm and Frank 2017). Another is about whether it would be conceivable, possible, and desirable to create a sex robot able to consent—or not consent—to sex with a human (Frank and Nyholm 2017). Both of those contributions of ours were responses to the main topic that has so far dominated the ethics of sex robots, namely, a powerful feminist critique according to which the development of sex robots might lead to yet another way in which sex partners might be objectified and thereby not receive the respect and concern due to all potential sex partners.

The worry is partly that the very idea of creating sex robots will symbolize an objectifying and undignified image of what a sex partner—and particularly a female sex partner—is and should be. The worry is also partly that if people have such objectifying and morally problematic attitudes towards sex robots, those attitudes might then transfer over to how potential human sex partners are viewed and treated. Kathleen Richardson—and with her, the campaign to ban sex robots—is a driving force behind this feminist critique of sex robots (Richardson 2016). Our above-mentioned papers aimed to investigate conditions under which

relations between humans and sex robots could avoid this type of objectification, by instead setting a good example for how humans and their sex partners should relate to each other.

Other interventions in the sex robot ethics debate have featured other types of possible responses to the type of criticisms offered by Richardson. For example, John Danaher has investigated ways in which human interactions with sex robots might come to symbolize something positive rather than something negative (Danaher 2017). Together with Brian Earp and Anders Sandberg, Danaher has also offered a sustained critique of the arguments offered by the above-mentioned campaign against sex robots, arguing that the reasons for why we ought to campaign against sex robots are unconvincing (Danaher, Earp, and Sandberg 2017). Neil McArthur, in turn, has presented a positive case in favor of sex robots. He argues that sex robots could help to serve the sexual needs of, and offer companionship to, lonely people. Moreover, McArthur also argues that sex robots could serve as a form of transition therapy for people who have had bad sexual experiences. They may want to return back into the sexual community via a type of sexual interaction where they can control things and thereby gradually ease their way back into human sexual interactions (McArthur 2017); we will return to this idea below.

There has also been recent discussion of other ethical questions related to sex robots—such as whether a natural law approach to sexual ethics would necessarily rule out sex between humans and robots or not (Goldstein 2017). But as far as we know, none of these contributions to the ethics of sex robots have so far tackled the issue investigated in this article, namely, whether and under what conditions it might be ethically problematic to try to create sex robots to which humans would be emotionally attached. We thus view the present discussion as filling a gap in the literature on the ethics of sex robots and as an attempt to move the discussion forward.

There has been some discussion on a more general level about whether social robots to which people might form emotional attachments might be ethically problematic. As we already mentioned above, Scheutz thinks that this general prospect—i.e., the prospect of emotional bonds between humans and social robots—is a "dangerous" one. Additionally, in a set of "principles of robotics" suggested by Margaret Boden and colleagues, one of the suggested principles they set out bears directly on the topic we wish to discuss here. The fourth of the proposed principles reads as follows:

Robots are manufactured artefacts. They should not be designed in a decep-
tive way to exploit vulnerable users; instead their machine nature should be
transparent. (Boden et al. 2017)

We see this suggestion as insightful and as highly relevant to this topic. How-
ever, we also think that this rule conflates different ethical considerations that are
strictly speaking better treated as separate ethically relevant concerns—namely,
on the one hand, the issue of exploiting vulnerable people, and, on the other hand,
the issue of deceiving people. These small reservations aside, we do think that the
strictures against exploiting vulnerable people and against deceiving people are
worth taking very seriously. They have not yet received attention within the do-
main of sex robots that we are focusing on here. For that reason, these two ethical
considerations constitute the first two of the three ethical considerations related to
our topic that we will now turn to.

2. Ethically Sensitive Conditions for Sex Robots that Create Emotional Bonds

2.1. Exploiting emotional vulnerabilities for personal, i.e., financial gain

Let us return to Davecat for a moment. One of Davecat's above-mentioned media
appearances was on an American Television show called "my strange addiction,"
on which it was suggested that Davecat's attachment to his sex doll wife Sidore
is a form of addiction.[9] An addiction specialist interviewing Davecat on the show
says to Davecat that addiction is usually a matter of trying to fill some sort of void,
or of treating some sort of pain. He then asks Davecat what void he is trying to fill
or pain he is trying to heal. Davecat responds "loneliness." In another clip, Dav-
ecat says that while he does not at all mind being alone, he minds loneliness. His
synthetic wife Sidore helps him to get over his feeling of loneliness, so that he can
be alone without having to feel lonely. Loneliness, which can occur regardless of
whether or not a person has social contacts and relationships, is an "aversive" state
that is persistent, "distressing," connected with "sadness and anxiety" (McWhirter
1990, 418):

