Intellectual History Review

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rihr20

Mandeville on self-liking, morality, and hypocrisy

Sandy Berkovski

To cite this article: Sandy Berkovski (2022) Mandeville on self-liking, morality, and hypocrisy, Intellectual History Review, 32:1, 157-178, DOI: 10.1080/17496977.2021.2012885

To link to this article: https://doi.org/10.1080/17496977.2021.2012885

Published online: 08 Feb 2022.

Submit your article to this journal

Article views: 135

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=rihr20

INTELLECTUAL HISTORY REVIEW
2022, VOL. 32, NO. 1, 157–178
https://doi.org/10.1080/17496977.2021.2012885

Routledge
Taylor & Francis Group

Check for updates

Mandeville on self-liking, morality, and hypocrisy

Sandy Berkovski

Department of Philosophy, Bilkent University, Ankara, Turkey

ABSTRACT
I explore Mandeville's account of moral judgement and its implications for the understanding of hypocrisy. According to Mandeville, we have a psychological need to like ourselves sufficiently, so as to carry on with our lives. Because our self-liking necessarily depends on the opinions others form of us, we are extraordinarily sensitive to praise and condemnation. The practice of moral judgement exploits this sensitivity. Hypocrisy is an intrinsic element of this practice.

KEYWORDS
Mandeville; self-love; self-liking; sociability; morality; hypocrisy

1. Introduction

Bernard Mandeville wasn't too fortunate with his readers. Some contemporaries, like Berkeley, dismissed him as a callous heretic who challenged the very foundations of morality and society. Others, like Rousseau, might have been influenced by him, but never acknowledged his influence. In the twentieth century, F.B. Kaye, who published the celebrated critical edition of the *Fable of the Bees*, seemed absolutely determined to deny him any major originality of thought. Although evidence for this, by Kaye's own admission, could be circumstantial at best, Mandeville was said to have borrowed all his ostensibly novel ideas not only from Hobbes, Spinoza, or Machiavelli, but also from La Rochefoucauld, Bayle, Gassendi, Nicole, and other French writers. Later in his influential discussion, Arthur Lovejoy declared Mandeville both "less original" and "even less unorthodox than he seems,"[1] and the author of a lengthy monograph on Mandeville repeatedly warns us that he was a "satirist" and a "wit" who "amused himself" with paradoxes, a "writer of light verse" who also liked to "write dialogues," who wanted to have "fun at the expense of the clergy and the universities," and who "liked to treat an argument like a pot of tea."[2]

One wonders what these commentators might say about Plato's dialogues. In fact, this familiar perception of Mandeville is similar to that ancient hobby reported by Diogenes Laertius of tracing every Platonic thought worth mentioning to Pythagorean doctrines. Having said that, I do not wish to mount a full-scale defence of Mandeville's originality. In part this is a thankless task, given the lack of historical evidence, and in part the problem no longer exists anyway, as more recent scholars treat him as an original and creative thinker whose views fully deserve serious scrutiny. My very limited purpose here is to develop just one line of argument that can be extracted from Mandeville's writings.

CONTACT Sandy Berkovski sandy.berkovski@gmail.com
© 2022 International Society for Intellectual History

Accordingly, in Section 2, I give an account of Mandeville's views of morality and its links to self-love and self-liking. In Section 3, I address the implications of this account for understanding Mandeville's views of hypocrisy. I argue that what Mandeville termed "fashionable hypocrisy" is continuous with "malicious hypocrisy," and that both stem from the self-liking mechanism. I argue further that malicious hypocrisy lies at the heart of moral condemnation and moral praise.

2. Mandeville's big idea

What is the ground of moral judgement? What are the reasons involved in moral justification? La Rochefoucauld had this cryptic aphorism in the first edition of the *Maxims* later removed:

> We never denounce vice or praise virtue except out of self-interest.[3]

In the context of La Rochefoucauld's thought, this maxim is quite unexceptional: after all, there is hardly anything we *don't* do out of self-interest. The root of moral condemnation is in self-interest, but so is the root of everything else in social life. Is there a further special connection between moral praise and denunciation and self-interest? La Rochefoucauld doesn't say.

Perceived similarities between Mandeville's claims in the *Fable* and elsewhere and many of La Rochefoucauld's maxims are nothing new. For example, the maxims V.2, V.11, or V.15 may be read as elegant précis of the main argument of the *Grumbling Hive*. Yet elegance is no substitute for substance. You can't read an argument off an aphorism. Without lengthy elaborations, the aphorisms would remain merely intelligent witticisms. In the present instance, however, it's not even clear what La Rochefoucauld meant exactly.

Whether inspired by La Rochefoucauld or not, Mandeville gave a systematic answer to our question. Morality, understood as a package of moral vocabulary, judgements, and practices, is an invention that can be traced, on one hand, to some people's pride and sensitivity to flattery and, on the other, to other people's manipulation of these psychological qualities. Here is a sketch of Mandeville's account that amalgamates the views in the first and second volume of the *Fable* and in the *Enquiry into the Origin of Honour*.

It begins with anonymous wise Lawgivers (or "politicians") who, once upon a time, pondered the possibility of human society.[4] As other commentators observed, we needn't take this talk literally.[5] Plausibly, it is a metaphor for a gradual emergence of society and its norms from the primeval condition of man. "Lawgivers" may be substituted – e.g. for "cultural evolutionary forces" – and the rest of the narrative paraphrased accordingly. For expository purposes, I will stick to the terms of Mandeville's metaphor. The Lawgivers, then, recognised and had to contend with two facts: the psychological fact that people are motivated by private gain even at the expense of others' interests, and the sociological fact that no peaceful, prosperous co-existence is possible among individuals so motivated. Their solution was to come up with a fiction. They wished to tweak human motivation in such a way that people would see satisfaction of others' interests as their own, that they would see contributions to society, even when made at the cost of personal losses, as being in their own interest. For this fiction to work, the Lawgivers had to appeal to the people's instinct of "self-liking." Let's digress for a moment to briefly comment on this notion.

It is explicitly introduced in *FB* II.129–30, in which we were told that self-liking is an instinct common to animals and humans to value oneself above one's real worth. It is a necessary condition for the exercise of "self-love" similarly common to animals and humans; that is, of the instinct for physical self-preservation.[6] Every person, Mandeville reasons, is motivated by private interest, but only so far as he continues to like himself. What good are the benefits when they are gained by someone you don't like? Well, that someone may well be you. You can only be motivated to gain more and more (to satisfy your narrow interest more and more) when you also want to benefit *yourself*; when, that is, you "like" yourself, when you "esteem" or "value" yourself sufficiently. But the proposed link to "real worth" is problematic even by Mandeville's own rights. First of all, what *is* the real worth, and how can it be measured? Secondly, Mandeville holds that our self-liking – or better, its product: self-esteem – is fragile and highly sensitive to the opinions of others.[7] But why believe that others should rate us above that fixed quantity of real worth? Hobbes gave a neat solution to this problem when he equated our self-worth with the aggregated evaluations of ourselves by others, but there is no evidence for attributing this idea to Mandeville.[8]

Instead, it is more promising to stress Mandeville's other line of thought, that people should hold themselves dear so as to be able to continue living.[9] Unless their self-esteem is at a certain level, people will not be motivated to survive, to love themselves enough to satisfy their most basic interests. Our self-esteem is fragile, because our worth cannot be known with certainty either by introspection or by observation. It is not a quality simply given in experience. You may form a belief about how valuable you are, but you seek to confirm it by gathering the opinions of others; hence our concern for the opinions of others is a fundamental feature of our psychology. We have to like ourselves to a sufficiently high degree, and this is possible only with the right responses of others.

