

This document is the Author Original Manuscript.

Please cite the publisher's Version of Record:

S. Field (2014). 'Hobbes and human irrationality', *Global Discourse: An Interdisciplinary Journal of Current Affairs and Applied Contemporary Thought*, 5 (2), 207-220. Published with reply: Luke O'Sullivan, 'Sovereigns and citizens: a response to Field', *ibid.*, 221-224.

Reprinted as:

S Field (2015). 'Hobbes and human irrationality', in T. Nardin (ed.) *Rationality in Politics and its Limits*, Abingdon: Routledge, 31-44. With reply: Luke O'Sullivan, 'Sovereigns and citizens: a response to Field', *ibid.*, 45-48.

Hobbes and human irrationality

Sandra Field*

Yale-NUS College, Singapore

Hobbes's science of politics rests on a dual analysis of human beings: humans as complex material bodies in a network of mechanical forces, prone to passions and irrationality; and humans as subjects of right and obligation, morally exhortable by appeal to the standards of reason. The science of politics proposes an absolutist model of politics. If this proposal is not to be idle utopianism, the enduring functioning of the model needs to be compatible with the materialist analysis of human behaviour. In this paper, I argue that Hobbes's attempts to render his science of politics compatible with his materialism are only partly successful; a fuller compatibility is achieved in the political writings of Spinoza.

Keywords: Hobbes; Spinoza; materialism; reason; sovereign; peace

Hobbes's science of politics rests on a dual analysis of human beings: humans as complex material bodies in a network of mechanical forces, prone to passions and irrationality; and humans as subjects of right and obligation, morally exhortable by appeal to the standards of reason. Traditionally the secondary literature has displayed a greater interest in the latter analysis (Warrender 1957; Nagel 1959; Kavka 1986). However, this paper finds the former is equally worthy of attention. Specifically, I

* E-mail: sandra.field@yale-nus.edu.sg

claim that for Hobbes's science of politics to succeed, it must be compatible with the analysis of human beings as potentially irrational material bodies. This compatibility cannot be taken for granted: I argue that although Hobbes does attempt to constrain his science of politics by the results of his material analysis of human beings, he is only partly successful. I argue that critically examining the determination of political actors' behaviour in a web of causal relations should lead to a serious reconsideration of the conclusions of Hobbes's science of politics; a reconsideration carried out not by Hobbes within his own *oeuvre*, but rather in the work of Benedict de Spinoza.

§1

In this first section, I seek to establish a Hobbesian criterion for a successful science of politics.¹ This criterion will involve some reference to Hobbes's science of human beings, but determining this relation precisely is complicated by the fact, as I foreshadowed in my introductory remarks, that there are two different parallel analyses within the science of human beings.²

¹ For Hobbes, the terms 'philosophy' and 'science' are interchangeable (L ix.1 and table). Hence, his science of politics is equivalent to his civil philosophy.

I have used the following abbreviations for Hobbes's texts, derived from the texts' original titles, and I list them in an order corresponding to their original composition:

EL = *The Elements of Law* (Hobbes 1969)

DC = *De Cive* (Hobbes 2006)

L = *Leviathan* (Hobbes 1994)

DCo = *De Corpore* (Hobbes 1656)

DH = *De Homine* (Hobbes 1972)

I have avoided using the Molesworth edition of Hobbes's writings, given the inadequacies of those texts. See the discussions in Tuck (2006, xlviii) and Curley (1994, lxxi).

² Although Hobbes's language is gendered, his conceptual analysis is to a remarkable degree ungendered, particularly by comparison with other canonical figures from the history of political thought – for instance, consider Hobbes's refusal to naturalize the authority of men over women (L xx.4–5). Correspondingly, I have sought to reformulate Hobbes's arguments in gender neutral terms. This results in the infelicity of replacing Hobbes's 'science of man' with the less elegant 'science of human beings' or 'science of the human body'.