[L]oneliness is defined as an enduring condition of emotional distress that
arises when a person feels estranged from, misunderstood, or rejected by
others and/or lacks appropriate social partners for desired activities, partic-
ularly activities that provide a sense of social integration and opportunities
for emotional intimacy. (Rook 1984, 1391)

Now, let us suppose—as a thought experiment—that Davecat's synthetic wife was not a doll, but instead a robot. And let us further suppose that the manufacturer of the robot reasons that to increase their revenues, they not only sell the robot, but also make the customers continually pay for updates or upgrades. Suppose further that they come up with a effective way of "hooking" people on these robots; and that unless one pays for the updates every now and then, the robot will stop functioning properly, thereby not providing the emotional support the user has gotten attached to. This thought experiment illustrates the first type of condition or scenario which, as we wish to suggest, raises ethical worries regarding the aim of creating a sex robot evoking emotional attachment.

The worry here is that vulnerable people, e.g., lonely people, will be systematically targeted—in this scenario their emotional vulnerabilities are exploited or taken advantage of for the sake of the financial gain of the company creating and updating the robot. If the manufacturer of a robot tries to make people "addicted" to the robot—after which this addiction is then used as a steady source of further income for the company—this is, in our view, one type of circumstance under which the aim of creating emotional attachments to a sex robot can be sensibly seen as problematic from an ethical point of view. Users may come to depend on the technology to fulfill some of their most profound emotional needs.

One might ask here whether the idea of addiction, or something sufficiently like it, can be meaningfully used in this general context. The answer depends on whether an attachment of the general sort usually associated with a romantic relationship could ever be thought of as an addiction. Bioethicists Brian Earp, Olga Wudarczyk, Bennett Fodder, and Julian Savulescu certainly think so (Earp et al. 2017). They argue that people can become "addicted to love," and that there may be a need, in the case of some people, to develop medications that can be used to treat these types of emotional addictions (ibid.; see also Earp, Sandberg, and Savulescu 2015).

The claim that such addictive attachments ought to be treated with medications we will set aside here. But we agree with the suggestion from Earp et al. (2017) that it can sometimes make sense to view some emotional attachments which the subject experiences as romantic attachments as being addictive or addiction-like in nature. At the very least, a person can become so emotionally attached to the object of their affection that they behave in what others regard as unreasonable ways in order to hold on to what they are so strongly attached to. We think that this could happen to lonely people who form a very strong emotional attachment to a sex robot programmed to create an emotional attachment in users.

Such lonely people might turn to sex robots as an act of desperation, feeling that they have no other choice. Accordingly, in order to make the attempted creation of socially interactive sex robots that people will form emotional bonds with ethically acceptable, the designers of such robots need to find a way of avoiding the charge that their products are poised to exploit lonely and desperate people.[10]

2.2. Deception

The second potential ethical worry we wish to discuss also derives from Boden et al.'s (2017) fourth rule. It concerns the potential for deception. As we saw above, Boden et al. claim that artefacts should not be designed in a deceptive way to obscure their machine nature. However, in the case of sex robots this clearly presents a puzzle. To the extent that the designers make the machine nature of the robot explicit or obvious, this presumably diminishes the appeal of the robot as an entity someone might want to spend time with, have sex with, and even express affection towards. As noted above, designers of sex dolls and robots are continually researching and advertising improvements in things like skin texture and temperature, speech quality, etc. These are features are specifically intended to simulate the human body and human communication.

Notably, a parallel objection involving deception arises in the literature surrounding the use of care and companionship robots for the elderly or home bound (Sparrow and Sparrow 2006; cf. Coeckelbergh 2010, 2012).[11] Such arguments tend to take a broadly Kantian form. Specifically, Robert Sparrow and Linda Sparrow have argued that the use of robotic caregivers may be problematic from the perspective of basic Kantian respect for persons as ends in themselves because of the potential deception involved. The behavior and outward features of the robots are likely to evoke in care-receivers false beliefs about what the robots feel or think about them, i.e., that they do indeed value their well-being etc. (Sparrow and Sparrow 2006). Beyond the respect-focused objection, Sparrow and Sparrow also argue that this potential deception is a problem from the perspective of people actually getting what they need from the robots:

> [S]uch deception is a bad thing because our preferences are unlikely to be met, our interests advanced, or our wellbeing served, by illusions. What most of us want out of life is to be loved and cared for, and to have friends and companions, not merely to believe that we are loved and cared for, and to believe that we have friends and companions, when in fact these beliefs are false. (Sparrow and Sparrow 2006, 155)

Presumably, the idea here is that this strong desire to be loved and cared for—and to have friends and companions—is also a reason why people might come to falsely believe that robot companions truly do love and care for them. A strong desire for some outcome can easily motivate us to disregard evidence against the likelihood of this outcome and to instead primarily focus on confirming evidence in favor of the status quo. This is sometimes referred to as confirmation bias.