To return to the Lawgivers, with their fiction they exploited this enduring concern with self-liking. To ensure co-existence, they rewarded those individuals who would serve public interest at the expense of their own private interests.[10] Generally, this could be achieved in two ways. They could alter the psychological makeup of people and weaken self-love itself: that is, people could be changed in such a way as to care for others at the expense of their private interests. We put aside this possibility, however, because we assume that people in society are not fundamentally different in their motivation from people outside our society. Alternatively, the Lawgivers could manipulate their subjects. They could fool them into contributing to society. For example, people could be rewarded with material objects. However, there are not enough material rewards available for each time a person acts in the public interest, so the Lawgivers chose another strategy: to reward citizens with "flattery," "applause," and "honour."[11]

When you contributed to society at the expense of satisfying your interests, you would then be praised. But praised how? What is the content of flattery that the Lawgivers are supposed to offer individuals who contribute to the public good? Given what was said earlier about the actual motives of men, people were divided into two categories. In the first category were those who managed to conquer at least some of their self-regarding desires – namely, those relevant on the occasion – and acted against them. In the second category were those who acted on their self-regarding desires. The former, says Mandeville, were declared the "finest," "most beautiful and valuable of their kind."

Those who failed to act for the public good were declared to be the "Dross of their Kind," little different from animals.

But, to come to a very critical question, why would anyone be responsive to this kind of flattery? If flattery is not a material good, what good is it? Well, we already know one answer: flattery boosts self-liking. Next to the satisfaction of material needs, people need to have a high self-esteem, which requires confirmation of the others' opinions of ourselves. Thus, when I flatter you, I indicate that I like you. This liking would translate into your higher self-esteem. But this cannot be a complete answer. First, it is unrealistic. We don't simply walk around recording the opinions of others in the hope of improving our self-esteem. If you approach a crowd of unfamiliar people and shout "do you like me?" to them, multiple ayes won't change your view of yourself. If you write a tweet liked at random by a thousand people, you won't necessarily like yourself more either.

If any kind of approval statistic should have an effect on your self-esteem, likes must be understood as signs of something else going on under the surface. You must assume that they were given for the right reason. Bare numbers won't cut it. In fact, Mandeville makes this very assumption, even if covertly. Rather than merely registering their approval, the Lawgivers had to declare those other-regarding desires "good" or "fine." In addition to making a factual judgement "we like you," they also had to introduce a value judgement "your action is good. *You* are good." There is a further assumption too: the value of your action should also be linked to the value of your self. Others might like your action, and you might, for this reason, approve of your action with more confidence. We assume, therefore, that the approval of actions must be able to increase one's approval of themself.

Let's spell this out a little more. Your action is liked by others (e.g. the Lawgivers) for the reason that it is valuable; valuable, that is, according to *others'* putative values. It is also a reason why, on this occasion, others like you and not just your action. You become aware of their liking, but, in order for this to boost your self-esteem, their values (relevant to the action in question) and yours must coincide. Thus, if the Lawgivers' flattery spreads across the community, and if it is effective in a wide range of situations, it must be accompanied by a community-wide change in values. Whatever values you may have had prior to their flattery – if indeed you had them at all – you now accept the values held by the Lawgivers. The same goes for the people around you, who are supposed to express their approval of the actions done for the public good. In other words, certain values, such as those that concern the superiority of people over animals, are *introduced* by the Lawgivers and adopted by the citizens. Because they are so adopted, approval of the action that benefits the community may boost the agent's self-esteem.

What, then, is Mandeville's conception of moral value? Some have argued that Mandeville is committed to rigorism and self-denial. A morally good action is performed when the agent acts against his own interest. Others thought that he was a utilitarian. After all, an action that benefits society is the one that is declared "virtuous" by the Lawgivers.[12]

I think the answer is more nuanced. The Lawgivers first inform the people that anyone able to control his animalistic urges and direct his action toward the public good is better than the one who is governed by self-regarding desires:[13]

> [The Lawgivers made] the People they were to govern, believe, that it was more beneficial for every Body to conquer than indulge his Appetites, and much better to mind the Publick than

> what *seem'd* his private Interest. [T]hey extoll'd the Excellency of our Nature above other Animals, and setting forth with unbounded Praises the Wonders of our Sagacity and Vastness of Understanding, bestow'd a thousand Encomiums on the Rationality of our Souls, by the Help of which we were capable of performing the most noble Atchievements. Having by this artful way of Flattery insinuated themselves into the Hearts of Men, they began to instruct them in the Notions of Honour and Shame; representing the one as the worst of all *Evils*, and the other as the *highest Good* to which Mortals could aspire. They indeed confess'd, that those impulses of Nature were very pressing; that it was troublesome to resist, and very difficult wholly to subdue them. But this they only used as an Argument to demonstrate, how *glorious* the Conquest of them was on the one hand, and how *scandalous* on the other not to attempt it.[14]

The persuasion strategy begins with some factual claims about the difficulties of self-control. It continues with the glorification of those who can control their animal natures, and with the shaming of those who cannot. Here the Lawgivers appeal to the two-stage mechanism of self-liking outlined earlier. This stage is complete when the Lawgivers manage to instil the sense of good aligned with self-control in the minds of their citizens. The last stage is to make people see that their *real* private interest is to benefit the public and to act on it.[15]

Why benefiting the public is in their own interest is, of course, explained again by the appeal to the two-stage self-liking mechanism: it is in your interest to gain the approval of others, at least if that approval is explained by your "better" nature. The Lawgivers, therefore, aimed to persuade people that certain qualities are "better," and to this extent they aimed to create a new system of values. We are not expressly told in what way and why these qualities are better. Commenting on this passage, Lovejoy describes the new values as dictated by a sense of "generic" or "racial" superiority of the human race over the animals.[16] Men were superior to animals chiefly due to the possession of reason and control over self-regarding desires. Though Mandeville takes no stand on these matters, this doesn't have to be the full explanation. After all, what is the intrinsic value of reason? Perhaps the Lawgivers could say further that action itself is possible when urges and desires are brought under rational control. Unless this is so, there is no choice and no action. On the other hand, since people think of themselves as agents, they should necessarily think of themselves as rational creatures with the ability to control urges and desires. Hence the Lawgivers' flattery of the rational capacity of citizens is persuasive because those citizens already see themselves as endowed with the ability of rational self-control; an ability that is not always easy to exercise, however.

Whatever the exact mechanics of their introduction, are the values of reason and control introduced at this first stage *moral*? Mandeville's answer is somewhat ambiguous. On one hand, he says opaquely, these flattery-based innovations by the Lawgivers contrived "the first Rudiments of Morality" by making people "useful to each other as well as tractable."[17] As people began to contribute to the public interest, began to cooperate with each other, they became more peaceful. But are we entitled to conclude that the reining of animal "impulses" deserves to be called "moral"? Says who? The Lawgivers need not call it "moral," as their aim at this stage is to tell the citizens what kind of people, from now on, will be praised or shamed. This they can do without using any moral predicates:

> We will like you and shower you with compliments if you resist your desires and temptations of sensual pleasure, and use your rational faculty. We will shame you if you fail to do so.[18]

The citizens should not call themselves "moral" either. Their condition changes only in this regard: that they learn how to earn praise from the Lawgivers and their peers.

Perhaps, then, we as philosophical observers should call such newly acquired self-control "moral." That would be the "origin of moral virtue" that we were supposed to discover. It would mean that moral discourse began with the Lawgivers' contrivance. But Mandeville does not take this step. All that the Lawgivers seem to have achieved with their contrivance is to persuade people that the promotion of public interest may be in their *own* interest. They showed their subjects how prudential calculations yield actions that benefit society. This kind of rhetoric need not be couched in moral terms. We need not assume that any newly invented moral discourse was part of the Lawgivers' persuasion.[19]

According to Mandeville, there is another momentous change brought about by the Lawgivers' innovations that does result in the emergence of moral discourse. But it is not due to the Lawgivers themselves. Mandeville describes it thus:

> It being the Interest then of the *very worst of them*, more than any, to *preach up* Publick-spiritedness, that they might *reap the Fruits* of the Labour and Self-denial of others, and at the same time indulge their own Appetites with less disturbance, *they agreed with the rest, to call* every thing, which, without Regard to the Publick, Man should commit to gratify any of his Appetites, VICE; if in that Action there cou'd be observed the least prospect, that it might either be injurious to any of the Society, or ever render himself less serviceable to others: And to give the Name of VIRTUE to every Performance, by which Man, contrary to the impulse of Nature, should endeavour the Benefit of others, or the Conquest of his own Passions out of a Rational Ambition of being good.[20]

We are given here the origin of the moral idiom. The "very worst" citizens introduce the language of virtue and vice with the express purpose to enjoy the contributions of others, i.e. of the better citizens. They stipulate the moral value of the actions of others, and for some unspecified reasons their stipulations are accepted. Cleaning up Mandeville's formulations just a little, the proposed stipulations work according to this schema:

> **(Good)**
>
> An action *x* is morally good if and only if *x* is aimed either to benefit the society or to control one's self-regarding desires.
>
> **(Bad)**
>
> An action *x* is morally bad if and only if *x* is aimed either to satisfy the agent's self-regarding desires at the expense of the society's interests or merely to indulge in those desires.