I start by laying out three characteristics of Hobbesian science in general: its materialism, its speculative character, and its practical orientation. Hobbes's science is materialist: he claims that there is in the universe nothing but body and its accidents, and the motion of bodies (DCo i.8, vi.5–6). Furthermore, there can be no cause of motion in a body except another 'Body Contiguous, and Moved' (DCo ix.7). As a corollary, human beings lack free will. This is because the mental processes of reflection and decision which lead to human actions are consequences of mechanical causes; they are 'motions of the mind' (DCo vi.6) following the interaction between the organs of sense and imagination and the material impetus they encounter. If materials forces are sufficient to generate a certain mental state which results in a bodily action, then the person does not have some immaterial power to resist this determination (L vi.1, xxi.4). For all motions that occur are determined to occur by prior motions; all events are linked to prior events by chains of mechanical causation. This means that propositions concerning future events are either necessarily true or false; if we tend to speak of future events as contingent, this is merely a reflection of our ignorance regarding the causes already in train (DCo ix.6, x.4–5).

Hobbes's science is speculative. For him, the knowledge derived from sense and from experience, and even from experimentation, fails to count as science; that knowledge merely reports the fact of what has occurred, and the prudence that builds on it only fallibly projects the tendencies of the past into the future. Such knowledge is useful and necessary to science, but nonetheless distinguishable from it. Science, by contrast, establishes definitions that grasp the generation (causes) and effects (consequences) of bodies according to their nature. With these definitions in hand, the truths of science are deductive, taking the conditional form, 'if this be x, then y follows' (EL I.iv.11, I.vi.4; L iii.7, v.1-2, v.17, ix.1; DCo i.2, vi.1, vi.13; Anstey 2005, 215; Johnston 1986, 51–52; Shapin, Schaffer, and Hobbes 1985, 7). Because of his speculative conception of science, there is an important methodological difference between Hobbes's sciences of two kinds of bodies:

For two chief kinds of Bodies, and very different from one another, offer themselves to those who search after their Generation & Properties; One whereof being the work of Nature, is called a *Naturall Body*; the other is called a *Commonwealth*, and is made by the wills and agreements of men.

And from these spring the two parts of Philosophy called *Naturall* and *Civill*.
(DCo i.9)

For natural bodies, science cannot simply stipulate its own definitions, but rather it needs to grasp principles ‘placed in the things themselves by the Authour of nature’ (DCo xxv.1), and via grasp of these principles it needs to reconstruct the bodies’ causes and investigate their effects. By contrast, with artificial bodies, the principle of the body lies in the intention or definition in accord with which humans attempt to produce it (DH x.4–5). The core definition of the artificial body of the commonwealth is an association for peaceful coexistence; an association for getting ‘out of that miserable condition of war’ (L xvii.1). Thus the science of politics needs first of all to seek the canonical causes of such a body. Frequently the actual commonwealth will not properly live up to this definition, and the intentions of its actual makers may even diverge from the canonical definition (for instance, they may be driven by glory rather than peace); but only secondarily does the science attempt to say what constitutes the actual inadequate and often war-torn commonwealths before us (L xxx.5). The logic of this science of artificial bodies might be clarified by comparison with geometry (DH x.5). To understand a circle, we need to understand it as a figure whose circumference is traceable by points equidistant from a central point; the science of circles is not first of all the science of actual imperfect circles in their imperfection.

Finally, Hobbes has a practical conception of science.

The *End* or *Scope* of Philosophy, is, that we may make use to our benefit of effects formerly seen, or that by application of Bodies to one another, we may produce the like effects of those we conceive in our minde, as far forth as matter, strength & industry will permit, for the commodity of human life. ... the scope of all speculation is the performing of some action, or some thing to be done. (DCo i.6; see also L xlvi.1)

Thus, even though a science may be trivially true when it states consequences of a definition, the science is not substantially successful if there is no way of relating it to the practical world of bodies within which humans live. Specifically, Hobbes claims

that a proper science of politics will explain the causes of war in such a way that human industry can in the future prevent war (DCo i.7; Johnston 1986, 53).

