

Hana Kalivodová, 'Manipulation and Moral Responsibility', Bakalářská práce

Posudek oponenta

This bachelor's thesis investigates the ethical problem of 'manipulation' as discussed in analytical philosophy since the pioneering work of Harry Frankfurt beginning in his 1969 article 'Alternate Possibilities and Moral Responsibility'. Frankfurt sought to defend a conception of moral responsibility that would survive the acknowledgement of metaphysical determinism. He did this by attacking the very assumption that the 'principle of alternate possibilities' is the key to moral responsibility. Frankfurt held that what is really at issue is not freedom, or the potential for performing an alternative action, but the agent's responsiveness to reasons and their ability to reflectively endorse their desires with 'higher order desires' that constitute their real self (discussed here pp. 13-15). The extensive discussion of moral responsibility that has followed often focuses on the perceived fact that 'manipulation' might nullify moral responsibility in this sense. If so, it would suggest that Frankfurt's 'real self' account is not defensible, because it has not isolated the special kind of capacity that is definitive of moral responsibility. In this context, Hana Kalivodová sets out to determine the extent to which moral responsibility is indeed compromised in cases of manipulation. Her discussion will also illuminate our conception of moral responsibility itself, she argues, because 'some undeniable ambiguity is inherent to the concept of moral responsibility' and 'manipulation cases show us the limits which can at least in a negative way say something about it' (p.23).

The author begins by providing a brief but instructive characterisation of recent developments in the discussion of free-will and determinism and the relation of the concept of moral responsibility to these developments. (pp. 8-13). In the main body of the thesis, she explores three significant approaches to moral responsibility that are relevant to the question of manipulation: these are (i) Frankfurt's own 'real self account'; (ii) Alfred Mele's historical account, which emphasizes how the subject arrived at their moral make-up; and (iii) Manuel Vargas' forward-looking account that appeals to the cultivation of moral agents, or 'building better beings'. These three contrasting accounts are well-chosen as they allow the author to compare their strengths and weaknesses and thereby display the complex and perhaps ambiguous character of moral responsibility itself (p. 87).

A significant omission here is a clear and unequivocal characterisation of 'manipulation'. The author introduces this central concept by invoking the influence of advertising and a case of Czech terrorism, which she says was inspired by online disinformation (pp. 15-16). But no clear and systematic attempt is made to define manipulation in distinction to (morally acceptable) influence and 'nudging' on the one hand, and (morally unacceptable) force, threats, and coercion on the other hand. Ultimately, we are told that 'manipulation is not easy to specify, but while it is difficult to find some common features of it, it doesn't mean that there is nothing' (p. 16). The danger here is that we will be left to our own private feelings as to what counts as manipulation and what does not, and these feelings may differ.

In the event, the author's discussion concentrates primarily on cases of 'global manipulation', such as evil neurosurgeons implanting decisions and even whole value-systems into a subject ('While Sally slept, Chuck's values were implanted into her system' p. 75), or God-like powers being used to preprogramme a zygote to perform an act exactly 30 years later (p. 19). It is questionable whether these thought-experiments, which resemble the 'deceiving God' or the 'evil demon' of Descartes' *Meditations*, have much in common with the cases of manipulation by advertisers and by agents operating in social and mass media, i.e. the cases that the author initially

used to introduce and illustrate the phenomenon of manipulation. Advertising, social media and mass media—for example, Fox News (p. 69)—cannot conceal their agency and influence on the subject, nor can they deprive the subject of the ability to critically assess and reflect on the decisions and acts that they promote. The author seems to recognise these crucial differences between the two senses of manipulation when she allows that in the ‘global’ cases there is both no knowledge of the threat of manipulation nor is there any access to the necessary ‘tools to avert its consequences’ (p. 89). There is therefore a question as to whether the attempt to address questions of media and internet manipulation in the conclusion (see particularly pp. 90-91) is really warranted by the main body of the text. The reader may feel that there are just two different senses of manipulation in play here and not much can be learnt about one from a discussion of the other.

There are several technical problems the thesis that should be mentioned. Among these is the disparity between the page numbers in the ‘Contents’ and the actual page numbers of the different sections. A reader looking for Part 4, ‘Historical condition for moral responsibility’, for example, is pointed to page 49 in the table of contents when the section in fact begins on page 71 of the text. The philosophically interesting section ‘Personal Identity’ (pp. 84-85) doesn’t appear in the Contents at all. In this connection it should also be said that this section fails to do justice to a key question that arises for cases such as the implantation of a whole value system in a person while they sleep (are they really the same person when they awake?). Occasionally there are puzzling compressions—on page 13 I was not sure what was meant by ‘Mele writes that the dif a compulsive hand washer’ and so I missed the point being made. It should also be observed that the thesis substantially overruns the maximum number of pages for a bachelor’s thesis (*‘Rozsah bakalářské práce je stanoven od 40 do 80 normovaných stran’*). There are 96 pages here and the pages themselves are in a format that exceeds the length of *‘normostrany’*.

But my overall evaluation of this bachelor’s thesis, despite the limitations mentioned, remains positive. The question of manipulation and moral responsibility is one that has not been widely discussed or taught in the Czech context, at least as far as I know, and the author should be commended for her initiative and enterprise in researching this difficult problem area and providing a good and reasonably clear survey and assessment of the principal standpoints. The bibliography displays uncommonly wide reading at this level, and the many texts are properly used in the author’s discussion. The author demonstrates that the three principal theories considered give philosophically interesting and divergent answers to the possibility of moral responsibility in globally manipulated agents (‘yes, no, maybe’ for Frankfurt, Mele and Vargas respectively as she memorably puts it at p. 87). It should also be mentioned that the author writes throughout in good clear English.

Jednoznačně navrhuji práci k přijetí. Doporučuji před obhajobou alespoň známku *velmi dobře*.

31.8.21, v Praze,

doc James Hill PhD