Broadly put, actions are distinguished according both to their effects and their motivations. Actions motivated by other-regarding desires are good, those motivated by self-regarding desires are bad. Actions that benefit society are good, those that benefit the agent at the expense of society are bad. The great novelty of Mandeville's idea is, of course, not this conception of virtue as self-denial. It is the *use* that the idiom of virtue is put to. Why are the terms "good" and "bad," "virtue" and "vice," introduced in the first place? To "preach" and to "reap the fruits," Mandeville tells us. The moral idiom is a rhetorical instrument used by the "worst" of mankind to manipulate the others into benefiting society, and indirectly, the worst themselves. The manipulation mechanism builds on the self-liking mechanism and involves

rewards and punishments. "Virtuous" actors are met with approval from their peers, "vicious" ones are met with disapproval. Approvals and disapprovals can punish and reward because people source the opinion of others to confirm their views of themselves. It is assumed that a low self-esteem is associated with suffering, a high one with pleasure.

Now, there is a fly in the ointment here; two in fact. We are never told why anyone would accept (Good) and (Bad) that, after all, come from the "worst" of mankind. Why would the rest of the citizens care about the opinions of those manipulators in the first place?[21] Secondly, how could the manipulators "manipulate" and benefit themselves merely by introducing a novel idiom into the discourse? As for the first question, a plausible explanation is that the manipulators certainly never presented themselves *as* manipulators. By dispensing praise and condemnation, they presented themselves as caring about public utility and proper self-control. Equally, they never meant to challenge the already extant valuational code established by the Lawgivers. Rather, they meant to turn its acceptance by other citizens to their own advantage.

By the same token, (Good) and (Bad) do not introduce any new concept. We may read them rather as reductive definitions of "moral" (or "virtue"), whereby the left-hand side is a handy abbreviation of the right-hand side of the biconditionals. But the manipulators did invent a new use for the already available concepts and corresponding terms. They intended to use them on public occasions with two effects. They shunned and shamed the putative violators of the valuational code, and by doing so distracted the public's attention from their *own* failings. They praised those who putatively conformed to the valuational code, and by doing so affiliated themselves with those "good" people and improved their own reputation (whilst secretly, again, violating the code). This novel use was meant to serve, therefore, the manipulators' hypocrisy. I return to this in Section 3.

We may now restate Mandeville's view and re-interpret La Rochefoucauld's Delphic aphorism as follows. When we judge an action moral or immoral, we describe it as performed with a certain kind of intention specified in (Good) and (Bad). But crucially, that is not *why* we make such a judgement in the first place. Our purpose is not descriptive and classificatory. The judgement is made against the background assumptions of what actions people normally approve and disapprove. When, therefore, we say that a person acted morally, we convey the information that we, and also possibly some others, approve of his actions. As the agent enjoys the approval of others (because of the self-liking mechanism), he is encouraged to perform this and similar actions again. Analogously for the situation of immoral actions. Finally, I have said that "we" make moral judgements. That's not entirely accurate. Not just any kind of person, on any kind of occasion, is likely to issue a moral judgement. It is, rather, the people who themselves do not normally act for the benefit of society and instead are busy satisfying their narrow private interests.

Thus, an interesting novel idea that Mandeville elaborates in the *Enquiry into the Origin of Moral Virtue* is the function of moral judgement. Quite apart from the question of what constitutes virtue and moral value and how they relate to flattery and self-liking, Mandeville also asks who in fact uses moral judgements, when, and why. Moral judgement is a rhetorical device of manipulation. It is designed to dispense praise and condemnation with the view of serving the narrow self-interest of the judge.

If this reading is correct, then one familiar problem in the Mandevillean scholarship may be seen in a new light: namely, the distinction between "genuine" virtue and apparent, "counterfeited" virtue. The former requires conquest of passions and realignment of desires: a person is genuinely virtuous if he acts on other-regarding desires. Apparent virtue has no such requirement. It is enough to convince, to signal to the audience that one acts for the benefit of others. Genuine virtue is very hard or impossible to achieve, but apparent virtue can readily be mastered.[22] The problem here is why we hold on to this very strict notion of real virtue in the first place. Why, in particular, is it socially useful to distinguish between two kinds of virtue?

Examining this question, Christian Maurer gives an historical explanation.[23] On one hand, Mandeville is strongly motivated to reject Shaftesbury's views in which moral virtue is aligned with conformity to the norms of politeness. This, Maurer reads Mandeville as saying, delivers no more than apparent virtue(s). On the other hand, Mandeville was influenced by the French Augustinian and Calvinist traditions that stressed the postlapsarian condition of mankind. Genuine virtue requires self-denial and is, for that very reason, unattainable for us. To be virtuous is incompatible with our postlapsarian corrupt nature. This reading, I think, makes Mandeville's argument somewhat uninteresting. In effect, we have to concede – admittedly with some textual support – that Mandeville simply relayed his pre-conceived beliefs. Rhetoric and theology were the reasons behind his distinction. Even if this were so, could we attribute to Mandeville arguments, rather than stipulations, that could be defended on independent grounds? That's the approach I have taken here. As Maurer observes, we may then run the danger of being anachronistic. Perhaps this is a risk we should be willing to take.

Another explanation is given by Robin Douglass, who argues that people need an unattainable ideal of virtue that can be emulated only imperfectly in practice. The project of sociability initiated by the Lawgivers depends on the introduction of this ideal.[24] Why, however, couldn't the Lawgivers be satisfied with apparent virtue; that is, with an utilitarian standard of benefiting society, never mind the motives? Douglass seems to argue that the idea of "real virtue" is necessary in order to motivate people to do better. Praise of others encourages them to "imitate" the ideal of virtue that "cannot be attained."[25] This solution has a decidedly non-Mandevillean flavour, however: though the fundamental origin of their motives is external praise, at the end of the day many people, on many occasions, are motivated in their behaviour by the pursuit of virtue! The gap between Mandeville and his opponents like Hutcheson hasn't closed, but still narrowed significantly.

Hector Monro offers a more straightforward gloss on the passage above.[26] He claims that real virtue is necessary, as flattery targets the ability to subdue passions and distinguish yourself from animals. People's self-liking is guaranteed to increase when they compare themselves to animals, who allegedly cannot subdue their passions. Though this is what Mandeville himself says in the quotation above, it isn't too convincing: animals are not in the habit of benefiting creatures other than their kin.[27] People may sufficiently be distinguished from animals just by the scope of their socially useful actions. In any event, once people have acquired some capacity of self-control, their differentiation from animals would be beyond dispute. At a later stage of social development, praise couched exclusively in terms of the animal–human contrast would lose its effectiveness.

On the interpretation I have offered we get a somewhat different explanation. Socially useful behaviour is reinforced by praise; but we take praise seriously and enjoy it only when its object is our motives. In dispensing their praise, the others are supposed to judge what we *are*, and not merely the effects of our actions. That is true when the Lawgivers praise citizens for their mastering of self-control, and it is true also at a later stage when the moral predicates of "virtue" and "vice" are introduced by the "very worst" people to manipulate the rest. At that later stage, we need not assume that people compete with animals for praise: telling them that they are better than animals would be at best faint praise. They do, however, compete with other people. Therefore, on one hand, when they receive praise couched in moral terms, they have to assume that their motives are praiseworthy and virtuous. On the other hand, at this stage of social development we should assume that many people are able to perform actions useful for society at large. Then we expect that those people would also be praised for their performance. But then, if you are praised for being virtuous, this praise should be devalued in the presence of many other similarly virtuous agents.