If we put this common bias together with people's well-documented tendency to attribute human-like properties to anything that looks human or behaves in a human-like way (Ruijten under review), it is not hard to believe that people might often come to believe—either on a conscious level or on a more intuitive and subconscious level—that certain robots have much more of an inner life and many more capabilities than they really have. In other words, the combination of (i) a strong motivation to believe that we are being cared for (prompted by our strong desire to be cared for) and (ii) our tendency to anthropomorphize may likely dispose many of us to be deceived by social robots.

Designers and developers of love robots are likely to be well versed in the psychological literature and affective computing research, and will rely on these findings to create robots that users will find desirable, with which they will emotionally bond, while forming the belief that the robot reciprocates at least some of these feelings. It is reasonable to assume that developers realize that users will form false beliefs of the sort just discussed above. Whether this explicitly constitutes lying to users, merely misleading them, or failing to correct their false beliefs, is up for debate. As Jennifer Saul has argued, these are morally distinct categories, and may entail different strengths of duties associated with them (Saul 2012). But taking a broadly Kantian perspective on the wrongness of lying, all three categories of deception could be thought to fail to treat people as ends in themselves (cf. Sparrow and Sparrow 2006). In Christine Korsgaard's terms, what is required in this type of a communicative transaction in order for it to show that distinctive type of respect, is for it to ensure that all moral subjects have "knowledge of what is going on" and some "power over the proceedings" (Korsgaard 1996, 139). If one knowingly creates false beliefs in somebody that one knows they will act on, then one seemingly fails to do these things. Creating a robot that will induce in users' false beliefs about the robot's capacities is potentially to offend against this principle. This Kantian analysis is one common way of understanding worries about deception, autonomy, and respect. Although we don't wish to commit to this specific analysis here, we do expect that other moral frameworks will generate similar concerns.

2.3. A Way of Blocking Off Other Valuable Relationships?

We turn now to the last types of conditions under which we think that there is a moral objection towards creating sex robots with aim of making users become strongly emotionally attached to the robots. To introduce this last type of conditions, we will first consider another type of social robot, used for a different type of purpose: namely, Kaspar, a robot used to stimulate social interaction in autistic children.

The aim of this type of robot is to "open up" autistic children to social interaction with other human beings (Robins et al. 2004). As described by the researchers developing Kaspar, the idea is that while autistic children are very shy and non-communicative around other human beings—for example, by not wanting to look them in the eye or talk with them—the children may be more open to interacting with a robot that looks like a human. The hope is that by first interacting with this humanoid robot (Kaspar looks like a small child with very simple features), the children can then "switch over," so to speak, to interacting directly and perhaps even bonding with other human beings, such as their teachers at school or other children. Page 413: little bit below the middle of the page: there is a sentence that goes like this: "This robot could then be a means of opening up social, valuable relationships with other human beings. Initial findings with Kaspar the robot are promising. There appears to be some evidence that Kaspar does indeed help to open up some autistic children to social interactions with other humans, for example when the children look at the person operating the robot and point to Kaspar as a way of trying to establish Kaspar as object of joint attention with the robot operator (Wood et al. 2017; Robins 2018).

But imagine that upon starting to interact with these robots, the autistic children were to form a strong preference for interacting with Kaspar and other robots. Imagine that they would find this so much more comfortable than interacting with human beings (e.g., the teacher at school or other children) that they would become less open to human interactions than before. The introduction of this robot to these children would then block off valuable social relations that the children might otherwise over time have been able to develop with those other people. Presumably, that would be a very bad outcome. Indeed, at a recent presentation of his research with Kaspar and autistic children, robot developer Ben Robins explicitly pointed out that a key ethical aim he and his team have is precisely that Kaspar should be a path to human-to-human social interactions, and not a substitute for them (Robins 2018).

We would like to suggest that a similar concern is worth taking seriously in the case of sex robots created to make their users emotionally attached to them. If these robots were successful enough in making their users bond with them, and if this would come at the cost of blocking off valuable human-human relationships that the users could otherwise have, this could sensibly be viewed as a further circumstance that could make the aim of creating a sex robot that people will want to have as a companion ethically questionable (cf. Turkle 2011 on the "robotic moment").[12]

This worry is premised on the assumption that a romantic and intimate relationship with another human being—at least at this stage of the development of humanoid robots—would be a much richer and valuable type of relationship than a scenario in which a person interacts with a sex robot as if that sex robot were his or her romantic and intimate partner. If a person becomes so attached to such a sex robot that he or she will feel that they need to be "true" to that robot—so that it would be "cheating" on the robot to, say, go on dates with other humans—the introduction of this robot would in effect block off what would presumably be a more valuable and more multi-dimensional relationship that the person could otherwise potentially develop with other humans. At least at this current stage, a sex and companionship robot to which a person gets so emotionally attached that the person would hesitate to seek human romantic partners could sensibly be thought to impoverish that person's life by serving as a poor substitute for what would be a fuller and more multidimensional type of relationship (cf. Danaher 2018).

