

## Reflection on the Moral Anthropomorphism of Robots

*Tsung-Hsing Ho*

Department of Philosophy, National Chung Cheng University

E-mail: tsunghsing@ccu.edu.tw

### Abstract

If robots are to function automatically, without human supervision, as depicted in sci-fi imagination, then we must ensure that robots not commit moral wrongs. According to the behaviourist conception of moral agency, if robots, assessed purely on the basis of behaviour, perform as morally as humans, they can be considered moral agents. This naturally leads to moral anthropomorphism: the position that whatever moral standards apply to humans apply equally to robots. I argue against moral anthropomorphism. In light of P. F. Strawson's insights into interpersonal relationships and reactive attitudes, and drawing on paternalist actions as examples, I argue that robots, being not persons, are unable to participate in interpersonal relationships, and therefore their paternalist actions towards humans ought to be less permissible than humans'.

**Key Words:** robots, artificial intelligence (AI), moral agency, moral anthropomorphism, reactive attitudes