Therefore, to sustain your self-liking and to interpret only the praise directed at *you* as valuable, you would have to discount the praise given to other people. A good way to do exactly this is to assume, and also claim publicly, that you are praised for really having virtuous motives in performing socially useful actions, while others are praised when they manage to fake their virtuous motives. The two different conceptions of virtue emerge, rooted in the different ways we evaluate our own performances and the performances of others.

Finally, and somewhat confusingly, Mandeville also recognises a separate category of "refined" virtue that, he admits, is not part of the Lawgivers' innovations. It is the virtue of performing good for the sake of good "in silence."[28] To practise this virtue you have to be satisfied to perform morally good actions in privacy when no external reward like praise is forthcoming. Mandeville doesn't explain how, or why, we acquire the conception of this virtue. But he insists that, as things stand, even practitioners of this virtue are driven by pride in the "contemplation of their worth." Self-liking, therefore, is the main driving force behind this form of virtue as well. On the other hand, the difficulty for Mandeville would be to explain how the exercise of this virtue would increase our worth in our own eyes if, by assumption, it is not accompanied by praise (that could confirm our increasing worth). I return to this issue in Section 3.

3. Hypocrisy

In this section, I explore the links between Mandeville's views of morality and self-liking, understood in accordance with the interpretations in Section 2, and his views on hypocrisy. I argue that there is continuity between what Mandeville called "fashionable" and "malicious" hypocrisy, and that it can be traced to the manipulation of others, in turn rooted in self-liking.

Mandeville describes two kinds of hypocrisy in the following key passage in the *Enquiry into the Origin of Honour*:

> *Cleomenes.* There are two Sorts of Hypocrites, that differ very much from one another. To distinguish them by Names, the One I would call the Malicious, and the Other the

> Fashionable. By malicious Hypocrites, I mean Such as pretend to a great Deal of Religion, when they know their Pretensions to be false; who take Pains to appear Pious and Devout, in order to be Villains, and in Hopes that they shall be trusted to get an Opportunity of deceiving those, who believe them to be sincere. Fashionable Hypocrites I call those, who, without any Motive of Religion, or Sense of Duty, go to Church, in Imitation of their Neighbours; counterfeit Devotion, and, without any Design upon others, comply occasionally with all the Rites and Ceremonies of Publick Worship, from no other Principle than an Aversion to Singularity, and a Desire of being in the Fashion. The first are, as you say, the worst of Men: but the other are rather beneficial to Society, and can only be injurious to themselves.
>
> *Horatio.* Your Distinction is very just, if these latter deserve to be call'd Hypocrites at all.
>
> *Cleomenes.* To make a Shew outwardly of what is not felt within, and counterfeit what is not real, is certainly Hypocrisy, whether it does Good or Hurt.
>
> *Horatio.* Then, strictly speaking, good Manners and Politeness must come under the same Denomination.[29]

To make sense of these claims, let's consider first the "fashionable" hypocrites. To understand what they are after, and why their behaviour should be classified as hypocritical, we begin with self-love. If, as Mandeville argues, it is the governing motive of all behaviour, is there a place left for other-regarding desires?

According to what we might call "strong reductionism," all other-regarding desires are, on a proper analysis, an illusion.[30] Certainly, sometimes you *act* in the interests of others and benefit them. But when you so act, you are not *motivated* to benefit them. All actions are construed as involving a sequence of acts and a single motive. Some of your acts in the sequence benefit others, but all of them are motivated by the same motive: your self-regarding desire. Consider exchange: when you buy food at the shop, your complex action necessarily has the effects of benefiting others. Conceivably, you are generally aware of these effects, and you welcome them. Yet it is also plausible that you buy food to benefit yourself, and that the whole action is sufficiently motivated by the desire to alleviate your own hunger. There is no need to introduce an intermediate other-regarding desire to benefit the shopkeeper, say, to explain why you came to the shop etc. The strong reductionist sees all actions as conforming to this kind of model.

Mandeville was not a strong reductionist. He was perfectly happy to recognise the reality of other-regarding desires and their motivating force. His position is "weak reductionism": every complex action involves a hierarchy of motives corresponding to the causal sequence of acts to be performed. The ultimate motive will be self-love, and the ultimate element in the sequence will be satisfaction of a self-regarding desire. Among the intermediate motives there may well be other-regarding desires. Take pity. Mandeville devotes several moving passages to the suffering of animals.[31] We have a "real passion" in our nature to alleviate their suffering. That is why on some occasions, and quite inconsistently, we abhor any harm done to them. We can be genuinely motivated by pity to abstain from food if it entails killing an animal. But is pity, at least on those occasions, the governing desire in the motivational hierarchy? Having acknowledged its power, Mandeville insists that pity is subordinate to self-love. Ultimately, we act to alleviate our own suffering. Pity makes us bothered by the suffering of animals and other creatures. We genuinely suffer when we have to observe *their* suffering. It is our own suffering that drives us to reject the suffering of others. As another example,

consider virtue for its own sake. People derive pleasure from performing good acts in private chiefly because this increases their self-esteem.[32] This, Mandeville says, is a "certain sign of pride" evident upon reflection.

Hence, on Mandeville's weak reductionist view, at the top of the motivational hierarchy we have to posit self-love, even if in order to satisfy it we also have to be motivated by other desires. Compared to the strong reductionist explanations, the reality of motives is here in reverse. According to strong reductionism, you are quite aware of your self-regarding desire (say, hunger satisfaction in the case of economic exchange) and unaware of your other-regarding desires; certainly so, because they don't exist to begin with. According to weak reductionism, on the other hand, you are aware of your self-regarding desires and unaware of your self-love. The latter, however, is so not because self-love is not real, but rather because you are not likely to have an insight into your own psychology. It takes a philosophical psychologist to tell you how your motivation is in fact organised.

From this portrayal of human motivation, it follows that an exceedingly frequent condition of human actors is self-deception. In the course of their exposure to social life and becoming civilised, they learn to develop social motives and then act on them "spontaneously."[33] In their usual activities, possibly in *all* of them, they are motivated by self-love, but are very rarely aware of it. And on the occasions when they are motivated by overtly other-regarding desires, when their feelings for others are psychologically real, they again fail to see self-love as the origin of those motives and feelings. As in other cases, this self-deception is not caused by a special cognitive disability of the agents. Instead, the agents have evidence that their motive is self-love, but prefer to forget it at some stage, or to ignore it. Through habit and lack of effort, they develop the tendency to see themselves as driven by motives other than self-love. This is an effect of the socio-moral framework instituted by the Lawgivers. The society encourages people to care for others. Those individuals who act for the good of others are praised. They are praised, as we saw, not merely for the effects of their actions, but for the kind of individuals they are. Hence, so far as every person wants praise, he is strongly motivated to *create an impression* that he cares for others. What his real, ultimate motives are the person is not encouraged to discover. He is instead encouraged to ignore those deeper motives that would not earn him praise.[34]

Although self-deception among fashionable hypocrites is frequent, conscious deception is also possible. As Mandeville implies, this especially would be the case among the less-trained social actors. People may act out of an explicit motive of self-liking. You might go to a charity dinner either because it is simply de rigueur in your social milieu, or because you specifically wish to impress some guests. In the latter case of conscious deception, you would make a special effort to display your generosity, whereas in the former, in which you have trained yourself to develop the motive of generosity, no such effort is necessary. Still, on the outside you might appear the same in both cases.