Is this a realistic enough scenario that it is worth worrying about? In other words, even if the prospect of sex robots that would make people prefer them to other humans would be a bad thing in theory, is this something that we should take seriously in practice? This is hard to say. As ethical theorists, we are not ideally placed to make predictions about the exact likelihood of such a scenario. However, we do want to point out that, as we mentioned above, there have been worries about something similar to this in the technologically very advanced society of Japan about falling birthrates caused by a growing preference among some men for the above-mentioned "virtual girlfriends" rather than human girlfriends. For example, a 2010 survey by the Japanese Ministry of Health, Labor, and Welfare among young men noted a surprisingly low interest in sex among these young men. The hypothesized explanation of this was that many of these young men had come to prefer virtual girlfriends to the more complicated demands of interaction with human sexual and romantic partners (Rani 2013).

Our claim here is, as we might put it, hypothetical in nature: suppose that it were the case that sex robots had been successfully designed to create a strong emotional bond in users such that they interpret it as a romantic bond. And suppose that this goes on to create in these users a dominant preference for these robots over human partners. In that case, this would be problematic from an ethical point of view. It would be problematic and regrettable so long as the robots are unsophisticated enough that a "relationship" with them (rather than with a human partner) is lacking many of the rich and valuable dimensions and aspects that can be realized within a human romantic relationship. The robot, in effect, would under such a circumstance block off a more valuable type of relationship.

Furthermore, the human may not only be prevented from enjoying these richer aspects of human relationships. It may be that the interaction with the robot creates the impression that this is all there is and could be to a relationship. In this way, the robot relationship might in some cases not only block off the instantiation of human relationships, but also the very conception thereof. Moreover, human-robot relationships, given current technology, do likely not offer the opportunities for personal growth and virtue development that we associate with human relationships. As Mary Wollstonecraft argues, whereas romantic relationships between equals breed virtue, relationships between unequals encourage in the person of higher status vices, such as self-centeredness and a tyrannical and domineering personality (Wollstonecraft 2009). Presumably, a putative relationship between a human and a current sex robot would be highly unequal in nature—even if we can perhaps imagine more equal relationships in the distant future.

3. What Could Developers of Sex Robots Do to Respond to These Problems?

Having discussed the three ethical worries about the aim of creating sex robots users might bond emotionally with, we now turn to the question of whether there is anything designers of such robots could do in order to make their products less vulnerable to these ethical concerns. We will put forward two suggestions. The first suggestion relates both to the first and the third of the three ethical worries described above. The second suggestion relates specifically to the second ethical worry we have described. Before we turn to these suggestions, we should point out that in this section we are particularly interested in what designers of sex robots could do to avoid the ethical worries we have described. Whether our suggestions would also be attractive from a marketing or business model point of view is something that we set aside here.

Let us return first to the issue of exploiting desperate and lonely people, who may feel that they have no other choice, and who for that reason might be driven into the arms of sex robots appearing to love them. How might an ethically conscientious designer of an emotionally interactive sex robot possibly deal with this worry, supposing that they do not simply ignore this concern? We suggest that the first thing to do is to think about how one can avoid exploiting desperate people more generally. One crucial thing to note is that what makes a person desperate and, therefore, vulnerable to exploitation is often that they lack—or feel that they lack—other options that are more attractive than what is on offer from the party accused of exploiting them (Zwolinksi and Wertheimer 2016). For example, if the only viable employment option on offer is to work in a sweatshop, then a desperate poor person might go for this option, since they have no sensible alternative to this.

Of course, the first thing that a sweatshop owner would need to do in order to avoid the charge of exploitation is to make the working conditions in their sweatshop more worker-friendly. Suppose that they do this, but that they still do not offer any salary that can be seen as fair and reasonable. The valid charge of exploitation might then still remain in place. The next thing the owner of the sweatshop might do—and here we get to the general idea that is most relevant in this context—is the following. The sweatshop owner could offer some sort of training program that would give the poor worker skills that would help him or her to come to have other options, and therefore not be without any other alternatives. Since the worker could seek more gainful employment elsewhere, now that he or she has these new skills, it can be argued that if the worker continues to work in the sweatshop with the low salary, the charge of exploitation seems to potentially have been undermined, at least to some extent.

Let us now try to carry this over to the context of sex robots used by desperate people, who may feel that they have no better options than robots designed to appear to love them. Could designers of such robots do anything comparable to our imagined sweatshop owner above, in order to avoid being open to the charge of exploitation? We think that they could, and this could also help to ease the third type of worry described above, namely, the worry that these robots might deskill users to a point where the option of more valuable relationships are being blocked off.