I now turn to consider the science of human beings in detail. I claim that Hobbes offers two complementary scientific analyses of human beings: the first corresponds to the mechanistic causes and consequences of human beings as natural bodies, whereas the second treats the rational or moral consequences of those natural bodies (Johnston 1986, 45–46; Matheron 1986, 77–78). Hobbes cites Harvey’s discovery of the science of the circulation of the blood as the start of the science of the human body (DCo Epis.13). He begins his own ‘science of man’ by analysing the human body’s causes and effects in a network of material forces. He builds a systematic mechanistic account of the body’s function and effects, starting with optics and perception; then imagination and mental functioning; finally treating the genesis of passions, dispositions, and judgement (EL I.ii-iv, vii-x; DC i.1–6; L i-iii, viii, x–xiii; DH xi–xiii). It is a science of how human bodies for the most part behave; he characterises it as a study of ‘man as matter’ (L Intro.2–4). The substantive result of this science is the identification of some generally shared features of our material constitution, which he calls human nature. He claims that ‘of the voluntary acts of every man the object is some *good to himself*’. (L xiv.8) This is not yet a substantial uniformity, because ‘good’ means simply whatever a given person desires, and human desires vary widely, according to their constitution and experience (L vi.7). Nonetheless Hobbes claims there is one desire that human beings do share, one baseline preference ordering: a preference in favour of self-preservation (L xv.17), and specifically, avoidance of violent death. It is an ‘*absolutely certain postulat[e] of human nature*’ that ‘*each man strives to avoid violent death as the supreme evil in nature*’. (DC Epis.10; see also L xiv.29)

It is important to see that Hobbes makes no general attribution of rationality to human nature, not even of a self-interested economic variety. The human nature he identifies is merely a basic proto-rational passion, and the ‘natural reason’ belonging to human nature (DC Epis.10) is simply an immediate aversion, with a minimal associated elementary causal reasoning about how to satisfy it. No capacity is postulated to see or act on one’s own long-term interest, nor is there any claim that humans by nature pursue self-preservation according to means that will be effective. To the contrary, that pursuit tends to be limited by inadequate knowledge and

judgement, and derailed by pressing passions (DC iii.26; L viii, xvii.2). This is not to deny that humans sometimes do display reason. In the absence of free will, the passions opposing reason cannot simply be wished away, but sometimes they are materially overcome. Sometimes there is an unusual convergence of causes that gives rise to an exceptional individual who is durably rational, or to a moment of clarity in someone not otherwise durably rational. And sometimes structural factors can conduce to a greater degree of rationality in a population, for instance being in a stable commonwealth where social circumstances do not inflame the passions, and where careful education cultivates clear and correct thinking.

At a certain point in Hobbes's science of human beings, he shifts to consider the consequences of the human body in terms of reason rather than mechanics (DC i.7; L xiv). These two analyses are distinguished, although Hobbes does not mark the distinction as being particularly significant. Rather than showing how humans in fact tend to behave, Hobbes now outlines how reason dictates that humans ought to behave. This second science of the human body generates natural rights, and the duties of the laws of nature (EL xv.1; DC ii.1, iii.33; L xiv.3, xv.41). Hobbes starts from the premise that humans by their nature are motivated to work for their self-preservation and to avoid violent death. He stipulates that humans are not only entitled to pursue those motivations, but also that they ought to pursue them in accordance with correct reasoning. Hobbes argues that the first dictate of reason is to seek peace. He then builds up increasingly detailed and specific requirements on conduct, especially regarding covenants and agreements (EL xiv.14, xv-xvii; DC i.15, ii-iii; L xiv.4-33, xv).

This second science of human beings is prescriptive, not explanatory. To be sure, Hobbes is not totally pessimistic regarding the possibility of a population abiding by the laws of nature. Although reason *tout court* may be very rare to obtain (L v.17), reason enough to grasp the law of nature is more widely accessible, if preached and taught appropriately (DC iii.26; L xxx.6). Yet achieving and putting into play this rational acceptance of the law of nature is still not straightforward. When people attempt to determine the requirements of the law of nature, they often get it wrong (DC ii.1); for instance, think of Hobbes's infamous fool who believes it is acceptable to break promises whenever it suits her (L xv.4). And even if the understanding of the law of nature were correct, the problem of motivation would

remain. Hobbes insists that the laws of nature are contrary to humans' natural passions and are not generally followed simply for their own sakes (DC iii.27; L xiv.31, xvii.2). Even though the second science of human beings speaks in detail about rational conduct, that conduct is not readily or automatically obtained.