Therefore, I take it that the motivation of fashionable hypocrites, as described by Mandeville, may be characterised either by self-deception or by deliberate deception. But is it right, as Horatio wonders, to classify them as "hypocrites"? Cleomenes explains that the discrepancy between outward behaviour and what is "felt within" is a sufficient sign of hypocrisy. As a general formula, this won't do. As compellingly argued by Szabados (1979), in a wide variety of cases the inward–outward discrepancy is not hypocrisy.[35]

A patriotic British spy in Moscow who drinks to Putin's health misleads the audience about his motivation and so "counterfeits" his devotion to the Kremlin bosses. The spy is deceitful, but surely not "hypocritical"! Closer to our themes, a guest who dislikes the host's cooking, but gently and courteously praises it all the same, is exactly the fashionable hypocrite Mandeville has in mind. But is he, really? Calling him a "hypocrite" seems unduly harsh. Horatio's doubts were in order. The inward–outward discrepancy is perhaps a necessary condition of hypocrisy, but arguably not a sufficient one.[36]

In fact, earlier in the First Dialogue, Cleomenes himself denies the sufficiency of the inward–outward discrepancy:

> *Cleomenes.* [T]here could be no Religion so strict, no System of Morality so refin'd, nor Theory so well meaning, but some People might pretend to profess and follow it, and yet be *loose Livers*, and wicked in their Practice.
>
> *Horatio.* Those who profess to be of a Theory, which they contradict by their Practice, are, without Doubt, hypocrites.
>
> *Cleomenes.* I have more Charity than to think so. There are *real Believers* that lead Wicked Lives; and Many stick not at Crimes, which they never would have dared to commit, if the Terrors of the Divine Justice, and the Flames of Hell, had struck their Imagination, and been before them in the same Manner as they really believe they shall be; or if at that Time their Fears had made the same Impression upon them, which they do at others, when the Evil dreaded seems to be near. Things at a Distance, tho' we are sure that they are to come, make little Impression upon us in Comparison with those that are present and immediately before us.[37]

What is said here about religion should also hold good for morality. The non-hypocritical "loose livers" genuinely believe the injunctions of religion or morality, but lack a sufficiently strong motive to follow them through. Punishments and rewards promised by religion or morality are too uncertain for them to be motivating. One may doubt, of course, whether Cleomenes' descriptions are cogent. In the case of religion, it may be asked whether the loose livers should be classed as "real Believers," given that their religious reasons are too weak to propel them to action. What does their Christian *belief* in God consist in, if the Christian God is consistently powerless to make them act? By the same token, we might doubt whether they could be described as holding moral beliefs, if their alleged beliefs consistently fail to translate into action.

This worry aside, our present exegetical concern is to explain why fashionable hypocrites are any more hypocritical than loose livers: what is it in their character or behaviour that makes them hypocrites? Horatio seems to endorse the view that Cleomenes defends later in the book (in the passage quoted earlier), but here, Cleomenes demurs. To get a more accurate picture of hypocrisy, let us turn to the "malicious" hypocrites. Cleomenes describes them as building up their reputation through the inward–outward discrepancy, all in order to "deceive" at some later date. A malicious hypocrite publicly announces his moral commitments, while biding his time to get a chance to deceive others further. To count as hypocritical, the inward–outward discrepancy must be coupled with a further deception that it is meant to enable. But enable how? and what is this deception supposed to be about in the first place? After all, our British spy too misleads his Russian audience about his real motives in order to perpetuate a *further* deception.

Mandeville's explanations in this passage are cut short, but we can connect the dots if we identify malicious hypocrites with the "very worst" people of the *Enquiry into the Origin of Moral Virtue*, who made the moral idiom a common currency.[38] Recall that those people engage in a moralistic manipulation: hoping to profit from the efforts of other citizens, they use moral idiom to praise and condemn others depending on their contributions to society, whilst *themselves* failing to conform to the demands of virtue. By encouraging others to follow moral norms, they manipulate them to carry the burden of social investment. Those who comply are praised and rewarded. Those who refuse are shunned and ostracised. At the same time, our malicious hypocrites privately eschew moral norms themselves. Their public efforts, camouflaged as directed at a public good, are in fact directed at a personal goal. This goal, in general terms, is such that anyone *else* pursuing it would be condemned by the moralising manipulator for failing to conform to the moral norm. This, of course, was precisely the aim of the "very worst" people in the *Fable*: they wanted to satisfy their self-regarding desires whilst publicly condemning anyone for doing that and employing the characteristic moral idioms in their condemnations. Hypocritical praise is a mirror image of hypocritical condemnation. A hypocrite praises others for some moral act, but refuses himself to act so and instead prefers to act viciously when given a chance.

Putting all of this together, we have the following general formulae:

Hypocritical condemnation

x hypocritically condemns *y* for *F*-ing if: *x* claims that, in *F*-ing, *y* acts immorally, *x* claims or implies that anyone who is *F*-ing acts immorally, *x* himself is either secretly *G*-ing or is planning to *G* in the future at the expense of some morally good *H*-ing (where *G*-ing is either the same as *F*-ing, or is condemnable for the reasons sufficiently similar to the reasons cited by *x* in condemning *F*-ing).

Hypocritical praise

x hypocritically praises *y* for *F*-ing if: *x* claims that, in *F*-ing, *y* acts morally, *x* claims or implies that anyone who is *F*-ing acts morally, *x* himself declines to *G* in favour of some morally repugnant *H*-ing (where *G*-ing is either the same as *F*-ing, or is praiseworthy for the reasons sufficiently similar to the reasons cited by *x* in praising *F*-ing).

Malicious hypocrites hope to deceive others when given an "opportunity" in such a way that their deception will cover for an action condemnable for the same moral reasons they themselves now use to condemn or praise. Their malice is in having an ultimate immoral goal for which their present moralistic discourse is just a ploy.

To return to the earlier question: should fashionable hypocrites be properly accused of hypocrisy too? What are they accused of, exactly? The whole point of Mandeville's distinction seems to be to deny that, through their pretence, fashionable hypocrites mean to achieve a malicious goal. Let us look at them more closely, though. Suppose, as before, that you go to a charity dinner merely in order to fit in. Your friends attend, so do you. Suppose too that you do not pontificate about charity. You are not a malicious hypocrite. Still, fitting in *is* designed to advance your goals of gaining a better reputation. In turn, a better reputation is a means to protect and improve your social standing. And so on. When we ask why the individual developed a habit of performing such actions, we expect to discover that the reason was self-liking. True, a fashionable hypocrite does

not praise or condemn other people. Ostensibly he only minds his own business. Yet, just like malicious hypocrites, he uses outwardly moral behaviour to advance narrow personal interests; and, of course, compliance with the moral "fashion", with the public standards of right and wrong, is intimately related to malicious hypocrisy, for those standards themselves were set by malicious hypocrites. A fashionable hypocrite is an imitator, and a malicious hypocrite is a trendsetter.

Fashionable hypocrisy is thus continuous with malicious hypocrisy. It deserves to be called "hypocrisy" when we look at its practice as a whole and identify the central explanatory role played by self-liking, the desire to seek the approval by others. Their guilt is to "shew outwardly," for the consumption of others, what they don't have within. Then we can explain why "loose Livers," aka "real Believers," are not hypocrites. Their loose behaviour is not motivated by praise from others and social standing. Persuaded by moral and religious precepts, they do wish to put them in practice. Alas, they fail. Theirs is a case of akrasia, of surrender to desires and urges against their best judgement. However we describe their motivational structure that leads them to violate the precepts they find most convincing, the decisive difference between them and hypocrites is the primacy of outward appearance for the latter, but not for the former. As with all people, according to Mandeville, praise by others should have some role in the explanation of the attitudes and behaviour of loose livers on *other* occasions. The point Mandeville is making is that you can live against your own beliefs, and so far as this is done away from the gaze of others, not on a public occasion, you will not qualify as a hypocrite. Of course, when loose livers begin to air their moral beliefs in public, they should rightly be accused of hypocrisy.