Our suggestion is partly inspired by McArthur's (2017) above-mentioned idea that sex robots could serve as a sort of therapy or transition tool for people wishing to return to the sexual community.[13] Suppose that the socially interactive sex robots could serve to train users to become better lovers and more sensitive partners. They could do so by offering users tips and tricks in the sexual domain,

by teaching users about human anatomy, and by making recommendations about how to interact with a partner. This could help to train the users in the intimate and romantic domain. Users could thereby become more eligible as sexual or romantic partners, and this could also significantly boost their confidence in this domain. As a result, the users of these sex robots could start feeling that they could be more successful when it comes to finding human partners, if they would wish to seek human partners rather than robotic partners. If robots could indeed train users to become better lovers and more sensitive partners, this would also make them more attractive as sexual and/or romantic partners for other humans.[14] Accordingly, the users could become less desperate in the sense of not having any other options to turn to other than these sex robots.

If designers made their socially interactive sex robots able to train users in these ways, so that users would come to also have—and feel that they have—other options, this would help to ease ethical worries that they are exploiting lonely and desperate people who feel that they have no other options. Moreover, it could also help to ease the third worry described above, namely, that the sex robots might block off the possibility of having more valuable relationships with human beings instead of with these robots. Having been trained by these robots to become better lovers and more sensitive partners, the users would not only have reason to feel that they have other options. They could also more easily move on to relationships with human partners.

What about the deception worry? Like Boden et al. (2017) suggest above, the most obvious solution is to take measures to make the robotic nature of the robots transparent, so that users are not mislead into thinking that the robots have more advanced capacities than they actually have. Think, for example, of the back of the head of the famous robot Sophia from Hanson robotics (Sharkey 2017). While Sophia's face looks very much like a human face, the back of Sophia's head is literally transparent and highly robotic in its appearance, so that it becomes obvious that Sophia is a robot.[15] Of course, when it comes to robots designed to interact with users in an emotionally engaging way, this is easier said than done. The kinds of cues and other means that would help to engage users' emotions—e.g., facial expressions and things the robots would "say"—would presumably work best if they are designed to be human-like rather than maximally robot-like.

Our suggestion here is inspired by two different things: procedures for informed consent within medical contexts and warning labels on products like cigarettes. We suggest that well-informed consent to interaction with a sex robot designed to appear to respond emotionally to their users could be achieved if the

robot every now and then would play a message reminding the user that robots function in different ways than humans do. There should not just be a text with a lot of complicated legal language—like when one downloads an app on a smart phone—that the user is asked to look at and consent to before they can start inter-acting with the robot (cf. Frischmann and Selinger 2018). Rather, it seems like a better idea to us to have the robot every now and then engage the user in discussion about differences between robots and humans.

The sex robot could say things such as "remember that I am a robot and that humans and robots differ in how they function" or "recall that as a robot, I operate in a very different way than a human does." And the robot could be made to say such things every now and then, not just the first time the user interacts with the robot. The socially interactive sex robot could be made to say such things often enough that the user would not forget it. In addition, the user could sometimes be asked to reaffirm their consent to using a robot that is programmed to elicit social responses. The robot could say, for example, "now that I have reminded you that I am a robot and that humans and robots function in very different ways, do you still consent to the interaction that you have with me?" or messages like that. Such verbal reminders and requests for ongoing consent could help—we suggest—to counteract the worry that sex robots designed to engage with their users emotion-ally would be overly deceptive.

4. Concluding Remarks

Sex robots designed to interact with their users in an emotionally engaging way can seem like a piece of science fiction, but they are part of a broader trend. They are part of a trend whereby robots are increasingly introduced into our lives, taking over many tasks previously performed by humans, and playing roles in our lives we could not easily have predicted. Many of the ethical concerns that arise in rela-tion to other social robots—such as care robots—will also be worth considering in relation to sex robots. Some of these worries, such as the worry that social robots might be deceptive, can even appear to cast doubt on the whole project of creating socially interactive robots. But it is important to proceed in a nuanced way. Just like we should not reject the whole idea of social robots because there are ethical concerns that can validly be raised about them, we should not be too quick to reject the whole idea of sex robots because there are ethical concerns of similar kinds that can be raised about them.

At the present, it is easier to come up with ethical worries about sex robots than it is to think of ways of allaying those worries. However, we are of the opinion that

rather than trying to ban all forms of sex robots, it is a better idea of investigate what types of sex robots, with what types of features, would be a more ethically acceptable alternative for designers of such robots to try to develop (cf. Richardson 2016; Danaher, Earp, and Sandberg 2017). That is why we have here not only discussed ethical objections to the aim of creating sex robots that people will engage with emotionally, but also suggestions about how to make such robots more ethically acceptable.