The science of politics is created both by and out of humans, and consequently needs to be built upon the science of human bodies (DC Pref.9; L Intro.2; DCo i.9). I take no position on the question of whether Hobbes's science of natural bodies entails his science of politics. (On that topic, see Watkins [1989].) Instead, I simply attempt to make precise how that former science might constrain the latter. I rule out two extreme possibilities in order to establish more clearly the account that I favour. A first attempt to specify the relation might claim that the science of politics is simply an extension of the first science of human beings: just as that science considers human matter in terms of the mechanical causes of its actual existence, including its imperfect rationality, so too a science of politics needs to consider the commonwealth as the entity it actually is, in all its malfunctions. But this proposal must be rejected, because it fails to appreciate the distinction between sciences of natural and of artificial bodies. As outlined earlier, the science of the artificial body of the commonwealth must first of all be a science of the ideal body which achieves the goal of peace, not of the imperfect body that we see before us. Indeed, *Leviathan's* chapter on the 'causes, generation, and definition' of the commonwealth starts by articulating '[t]he final cause, end, or design for which the commonwealth is established, namely peace' (L xvii.1; see also DC Pref.14–15).

A second and opposite attempt to specify the relation between the science of politics and the science of human beings might claim that the science of politics is simply an extension of the science of human beings as rational actors. For if we are considering an ideal commonwealth that is stably peaceful, why not also consider ideal causes – that is, rational human beings? Indeed, Hobbes sometimes characterises the science of politics simply as a science of rights and duties, and of the rule of just and unjust (DC Pref.1; L ix table; DCo i.7, i.9). But I claim that whilst the science of politics may include this, it needs also to include a model of politics that can support the outcome of an enduringly peaceful commonwealth, given what we know of actual non-ideal human nature. For without such a model, it fails the practicality requirement that I outlined earlier: that it provide action guidance to achieve a

humanly important purpose. Articulating to everyone their duties does not secure the practical purpose of achieving a peaceful commonwealth if we have reason to believe that it is implausible that people will live up to these duties. For this reason, the science of politics needs to build on the science of human matter (L Intro.2-4); it needs to be grounded in ‘the Appetites of Men, and the Passions of their Minds’ (DCo iv.7); it cannot simply rely on words and promises (L xlvi.35–36). Specifically, a model of politics which for its functioning relies heavily on a population of humans who are already rational would fail the test of practicality, because the science of human beings as material bodies tells us that it only even starts to be plausible that a population should be rational after a political order has been put in place to moderate the passions.

Between these two mistaken extremes, the required constraint that the sciences of humans beings place on the science of politics emerges, and I will articulate it via three characteristics. First, the science of politics is not simply the science of the body of the commonwealth as it materially exists. Rather, it is first and foremost a science of an intended goal: a stably peaceful society. Part of its task is to put forward a model of such a society. Second, the science of human beings as moral subjects of rational duty serves as a limiting criterion for that model: the constraint on subjects proposed in the model of politics should be rationally justifiable, such that rational subjects ought to acquiesce to it. Third, the science of human beings as material bodies also serves as a limiting criterion for that model: the model of politics must be such that it can plausibly peacefully endure, when populated not by perfect moral actors but by ordinary material humans. All these three characteristics are important, but for the purposes of this paper, I now focus only on the final one: whether the proposed model will materially and plausibly succeed in securing an enduringly peaceful commonwealth.

§2

Hobbes’s science of politics makes efforts to comply with this material criterion. In this section I distinguish two aspects of that compliance: the compliance of the modelling of subject behaviour, and the compliance of the modelling of sovereign behaviour. I argue that Hobbes is successful in respect of subjects but unsuccessful in respect of sovereigns. I will focus my discussion on the case of a sovereign who is a

monarch rather than an aristocratic or democratic assembly, both for simplicity's sake, but also because it is the canonical case in which Hobbes's science should work most straightforwardly.

The goal of Hobbes's science of politics is a peaceful political order (DC i.1; L xvii.1). An interrogation of Hobbes's political theory from the point of view of right or rationality might find it adequate simply to consider the covenantal moment of the establishment of the commonwealth. Hobbes offers two alternative canonical causal geneses: sovereignty by institution and sovereignty by acquisition. Subjects' fear of each other establishes a sovereign by institution: subjects mutually agree to defer to a single ruler, not because that ruler is antecedently frightening to them, but because they hope that the ruler will prevent them from harming one another. Subjects' fear of a powerful conqueror establishes a sovereign by acquisition: subjects each defer to a conqueror out of direct fear of harm at the conqueror's hands. Regardless of which genesis is followed, a covenant is argued to establish absolute right for the sovereign, and near-absolute concession of right from subjects (DC v.12, viii.1; L xvii.15, xviii.1, xx.1–3). This right, once established, remains in force unless the sovereign loses its power: for natural law stipulates that valid covenants be kept (L xv.1–3). The sovereign has no obligation to do or to avoid any particular conduct, although if his or her conduct does lead to a loss of power then loss of right follows (L xviii.8, xxix.23).