So far, I have followed Mandeville's sketch in the *EOH* passage above and assumed that malicious hypocrites are deliberate manipulators; but, in fact, Mandeville himself is aware that malicious hypocrisy is often much more complex. Most malicious hypocrites, and certainly the very successful ones, must be victims of self-deception. Just before that *EOH* passage, Cleomenes says:[39]

> [C]onsidering, that, to save Appearances, Hypocrites are at least as good as the sincere Men I have spoken of, it is impossible, that there should not be a great Shew of Religion among them, if there were but eight or ten of them sincere in every Hundred: And where such Pains should be taken to make the Men seem to be Godly; and this Point of outward Worship should be labour'd with so much Diligence and Assiduity, I am persuaded, that many even of those, who should be too wicked to be Hypocrites, and to counterfeit long, would sometimes, not only pray in good Earnest, but likewise, set on by the Examples before them, be transported with real Zeal for the Good of their Cause.[40]

Hypocrites *simpliciter* here are fashionable hypocrites. Those that are too wicked even to be called "hypocrites" are malicious hypocrites, as Mandeville explains in the passage already quoted. Malicious hypocrites, therefore, are supposed to pretend that they are virtuous. They are supposed to speak and behave as virtuous men, convincingly so. It takes a greater effort to convince when you don't believe yourself, so it pays off to convince yourself as well in what originally was just a ploy to secure a certain benefit. To speak and act convincingly, the malicious hypocrite will have to convince himself that he is virtuous and so re-interpret his actions as virtuous. In this sense, he comes to believe his own lies and so deceives himself.

Mandeville's simple argument linking self-deception to persuasion anticipates modern ideas. According to Stanislavski's influential system, acting and pretence are much more convincing when the actor becomes one with the character, when he "embodies" the character.[41] Presumably embodiment would be inconsistent with the continuing deliberate deception. Mentally, it would become too costly to be aware of your pretence and to maintain a distance between who you are and whose part you are supposed to play. Similarly, evolutionary psychologists stress the advantages yielded by "self-enhancing bias."[42]

This conclusion, moreover, sits well with Mandeville's account of self-liking. For, suppose that a malicious hypocrite does not deceive himself. That is, he intends to fool others into virtuous acts (say, charity work) while privately swimming in vice (say, bribery and hoarding) that he *himself* understands as being against his declared virtues. Like anyone else, the hypocrite desires praise. He expects to get it, of course, as he took care to cover up his vices. On the other hand, he receives praise *as* a virtuous man. That is clear to him and to his commenders alike. Now, from his perspective of a clear-eyed deceiver, he is not a virtuous man. He is rather someone who fools people, hence quite vicious. Therefore, those who praise him do so for a *wrong reason*. But then, as I argued earlier, their praise is idle for him. Normally we don't wish to be praised for wrong reasons by ignorant people or random people. Occasionally such praise can be tolerated and welcomed, but if *all* the praise that we receive is of such inferior quality, our self-liking would not increase and most probably would fall. Thus, to profit from praise and to increase his self-esteem, the hypocrite would have to convince himself that the praise is deserved, and that he really is a virtuous person.

But now, having stressed the possibility of their self-deception, haven't we taken malice out of malicious hypocrites? If their pontification and moralistic manipulation encourage socially useful practices, how bad can they be? Lecturing people about carbon emissions whilst flying a private jet can't be that bad, if you raise awareness about climate change. The line between fashionable hypocrisy, which is socially useful, and malicious hypocrisy now appears blurred. There is a tension in the passages quoted between the claim that malicious hypocrites *are* motivated by moral reasons ("pray in earnest") and the claim that they are merely pretending to do so ("pretend to a great deal of religion") and only scheming to deceive others to achieve their putatively villainous goals.

One way to resolve the tension is to say that malicious hypocrisy begins as a conscious manipulation of other people for the purpose of securing one's manifestly selfish ends, such as political power for its own sake. The perceived extra malice of malicious hypocrites has to do with the fact that their most important projects are plainly immoral, while moralistic rhetoric is their main instrument of success. Though the malicious hypocrite may come to believe his own preaching, this can't obscure the essential immorality of his goals.

Mandeville gestured at exactly this solution in his remarks on Cromwell's character.[43] Cromwell, Mandeville says, initially wanted popularity, and he "studied day and night" to gain it. He could join any side of the conflict to fulfil his political ambition. Everything else was just a means to that overarching goal. In this sense, Cromwell's motivation should be "one continued scene of hypocrisy."[44] But then, it's not surprising that he did some good, too: in the course of his long career, there would, very probably, be

occasions when the good of the country coincided with his ambition, hence we ought not to be surprised that Cromwell took good care of the English commonwealth; once he had made its well-being "inseparable" from his own.[45]

To complete the contrast between two hypocrisies, we might say that fashionable hypocrites have no such overarching goals. With them, it is in reverse: their dominant goals may be innocuous or even noble, but to get along with other people, to live in a society, they occasionally have to pretend to have feelings and beliefs they don't actually have.[46] They certainly don't use manipulation as their main instrument of advancement.

Let me wrap up this discussion by contrasting it with two recent readings of Mandeville's views of hypocrisy.

David Runciman credits Mandeville with the distinction between first-order and second-order hypocrisies:[47]

> First-order hypocrisy is the ubiquitous practice of concealing vice as virtue, which makes up the parade of our social existence. Second-order hypocrisy is concealing the truth about this practice, and pretending that the parade itself is a form of genuinely virtuous, and therefore self-denying, behaviour. We may need to hide the truth *about* ourselves in order to get by in this world, but we oughtn't to hide the truth from ourselves that this is what we are doing.
>
> [One can] move from the first-order hypocrisy of faking virtue to the second-order hypocrisy of trying to find a way to make that fake virtue appear real.[48]

As stated, the distinction isn't fully clear. If we faked virtue, would we not also pretend that our fakery is genuine, and so try to make its product – i.e. fake virtue – appear real? Otherwise, what's the point of fakery? One possibility is that fakery itself is widely understood to be fake, so that it doesn't mislead anyone. Kant took this view in *Anthropology*[49] 7:152. Runciman, however, intends this distinction to turn on self-deception. First-order hypocrisy characterises politeness. It is widespread in commercial societies in which politeness prevails.[50] Outwardly, acts of politeness are expected to conform to the demands of virtue. Inwardly, these acts are immoral, being driven by self-love. A polite person, though he pretends to be virtuous and act out of virtuous motives, does not mislead *himself* about the nature of his acts. He is quite aware that this is just a pretence, and that his actions are polite, and no more than that. Second-order hypocrisy begins when that same person "forgets" the fact of his pretence and comes to believe, erroneously, that politeness is itself a virtue.[51] A distinctive mark of the second-order hypocrite is his self-deception; that is, his wilful ignorance about self-love as the ultimate motive of his actions.[52]

Yet, as Runciman admits, the distinction can't turn on self-deception alone.[53] We expect ordinary polite individuals, i.e. the first-order hypocrites, to be carried away by their politeness and to mistake it for real virtue, too. I agree: consider e.g. politeness manuals that often attribute special moral significance to their injunctions. A truly polite person is taken to be a moral person. This tendency is evident in the earliest such manuals.[54]

Still, Runciman insists, the first- and second-order distinction *must* be attributed to Mandeville, as he clearly wished to distinguish between a useful hypocrisy and a hypocrisy that must be avoided.[55] One might think that Runciman's orders of hypocrisy, rather, correspond to fashionable and malicious hypocrisies, because, precisely with the latter pair, Mandeville was explicit and emphatic about "good" and "bad" hypocrisies.

Runciman considers this possibility and remarks that, like first-order hypocrites, malicious hypocrites exploit "the ubiquity of first-order hypocrisy in order to conceal sinister motives."[56] As far as I can see, this doesn't follow from anything that Mandeville says about the fashionable/malicious distinction. Malicious hypocrites need not rely specifically on the norms of politeness and obedience to customs to pontificate and manipulate, unless in a trivial way: everyone in the society conforms to the rules of politeness to conduct his business. According to Mandeville, they are similar to fashionable hypocrites, hence deserve to be called "hypocrites" merely because of the inward–outward discrepancy, as explained earlier.[57]

In any event, Runciman ultimately rejects the parallel. Malicious hypocrites deliberately deceive, while fashionable hypocrites are, by definition, victims of self-deception.[58] Thus the two distinctions don't match, because, as just explained, both first- and second-order hypocrites may deceive themselves. As we saw earlier, however, Mandeville explicitly argues that malicious hypocrites are by and large victims of self-deception. This may be supported by a general argument, too.

