

The Ethics of Robot Servitude

Stephen Petersen
steve@stephenpetersen.net

Department of Philosophy
Niagara University

Center for Cognitive Science
University of Buffalo
25 October 2006

Outline

- 1 Engineered robot servitude
 - The question
 - My answer
 - Motivation for the view
- 2 The objection
 - Engineered human servitude
 - Kant and EHS
 - Aristotle and EHS
 - Mill and EHS
- 3 Underview
 - Caveats and disclaimers
 - Summary

The question

- Suppose we could build creatures
 - with intelligence comparable to our own
 - who by design want to do tasks that we find unpleasant.
- *May we?*

The question's significance

This question is

- natural,
- engaging,
- neglected,
- and important.

My answer

I say yes, it's permissible to build robot servants.

- Contrary to the literature
- *And* contrary to my expectations

I do *not* mean to defend:

- robot volunteers
- robot *slavery*
- “post-identity” modification

Thesis

Instead, I defend the permissibility of *Engineered Robot Servitude* (ERS).

ERS

The building and employment of non-human persons who desire, by design, to do tasks humans sometimes find unpleasant or inconvenient.

Motivation for the view

An emerging consensus on intelligence:

Intelligence

Roughly, a level of adaptability in attaining goals.

- The hardwired goals could be (almost) anything
- Compare retriever dogs

The objection

If ERS is permissible, then it seems so should be *Engineered Human Servitude* (EHS):

EHS

The engineering and employment of *human* persons who desire, by design, to do tasks (typical) humans sometimes find unpleasant or inconvenient.

Compare the “delta” caste in Aldous Huxley’s *Brave New World*.

Engineered human servitude

- It turns out hard to say *why* EHS is wrong.
- For a variety of plausible explanations, I'll argue that either
 - it fails to explain the wrongness of EHS, or
 - it's disanalogous with ERS.

Kantian ethics in one slide

- Argues for a categorical imperative
- “Act as if the maxim of your action were to become by your will a universal law of nature”
- “So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means.”
- Act *autonomously* (“reason must regard itself as the author of its principles independently of alien influences”)

A dilemma for Kantians

Does the creature in question have potential for autonomy, or not?

- If no, what's the problem?
- If yes, what's the problem?

Our treatment of the non-autonomous

“EHS desensitizes us”

- Servile robots cannot have autonomy, but
- “He who is cruel to [robots] becomes hard also in his dealings with men.”

But . . .

- *Is* it cruel to have a laundrybot launder? Or a retriever fetch?
- More likely to confuse EHS with people than ERS with people
- Does not explain EHS wrongness, and maybe disanalogous with ERS

Autonomous robots

“EHS uses people as mere means”

- But must be used as *mere* means
- Permissible inclinations
- A “poor autonomy”: for later
- Does not explain EHS wrongness

Aristotelian ethics in one slide

- "... just as the good ... for a flautist, a sculptor, and ... in general, for whatever has a function ... seems to depend on its function, the same seems to be true for a human being, if a human being has some function."
- "Virtues ... we acquire just as we acquire crafts ... we become harpists by playing the harp. Similarly, then, we become just by doing just actions, temperate by doing temperate actions, brave by doing brave actions ..."

Aristotelian teleology

“EHS violates teleology”

- Underwrites the “it’s unnatural” intuition
- Design has given humans a determinate well-functioning
- Explains EHS wrongness, but disanalogous with ERS

Aristotelian virtues

“Engineering humans is not virtuous behavior”

- Suspiciously *ad hoc*
- Does not explain EHS wrongness

Mill in one slide

- "... actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness."
- "It is better to be a human being dissatisfied than a pig satisfied, better to be Socrates dissatisfied than a fool satisfied."

The heart of the matter

“EHS substitutes lower pleasures for higher”

- A fool satisfied in place of a Socrates dissatisfied
- The heart of the matter
- Captures Aristotelian and Kantian worries too

Population ethics

But *substituting* is tricky here . . .

- Same-person substitution not at issue
- If different-person, requires “same-number” cases
- Is merely *adding* a “fool satisfied” okay?
- Implies genetic *enhancement* obligatory
- Disanalogous (same-number case) or fails to explain (different-number case)

Partial interests

“EHS violates partial interests”

- Perhaps EHS harms not the creature, but those with partial interest in the creature (e.g. parents)
- Compatible with the *person-affecting restriction*

PAR

One can only compare well-being across two different circumstances by comparing the situations of particular people who are in both.

Partial interests

But . . .

- Robots don't have such parents
- Does not explain what's wrong with *Brave New World* vat cases
- Disanalogous (parents) or fails to explain (vats)

Caveats and disclaimers

- Such robots are of course still worthy of ethical respect
- Property a sticky question
- The line of ethical significance (a hunch)
- Rationalization a danger
- Err on the ethically safe side

Summary

- Permissible engineered robot servitude has *prima facie* plausibility
- Engineered human servitude is not obviously wrong or analogous
- The philosophical crux is in population ethics