However, the focus of my inquiry is not the situation of right, but rather the material constitution of the stably peaceful commonwealth; the material requirements to obtain obedient behaviour sufficient to secure peace. Again, I start by considering the establishment of sovereign power, and it appears that the covenantal moment by itself satisfies the material success requirement on the model of the commonwealth. By either genesis, the sovereign 'hath the use of so much power and strength conferred on him that *by terror thereof he is enabled to conform the wills of them all to peace at home and mutual aid against their enemies abroad*' (L xvii.13, emphasis added). The initial fact of mass obedience, once obtained even if just for a moment, constitutes a power for the sovereign. It gives the sovereign a great number of hands to deploy as it pleases. This then also generates ongoing power, because the initial compliance provides informants, guards, judges and executioners: in other words, it gives the sovereign a great capacity to police and punish disobedience. In the face of

this punitive capacity, there is strong pressure for each individual subject to obey. But this secondary obedience in turn deepens and reinforces the sovereign's capacity to deploy subjects' powers for its own purposes. Thus so long as power is established at one point in time, it appears the sovereign can behave as it pleases without disrupting its effective power (DC v.8; L xvii.13).

Yet on closer inspection, this mechanism proves to be seriously flawed. Hobbes acknowledges that the sovereign needs constant vigilance and needs to behave in a very particular way in order to secure peace. 'Many things are required to preserve internal peace, because ... many things conspire to disturb it'. (DC xiii.9) This is because even an overwhelming punitive incentive of violent death turns out not to be sufficient to secure widespread obedience. There is no challenge to the idea that people aim for a good to themselves, yet nonetheless their seditious passions can win out when their perceptions of good and bad are distorted. First, people may have theological reasons not to consider death as the greatest of evils. People can come to believe that eternal torment after death is worse than death itself, and that the point of law of nature is not life on earth but life after death (L xv.8). Second, even when people still fear violent death above all other things they may subscribe to a flawed doctrine of politics or have a flawed perception of politics which prevents them from always seeing sufficiently directly the link between obedience to the sovereign and avoiding death (EL II.viii.4–10; DC Pref.5–6, xii.1–8; L xviii.20, xxix). These distorted beliefs amplify the seditious passions which tend to simmer amongst the populace: the combination of the vainglory and ambition of the eminent, and the poverty and disaffection of the lowly (EL II.viii.1-3; DC xii.9-12). Discontent is dangerous: anyone 'not contented with their present condition' may be prone to 'stir up trouble and sedition' (L xi.4) out of hope to establish a political order in which they have a better life (Johnston 1986, 68–69, 93–94, 120–123).

Indeed, once matters reach this point there is no easy solution: simply heightening punitive measures risks provoking the rebellion it aims to quell. Hobbes instead recognises that the sovereign's achieving its ends without incurring the risk of rebellion requires a proliferation of carefully calibrated modes of rule, meticulously and consistently implemented. The suggestions Hobbes offers fall into four main categories. First, teach the correct doctrine of right, at the pulpits and the universities (EL II.ix.8; DC xiii.9; L xxx.4–14). Second, intervene in theological debates to re-

channel fear of the afterlife back into fears of this world (EL II.vi; L xxxviii). In these first two suggestions, both the doctrine of duty and theological doctrine are considered not merely in their truth but also as material factors in the model of the commonwealth: promulgating them is hoped to be effective in shaping conduct in a way that supports civil order (DC Pref.20–21). Third, eliminate subjects' disaffection, by being sure always to rule in a scrupulously transparent, fair, measured, and equitable manner, by helping those in material want, and by ensuring all public officials similarly abide by such high standards of behaviour (EL II.ix.3–6; DC xiii.10–11, xiii.17; L xxx.15–18, xxx.20–23). Fourth, deflate the emergence of other powers, especially those of subjects who show signs of ambition rather than loyalty (EL II.ix.7; DC xiii.12; L xxx.16, xxx.24). In sum, the sovereign must always rule according to exacting standards of reason and self-restraint if the political order is to stave off war and disorder. (See also Baumgold [1988, 101–109], Frost [2008, 156–165] and Tarlton [1978, 321–327].)