Notes

We wish to thank the journal's anonymous reviewers and the guest editors of this special issue for their helpful feedback.

1. truecompanion.com, accessed March 3, 2018.
2. https://www.youtube.com/watch?v=2MeQcI77dTQ, accessed March 30, 2018.
3. https://www.youtube.com/watch?v=AqokkXoa7uE, accessed March 30, 2018.
4. https://www.youtube.com/watch?v=0CNLEfmx6Rk, accessed March 30, 2018.
5. An anonymous reviewer asked whether our reasoning concerning sex-robot designers implies, by analogy, that sex workers who provide the "girlfriend experience" should be morally censured with reference to the arguments we discuss here. (What is meant by the "girlfriend experience" here are sex workers who not only sell sex, but also perform the role of a romantic partner.) We think that the ethics of human sex work and the ethics of sex robots created and sold by technology companies should be kept apart. In the case of human sex workers, there is typically a power imbalance that makes the sex workers very vulnerable in relation to their clients or to pimps. This dimension has an important impact on the moral status of what they do, making it problematic to directly carry over the sorts of ethical concerns that we discuss in relation to developers of sex robots to the case of human sex workers. Developers of sex robots are not exposed and vulnerable in the way that human sex workers are. For this reason the ethical situation of developers of sex robots is relevantly different from that of human sex workers.
6. Ethically speaking, the prospect of sex robots designed to appear to love their users is a highly ambiguous idea. To some people, this idea might seem shocking. To others, it might rather seem tragic, or perhaps tragicomic. Since this prospect is likely to awaken strong reactions in many people, we think that it is a good idea to try to carefully articulate and then critically assess of the ethics of this technological development.
7. https://www.youtube.com/watch?v=0CNLEfmx6Rk, accessed March 30, 2018.
8. Although calling this an "inherent danger" in his article, Scheutz does not develop a sustained argument for what the danger or ethical problem is. Instead he assumes that it is obviously risky for people start becoming emotionally attached to the robots with which they interact.

9. https://www.youtube.com/watch?v=kCjyILOOwUg, accessed March 30, 2018.

10. An objector might ask: is this any different from the myriad other technologies (or even more broadly consumer goods) whose sales are profitable only because they fulfill some very basic human social or emotional needs in a way such that users or consumers become dependent on them? It is probably not different to the extent that many products are intentionally designed to exploit users' vulnerability and continually profit off of it. This might make the creation of many consumer goods morally problematic. At the same time, creating a dependence on a technology to cope with loneliness and a person's deep psychological need for emotionally meaningful relationships seems especially pernicious. It not only preys on this need, but may offer a poor substitute. We should also not be too quick to compare relationship goods like love and companionship with mere consumer goods. After all, love and companionship are typically not seen as mere consumer goods. Romantic partners are valued as unique individuals and as ends, whereas most consumer goods are viewed as replaceable and as means to other ends (Nyholm 2015). Moreover, close personal relationships are deeply tied to health and well-being in a way that most consumer goods are not (Savulescu and Sandberg 2008).

11. Mark Coeckelbergh has argued that to successfully create empathetic bonds between humans and robots, the robots must be able to engage in what he calls "vulnerability mirroring," that is, the robot's capacity to appear to share certain properties with humans and other animals, like mortality, emotions, capacity for suffering and joy (Coeckelbergh 2010, 8).

12. Other factors that block people from meaningful romantic relationships, especially the larger social context, are arguably currently much more significant than the rise of sex robots. As one of the journal's anonymous reviewers pointed out to us, working conditions, rising costs of living, stress, and lack of leisure time may all contribute to blocking people from pursuing romantic relationships. Many of those factors are connected to the increasing role of technology in our lives. For instance, the fact that nearly 35 % of Korean couples do not have sex, or that there is a rise of extreme reclusion ("hikkomori") in Japan, has been linked to intensive use of the internet and related technologies. (Jozuko 2016) That there are problems like these in our lives more generally should not lead us to simply shrug our shoulders in relation to developing technologies like sex robots. Since engineers are consciously designing sex robots to play a certain role in users' lives, we here have an opportunity to think through the possible repercussions before their use becomes widespread, asking how they will interact with wide social, cultural, and economic forces.

13. Our suggestion is also partly inspired by an anecdote told by the Dutch playwright Simon(e) van Saarloos. In a play she had written, the main male actor had to interact with a very fragile doll that was his romantic partner in the play. The real-world romantic partner of this actor told Van Saarloos that the actor's interaction with

the fragile doll during the months of rehearsing and then during the performances of the play had actually made him into a better and more sensitive lover. This anecdote was recounted by Van Saarloos during a performance of *Robot Love, A Performative Symposium*, on September 15, 2018 in Eindhoven, the Netherlands. (An audio recording of this performance is available here at https://www.youtube.com/watch?v=OZREFoSB2Fo. Accessed December 7, 2018.)