Runciman further attributes to Mandeville the distinction between hypocritical politicians who occasionally dress up their policies in moral terms and politicians or political observers who accuse other politicians of hypocrisy and immoralism.[59] On the face of it, we have here fashionable hypocrites, who just go through the motions, and malicious hypocrites, whose main weapon is rhetoric built on moralistic censure. The former are men of business and action, the latter are demagogues occupied with relaying their (fake) virtues to the world. As much as it may be difficult to distinguish them with precision, self-deception is hardly the issue that divides them. So, it is unclear what these two categories have to do with the first-/second-order hypocrisies.

In general, therefore, the first-/second-order distinction is an odd lens through which to understand Mandeville's thought.

In his more recent analysis, Robin Douglass assumes the concept of hypocrisy as an inward–outward discrepancy between motives and behaviour.[60] He is then able to claim that hypocrisy, according to Mandeville, "pervades" social life. To have a chance to co-exist peacefully and cooperate with others, we have to mislead them about the persistent self-regarding motives behind our actions; hence our self-presentation is laced with hypocrisy. But if, as I argued, the inward–outward discrepancy is not a sufficient mark of hypocrisy, then this line of argument flounders. We may well be uniformly hypocritical in our self-presentation, but not because of the discrepancy itself. Rather, it is because of the continuity between fashionable hypocrisy and malicious hypocrisy. Hypocrisy involves manipulation of others to your own advantage, and our polite performances are another form of this manipulation.

Douglass also argues that we hold others to a higher moral standard than ourselves (namely, to the standard of "real virtue" I discussed earlier) and that, therefore, there is hypocrisy in our disapproval of others' "failings" to meet the standard of real virtue. I agree with the claim about the higher standard, but what is the argument for this? On one hand, Douglass says that, according to Mandeville, we are "naturally averse" to the self-love and pride of others. On the other hand, he also says that, according to Mandeville, in practice we rarely suspect self-regarding motives in the actions of others. But then why would we pass negative hypocritical judgements on them? If suspicion is indeed rare, we would often judge them "really virtuous" and thus rarely find fault,

hypocritically, with their actions. I agree with Douglass that the different standards by which we judge others and ourselves is an important and interesting claim in Mandeville, but to explain it, as suggested earlier, we have to appeal to the dynamics of self-liking: we necessarily hold others to a higher standard, lest the praise we expect to receive for our actions is devalued.

Finally, according to Douglass, malicious hypocrisy is a rather special case of hypocrisy, characterised by deliberate deception animated by "nefarious motives" to harm others. While I accept the claim about harm, I see malicious hypocrisy as a central and paradigmatic kind of hypocrisy; and I also believe that, in order to classify them both as "hypocrisies," we must stress the continuity between fashionable and malicious hypocrisy in terms of manipulation (though, again, I grant that Mandeville does not describe polite actors as aiming to harm others and that, generally, it is misleading to do so).

4. Conclusion

Let me summarise the main points of this paper. Mandeville offers us an account of the genesis of moral judgement. Self-esteem and sensitivity to praise are at its heart. But this familiar idea, I have suggested, must be coupled with another one: that moral discourse was initially conceived as an instrument of manipulation by free riders. The latter couched their praise in explicitly moral terms to encourage others to benefit society and bear the costs. Themselves, these "very worst" people, were content to pursue their narrow self-interest by enjoying the fruits of other people's efforts. Mandeville is somewhat opaque on the question of whether that is all that there is to morality. Thus e.g. this explanation doesn't cover the notion of "refined" virtue that is pursued for its own sake, nut he is explicit about the very specific purpose of manipulation that moral discourse, at least when used on public occasions, is supposed to serve. This idea may sound surprising or extreme, but it shouldn't be so for Mandeville's readers. After all, even language itself was created with the express purpose of manipulating others.

One implication of this view is that hypocrisy is intrinsic to the public expression of moral beliefs. This form of moral discourse was born in sin: it was created as a weapon of choice of the free riders. It continues to be so. Preaching morality, malicious hypocrites profess to endorse moral norms for the purpose of gaining people's admiration and, with that, more social prestige and possibly even greater political power. Fashionable hypocrites don't normally preach, but in other regards they are similar. Just like malicious hypocrites, they signal to others their virtuous motives in order to advance in society.

As I see it, an interesting novelty of Mandeville's account of hypocrisy is in identifying the figure of a malicious hypocrite. On one hand, as I have argued, he is nothing but the "very worst" person of *EOH* deploying moral judgement strategically for personal gain (including in the cases of self-deception). On the other hand, his hypocrisy is continuous with the hypocrisy of fashionable hypocrites and, indeed, explains why we classify the latter as "hypocrites."

Notes

1. See Lovejoy, *Reflections on Human Nature*, 170.

2. See Monro, *The Ambivalence of Bernard Mandeville*, 149, 157, 25, 46, 47.
3. La Rochefoucauld, *Collected Maxims*, I.151.
4. *FB* I.42ff. *FB* I = Mandeville, *The Fable of the Bees*, Vol. 1. *FB* II = Mandeville, *The Fable of the Bees*, Vol. 2.
5. See Lovejoy, *Reflections on Human Nature*, 176 and references therein.
6. *FB* II.133. This idea of self-love, expressed in pretty much the same terms, is already in the Stoics: Cicero, *De Officiis* I.11. See Cicero, *On Duties*, 6. It should be contrasted with Mandeville's earlier version in *FB* I.41, in which it stands for the satisfaction of desires and obtaining pleasure. Terminologically, "self-esteem" is best thought as a product of the passion of self-liking. An alternative terminology is rooted in a wider historical usage: Maurer, *Self-love, Egoism and the Selfish Hypothesis*, 8–12. However, as explained below, I disagree with Maurer (ibid., 61) that self-liking in *FB* II and Mandeville's *An Enquiry into the Origin of Honour* is best interpreted as "excessive pride." I also believe that self-love in *FB* II is the passion of self-preservation that *may*, under favourable circumstances, be accompanied by pleasure. Maurer has it in reverse: his view corresponds to the ideas in *FB* I. In *FB* II, there is, I think, a distinct change in Mandeville's thinking about self-love.
7. *FB* II, 133.
8. See Hobbes, *Leviathan*, X.16.
9. *FB* II, 136.
10. Here I agree with Kapust, *Flattery and the History of Political Thought*, 141.
11. *FB* I, 42, 261; *EOH* 40. *EOH* = Mandeville, *An Enquiry Into the Origin of Honour*.
12. For Mandeville as a rigorist, see Maurer, *Self-love, Egoism and the Selfish Hypothesis*, 74–5; Kapust, *Flattery and the History of Political Thought*, 140; Scott-Taggart, "Mandeville: Cynic or Fool?" 227; Maxwell, "Ethics and Politics in Mandeville"; 244, Kaye, "Introduction" in Mandeville, *The Fable of the Bees*, I, liii. For Mandeville as an utilitarian see Monro, *The Ambivalence of Bernard Mandeville*, 224; Heath, "Mandeville's Bewitching Engine of Praise," 226.
13. Quotations from the *Fable* follow Mandeville, *The Fable of the Bees*, Vol. 1; Mandeville, *The Fable of the Bees*, Vol. 2.
14. *FB* I, 42–3 (emphasis added; modified for readability).
15. Furthermore, in Mandeville's view, language itself was created for rhetorical purposes, to persuade others of the speaker's high worth, also to threaten them if they refuse to be so persuaded (*FB* II, 289). See Kapust, *Flattery and the History of Political Thought*, 146ff, for an extended discussion. Of course, as a theoretical claim, this is dubious even by Mandeville's own rights. Sensitivity to praise and criticism, the very conceptions of praise and self-esteem, are possible only in a very advanced condition of sociability. Language should have enabled that condition in the first place, with praise, for example, being mediated by linguistic means.
16. See Lovejoy, *Reflections on Human Nature*, 173ff.
17. *FB* I, 47.
18. Cf. *FB* I, 43.
19. There is also the notion of "refined virtue," similarly outside the scope of the Lawgivers' intervention. I return to it below.
20. *FB* I, 48–9 (emphasis added).
21. An anonymous referee raised this question.
22. See e.g. Runciman, *Political Hypocrisy*, 49; Heath, "Mandeville's Bewitching Engine of Praise," 210; Douglass, "Mandeville on the Origins of Virtue," 281.
23. See Maurer, *Self-love, Egoism and the Selfish Hypothesis*, 70–6.
24. See Douglass, "Mandeville on the Origins of Virtue," 285; Douglass, "Bernard Mandeville on the Use and Abuse of Hypocrisy," 5–6.
25. Douglass "Bernard Mandeville on the Use and Abuse of Hypocrisy," 5–6.
26. See Monro, *The Ambivalence of Bernard Mandeville*, 189–90. See also Lovejoy's comment on "racial superiority" above.
27. As observed already by the Stoics: Cicero, *De officiis* I.11.