To this point, I have considered the problem of bringing subjects to behave in the manner required for Hobbes's model of politics to be plausible. Hobbes's solution relies on prudent rule by the sovereign. I now turn to consider the sovereign's own behaviour from the perspective of the material science of human bodies. In the Introduction to *Leviathan*, Hobbes notoriously compares the creation of the commonwealth by human beings to God's creation of man (L Intro.1–4). The hubris of this comparison sits uncomfortably with Hobbes's admission that human beings are not gods, but to the contrary they are always simultaneously matter, parts of God's creation, and weak and irrationality-prone parts at that. What is the relationship between the image just sketched of a sovereign who wisely and reliably behaves in the way required to sustain the civil order, and the image derived from Hobbes's science of human beings as matter, applicable to the sovereign as to any other human, which shows humans to be prone to irrationality and short-sightedness? Indeed, this is a topic that has not been broached in the contemporary literature. Even those reconstructions of Hobbes's political theory which are very attentive to material considerations with respect to subjects offer no account of how a sovereign might

have the wisdom or motivation sufficient to this end (Baumgold 1988, 118–119; Frost 2008, 156–165; Tarlton 1978, 321–327).³

Hobbes offers two answers, but I will argue that neither one is satisfactory. The first answer from the text is seen in Hobbes's recommendation that we evaluate his science of politics presuming this rational sovereign behaviour, and that we place the wicked conduct of the sovereign exterior to the model. Tyrannical rule 'is in fact in every kind of commonwealth a great disadvantage where it occurs (it is the occurrence that is the disadvantage, not the possibility that it may occur), but the fault is the Ruler's, not the Régime's' (DC x.17; see also DC vii.4). Tyrannical rule is not the fault of the model of politics, but of those 'who misuse their authority for their own greed when they are appointed to such a position of power' (DC vi.17). Hobbes does not elaborate why the bad conduct of the sovereign is irrelevant to the evaluation of his model of politics, but I reconstruct the logic by an appeal to the character of the science of politics as a science of artificial bodies. Recall that such a science identifies the principles governing the body under consideration not by appeal to its existence as a natural object, but by appeal to the artificer's intention. But surely this means that the artificer is placed external to the body being made. If we understand the science of politics as addressed to the sovereign as artificer – and indeed, the text seems to operate with presumption that the sovereign, 'he that is to govern a whole nation' (L Intro.4), is the artificer⁴ – then the statement of the sovereign's actions is not a statement of how the sovereign is likely to behave. Rather, it is a statement of advice to the sovereign as an artificer who stands prior to and apart from the body being artificed. In other words, Hobbes's science of politics, as it applies to the sovereign's conduct, is akin to the older genre of 'mirror for princes': it is not nor should it be

³ It is true that the Hobbesian sovereign is obliged to God to rule prudently (see L xxx.1). But this does not address my worry regarding material causality. Sovereigns may well have such an obligation to God, but the question is whether they will have the wisdom or motivation to fulfil this obligation.

⁴ Matheron (1986, 77–78) also argues from a Spinozist perspective that having human beings modelled simultaneously as matter and artificer is a problem for Hobbes's system, but he reads the problematic duality as applying at the level of political subjects. To the contrary, I have claimed that the problem at the subject level is resolvable; rather, the intractable difficulty arises when the dual modelling is applied to the sovereign.

judged as a material science of the sovereign's behaviour. To the contrary, the science of politics should be understood as a doctrine of rationality, obligation, or duty of the sovereign, given his or her commitment to artificing the kind of body that is the commonwealth. The merit of a political model should be judged as it would be if the rulers were prudent and good; the possible wickedness of a ruler in no way impugns the model because the ruler ought not have behaved in that way, and in a certain sense is not part of the model.

I argue, however, that this first answer alone is entirely inadequate. The sovereign is not an external artificer who acts at a point in time, separable from the artificed body, like a transcendent God who creates the world in the course of a week. Rather, as I have demonstrated, sovereignty is an ongoing active role within the model without which the model doesn't function. It is this temporal extension which makes it necessary to consider the sovereign as an element within the model of politics, subject to the material plausibility requirement. To put this another way, suppose the doctrine of the sovereign's rational duty could not be complemented and undergirded by a material account of how the conduct required of the sovereign might plausibly be achieved. This amounts to supposing that it is predictable that a part of Hobbes's scientific model will not be able to do what is required for the model to function. In this case, the political model would fail Hobbes's practicality requirement.