14.   An anonymous reviewer suggested that educational sex robots could also potentially be used to train human beings to respect consent. We find this to be an interesting suggestion. We discuss robot sex and consent in some detail elsewhere (Frank and Nyholm 2017).

15.   Incidentally, Sophia has recently been accused of being deceptive, by Noel Sharkey and others who argue that Sophia is misleadingly presented as having much more sophisticated AI than the robot actually has. (Sharkey 2017) As this helps to illustrate, a robot's being transparent in one way (e.g., the back of Sophia's head) does not guarantee that the robot is also transparent in other ways (cf. the accusation that Sophia's interaction with people is deceptive).

References

Beck, Julie. 2013. "Married to a Doll: Why One Man Advocates Synthetic Love." *The Atlantic* https://www.theatlantic.com/health/archive/2013/09/married-to-a-doll-why-one-man-advocates-synthetic-love/279361/. Accessed March 30, 2018.

Boden, Margaret, Joanna Bryson, Darwin Caldwell, Kerstin Dautenhahn, Lilian Edwards, Sarah Kember, Paul Newman, Vivienne Parry, Geoff Pegman, Tom Rodden, Tom Soreell, Mick Wallis, Blay Whitby, and Alan Winfield. 2017. "Principles of Robotics: Regulating Robots in the Real World." *Connection Science* 29(2): 124–29. https://doi.org/10.1080/09540091.2016.1271400

Breazeal, Cynthia. 2002. *Designing Sociable Robots*. Cambridge, MA: MIT Press.

Coeckelbergh, Mark. 2010. "Artificial Companions: Empathy and Vulnerability Mirroring in Human-Robot Relations." *Studies in Ethics, Law, and Technology* 4(3): 1–17. https://doi.org/10.2202/1941-6008.1126

Coeckelbergh, Mark. 2012. "Are Emotional Robots Deceptive?" *IEEE Transactions on Affective Computing* 3(4): 388–93. https://doi.org/10.1109/T-AFFC.2011.29

Danaher, John. 2017. "The Symbolic Consequences Argument in the Sex Robot Debate." In *Robot Sex: Social and Ethical Implications*, ed. John Danaher and Neil McArthur, 103–32. Cambridge, MA: The MIT Press.
https://doi.org/10.7551/mitpress/9780262036689.003.0007

Danaher, John. 2018. "Embracing the Robot." *Aeon* https://aeon.co/essays/programmed-to-love-is-a-human-robot-relationship-wrong Accessed March 30, 2018.

Danaher, John, Brian D. Earp, and Anders Sandberg. 2017. "Should We Campaign Against Sex Robots?" In *Robot Sex: Social and Ethical Implications*, ed. John Danaher and Neil McArthur, 47–72. Cambridge, MA: The MIT Press.

Earp, Brian D., Anders Sandberg, and Julian Savulescu. 2015. "The Medicalization of Love." *Cambridge Quarterly of Healthcare Ethics* 24(3): 323–36. https://doi.org/10.1017/S0963180114000206

Earp, Brian D., Olga Wudarczyk, Bennett Fodder, and Julian Savulescu. 2017. "Addicted to Love: What Is Love Addiction and When Should It Be Treated?" *Philosophy of Psychiatry and Psychology* 24(1): 77–92. https://doi.org/10.1353/ppp.2017.0011

Frank, Lily, and Sven Nyholm. 2017. "Robot Sex and Consent: Is Consent to Sex between a Robot and a Human Conceivable, Possible, and Desirable?" *Artificial Intelligence and Law* 25(3): 305–23. https://doi.org/10.1007/s10506-017-9212-y

Frischmann, Brett and Evan Selinger. 2018. *Re-Engineering Humanity*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316544846

Goldstein, Joshua D. 2017. "Was It Good for You Too? The Natural Law Tradition and the Paradoxical Good of Sexbots." In *Robot Sex: Social and Ethical Implications*, eds. John Danaher and Neil McArthur, 173–200. Cambridge, MA: The MIT Press.

Hauskeller, Michael. 2017. "Automatic Sweethearts for Transhumanists." In *Robot Sex: Social and Ethical Implications*, eds. John Danaher and Neil McArthur, 203–219. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/9780262036689.003.0011

Jozuka, Emiko. 2016. "Why Won't 541,000 Young Japanese Leave the House?" *CNN World News*: https://edition.cnn.com/2016/09/11/asia/japanese-millennials-hikikomori-social-recluse/index.html. Accessed December 7, 2018.