28. *FB* I, 57.
29. See Mandeville, *An Enquiry Into the Origin of Honour*, 201–2.
30. Not to be confused with the semantic reductionism of (Good) and (Bad) described in Section 2.
31. *FB* I, 173–81.
32. Ibid., 57.
33. *FB* II, 139.
34. The situation described here conforms to the extant interpretations of self-deception. See e.g. von Hippel and Trivers, "The Evolution and Psychology of Self-Deception," 4; Trivers, *The Folly of Fools*, 9; Bach, "Self-Deception." Again, as in other cases, we need not assume that the agents have a conscious, explicit belief that they try to suppress.
35. See Szabados, "Hypocrisy." As Szabados also shows, Mandeville is in good company: a similar error was committed in Ryle, *The Concept of Mind*.
36. The same problem vitiates, e.g. the concept of hypocrisy offered in Hundert, *The Enlightenment's Fable*, 178.
37. *EOH*, 34–5 (emphasis added).
38. See the quote from *FB* I, 48–9, above.
39. See also the remark on "good Hypocrites" in *EOH*, 208.
40. *EOH*, 200–1.
41. See Whyman, *Stanislavski: The Basics*.
42. See von Hippel and Trivers, "The Evolution and Psychology of Self-Deception," 4–5. By the same token, influential modern accounts of hypocritical criticism that correspond to Mandeville's "malicious hypocrisy" trace its moral violation to self-deception and to the failure to examine one's own character and past actions. See Wallace, "Hypocrisy, Moral Address, and the Equal Standing of Persons"; Dover, "The Walk and the Talk."
43. *EOH*, 230–1.
44. *FB* II, 80.
45. *EOH*, 231. One might complain that this is still an oversimplified picture of a hypocrite. At some stage, Mandeville concedes, Cromwell *sincerely* believed his own rhetoric, yet he still had immoral goals and "blackest designs" (ibid., 231). But has he not, at this later stage, morphed into a single-minded fanatic, though he may have started as a hypocrite? This issue also bears on the historical accuracy of Mandeville's description of Cromwell, who, we are told, "never wavered or hesitated" throughout the king's trial (Firth, *Oliver Cromwell and the Rule of the Puritans in England*, 231). For changing perspectives on Cromwell, see Davis, *Oliver Cromwell*, 44–64.
46. *EOH*, 202.
47. The same distinction in the same terms is in Ryle, *The Concept of Mind*, 155; that is, oddly, not mentioned by Runciman.
48. See Runciman, *Political Hypocrisy*, 53–4, 70.
49. See Kant, *Anthropology from a Pragmatic Point of View*, 264.
50. Ibid., 51. In earlier times, religiosity could play the role of politeness, as Mandeville observes in *EOH*, 232.
51. See Runciman, *Political Hypocrisy*, 51.
52. See Runciman, *Political Hypocrisy*, 52.
53. See Runciman, *Political Hypocrisy*, 54.
54. See Erasmus, "On Good Manners for Boys."
55. Runciman, *Political Hypocrisy*, 52.
56. Ibid., 57.
57. See the quote from *EOH*, 201–2 above and the discussion there.
58. Runciman, *Political Hypocrisy*, 58, 63.
59. Ibid., 65.
60. See Douglass, "Bernard Mandeville on the Use and Abuse of Hypocrisy," 2.

Acknowledgments

I am grateful to three anonymous referees for useful comments on an earlier draft of this paper. I also wish to thank Andreas Blank for his generous help and encouragement.

Notes on contributor

Y. Sandy Berkovski is Assistant Professor of Philosophy in Bilkent University. His research interests straddle the issues in philosophy of language, moral psychology, and philosophy of science.

ORCID

Sandy Berkovski http://orcid.org/0000-0001-8926-5758

Bibliography

Bach, K. "Self-Deception." In *The Oxford Handbook of Philosophy of Mind*, edited by A. Beckermann, B. P. McLaughlin, and S. Walter (pp. 781–796). New York: Oxford University Press, 2009.

Cicero. *On Duties*. Edited by M. T. Griffin and E. M. Adkins. Cambridge: Cambridge University Press, 1991.

Davis, J. C. *Oliver Cromwell*. New York: Oxford University Press, 2001.

Douglass, R. "Bernard Mandeville on the Use and Abuse of Hypocrisy." *Political Studies*, 1–18. doi:10.1177/0032321720972617.

Douglass, R. "Mandeville on the Origins of Virtue." *British Journal for the History of Philosophy* 28, no. 2 (2020): 276–295.

Dover, D. "The Walk and the Talk." *Philosophical Review* 128, no. 4 (2019): 387–422.

Erasmus, D. "On Good Manners for Boys." [1530]. In *Collected Works of Erasmus*, edited by J. K. Sowards (pp. 269–289). Vol. 25–26. Toronto: University of Toronto Press, 1985.

Firth, C. H. *Oliver Cromwell and the Rule of the Puritans in England*. London: G. P. Putnam's Sons, 1900.

Heath, E. "Mandeville's Bewitching Engine of Praise." *History of Philosophy Quarterly* 15, no. 2 (1998): 205–226.

Hobbes, T. *Leviathan*. Edited by C. B. Macpherson. London: Penguin Books, 1981.

Hundert, E. G. *The Enlightenment's Fable*. Cambridge: Cambridge University Press, 1994.

Kant, I. "Anthropology from a Pragmatic Point of View." In *Anthropology, History and Education*, edited by G. Zöller, and R. B. Louden. Cambridge: Cambridge University Press, 2007.

Kapust, D. J. *Flattery and the History of Political Thought*. Cambridge: Cambridge University Press, 2018.

La Rochefoucauld, F. *Collected Maxims and Other Reflections*. New York: Oxford University Press, 2007.

Lovejoy, A. O. *Reflections on Human Nature*. Baltimore: The Johns Hopkins Press, 1961.

Mandeville, B. *An Enquiry Into the Origin of Honour*. London: Frank Cass & Co, 1971. [= *EOH*].

Mandeville, B. *The Fable of the Bees*. Vol. 1. Oxford: Clarendon Press, 1924. [= *FB* I].

Mandeville, B. *The Fable of the Bees*. Vol. 2. Oxford: Clarendon Press, 1924. [= *FB* II].

Maurer, C. *Self-Love, Egoism and the Selfish Hypothesis*. Edinburgh: Edinburgh University Press, 2019.

Maxwell, J. C. "Ethics and Politics in Mandeville." *Philosophy (London, England)* 26, no. 98 (1951): 242–252.

Monro, D. H. *The Ambivalence of Bernard Mandeville*. Oxford: Clarendon Press, 1975.

Runciman, D. *Political Hypocrisy*. Princeton: Princeton University Press, 2008.

Ryle, G. *The Concept of Mind*. Abingdon: Routledge, 2009.

Scott-Taggart, M. J. "Mandeville: Cynic or Fool?" *Philosophical Quarterly* 16, no. 64 (1966): 221–232.
Szabados, B. "Hypocrisy." *Canadian Journal of Philosophy* 9, no. 2 (1979): 195–210.
Trivers, R. *The Folly of Fools*. New York: Basic Books, 2011.
von Hippel, W., and R. Trivers. "The Evolution and Psychology of Self-Deception." *Behavioral and Brain Sciences* 34 (2011): 1–56.
Wallace, R. J. "Hypocrisy, Moral Address, and the Equal Standing of Persons." *Philosophy & Public Affairs* 38, no. 4 (2010): 307–341.
Whyman, R. *Stanislavski: The Basics*. Abingdon: Routledge, 2013.