Korsgaard, Christine. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511554476

McArthur, Neil. 2017. "The Case for Sex Robots." In *Robot Sex: Social and Ethical Implications*, ed. John Danaher and Neil McArthur, 31–46. Cambridge, MA: The MIT Press.

McWhirter, Benedict T. 1990. "Loneliness: A Review of Current Literature, with Implications for Counseling and Research." *Journal of Counseling & Development* 68(4): 417–22. https://doi.org/10.1002/j.1556-6676.1990.tb02521.x

Nyholm, Sven. 2015. "Love Troubles: Human Attachment and Biomedical Enhancements." *Journal of Applied Philosophy* 32(2): 190–202. https://doi.org/10.1111/japp.12085

Nyholm, Sven, and Lily Eva Frank. 2017. "From Sex Robots to Love Robots: Is Mutual Love with a Robot Possible?" In *Robot Sex: Social and Ethical Implications*, ed. John Danaher and Neil McArthur, 219–45. Cambridge, MA: The MIT Press https://doi.org/10.7551/mitpress/9780262036689.003.0012

Rani, Anita. 2013. "The Japanese Men Who Prefer Virtual Girlfriends to Sex." *BBC News Magazine*, http://www.bbc.com/news/magazine-24614830. Accessed March 22, 2018.

Richardson, Kathleen. 2016. "The Asymmetrical 'Relationship': Parallels between Prostitution and the Development of Sex Robots." *ACM SIGCAS Computers and Society* 45(3): 290–93. https://doi.org/10.1145/2874239.2874281

Robins, Ben. 2018. "Kaspar as a Therapeutic Tool: Encouraging Social Interaction Skills in Children with Autism." Keynote lecture at the *Social Robotics: The Future of Human-Machine Relationships* research meet, Eindhoven University of Technology, January 25, 2018.

Robins, Ben, Kerstin Dautenhahn, Rene te Boekhorst, and Aude Billard. 2004. "Effects of Repeated Exposure to a Humanoid Robot on Children with Autism." In *Designing a More Inclusive World*, ed. Simoen Keates, P. John Clarkson, Pat Langdon and Peter Robinson, 225–36. London: Springer.
https://doi.org/10.1007/978-0-85729-372-5_23

Rook, Karen S. 1984. "Research on Social Support, Loneliness, and Social Isolation: Toward an Integration." *Review of Personality & Social Psychology* 5: 239–64.

Ruijten, Peter. "Perceptions of Human-Likeness in Human-Robot Interaction Research." Unpublished manuscript.

Saul, Jennifer. 2012. *Lying, Misleading, and What is Said: An Exploration in Philosophy of Language and in Ethics*. Oxford: Oxford University Press.
https://doi.org/10.1093/acprof:oso/9780199603688.001.0001

Savulescu, Julian, and Anders Sandberg. 2008. "Neuroenhancement of Love and Marriage: The Chemicals between Us." *Neuroethics* 1(1): 31–44.
https://doi.org/10.1007/s12152-007-9002-4

Scheutz, Matthias. 2012. "The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots." In *Robot Ethics: The Ethical and Social Implications of Robotics*, ed. Keith Abney, Patrick Lin and George A. Kerkey, 205–23. Cambridge, MA: The MIT Press.

Sharkey, Noel. 2017. "Mama Mia, It's Sophia: A Show Robot or Dangerous Platform to Mislead?" *Forbes.com* https://www.forbes.com/sites/noelsharkey/2018/11/17/mama-mia-its-sophia-a-show-robot-or-dangerous-platform-to-mislead/#628612da7ac9. Accessed July 21, 2019.

Sparrow, Robert. and Linda Sparrow. 2006. "In the Hands of Machines? The Future of Aged Care." *Minds and Machines* 16(2): 141–61.
https://doi.org/10.1007/s11023-006-9030-6

Turkle, Sherry. 2011. *Alone Together: Why We Expect More from Technology and Less from Each Other*. New York: Basic Books.

Wood, Luka, Abolfazl Zaraki, Michal L. Walters, Ori Novanda, Ben Robins, and Kersin Dautenhahn. 2017. "The Iterative Development of the Humanoid Robot Kaspar: An Assistive Robot for Children with Autism." In *International Conference on Robotics*, ed. Abderrahmane Kheddar, Eiichi Yoshida, Shuzhi Sam Ge, Kenji Suzuki, John-John Cabibihan, Friederike Eyssel and Hongsheng He, 53–63. Dordrecht: Springer. https://doi.org/10.1007/978-3-319-70022-9_6

Wollstonecraft, Mary. 2009. *A Vindication of the Rights of Woman*. New York: Norton & Company.

Zwolinski, Matt and Alaan Wertheimer. 2016. "Exploitation." In *Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, https://plato.stanford.edu/archives/sum2017/entries/exploitation/. Accessed July 21, 2019.