This brings me to Hobbes's second answer. Although Hobbes's texts do not give nearly as much attention to the determinate causes of behaviour of rulers as they do to subjects, and even though his absolutism does not place any specifically crafted external constraint on the sovereign (as it did on subjects) to bring its behaviour in line with what the model requires, nonetheless Hobbes does make some effort to demonstrate the compatibility of the pattern of conduct he is requiring from the sovereign with an understanding of the sovereign itself as just another weakness-prone human body. Hobbes himself implicitly recognises the need to account for the plausibility of the modelled sovereign behaviour: he analyses the 'aptitude' of the various forms of sovereignty (monarchy, aristocracy, democracy) 'to produce the peace and security of the people' (L xix.4). He proposes that the rationality that is presumed will intrinsically be produced because of the political structure. This sketch of sovereign motivation is provided in two parts; however, the first part fails and the second part is weak.

The first part of the sketch of sovereign motivation consists in the frequently made claim that the sovereign will rule wisely because of the coincidence of interests and power between the sovereign and the subjects. The ‘strength and glory’ of the sovereign consists in the ‘vigour’ of its subjects (L xviii.20). Consequently, sovereigns will not rule tyrannically because they harm themselves by doing so, and even risk bringing about their own downfall. Will the sovereign kill its subjects to seize their wealth? Will the sovereign rob and kill some subjects in order to curry favour with others? Hobbes denies that the sovereign is likely to desire to do these things: ‘there is no reason why he would want to spoil his citizens, since that is not to his advantage’. (DC vi.13; see also EL II.v.1; DC xiii.2; L xviii.20, xxx.21) Commentators such as Pettit (2008, 128–129) often accept these claims at face value.⁵

Unfortunately, taken by itself this sketch fails. Ruling well may indeed aid the sovereign in avoiding the downfall of the regime in the long term; but the crucial question will be whether it is plausible that this fact will provide sufficient motivation to govern the sovereign’s conduct: whether the risk of downfall constitutes a sufficiently strong motive to outweigh the other and more immediate determinants of the sovereign’s action (passions and incentives). The *prima facie* reason to doubt its sufficiency is the highly exacting standards of rationality and restraint that the model of politics demands of the sovereign, as I laid out earlier. Indeed, in other contexts Hobbes is acutely aware of a slippage between true interest and actual motivation. ‘Men cannot divest themselves of the irrational desire to reject future goods for the sake of present goods (which inevitably entail unexpected evils)’. (DC iii.32; see also L xviii.20) I have shown that in the case of political subjects, subjects’ rationality is not relied upon for the sustainability of the model. The reason for this was that the material science of human beings shows that exhortation to behave rationally, without strong institutional, material, and punitive support, is insufficient for vast majority of humans. A similar principle must apply to the sovereign.

The second part of Hobbes’s sketch of sovereign motivation is more promising. He proposes that certain aspects of the sovereign’s social position generate immediate pressures that will reliably produce the requisite rational behaviour, even

⁵ This easy acceptance is especially curious against the backdrop of Pettit’s other writings (Pettit 1999), which insist on the need to keep rulers in check, precisely because their interests diverge from those of the people. I thank Geneviève Rousselière for this point.

without presupposing some strong intrinsic rationality. Hobbes observes that sovereigns are not only sovereigns, but also private individuals.

[W]hosoever beareth the person of the people, or is one of that assembly that bears it, beareth also his own natural person. And though he be careful in his politic person to procure the common interest, yet he is more (or no less) careful to procure the private good of himself, his family, kindred and friends, and for the most part if the public interest chance to cross the private, he prefers the private; for the passions of men are commonly more potent than their reason. (L xix.4)

The private passions are more immediate than the public, presumably because they tie in with the immediate affective pressures of bodily enjoyment and face-to-face social interactions. Nonetheless, Hobbes claims that in fact these align with public duty, and there are no systematic incentives for ruling badly. He considers the monarch's desire for personal glory and his or her proclivity to favour friends and family. Both are fairly immediate and pressing passions; and Hobbes claims that the first aligns with the common good, and the second is unlikely to cause harm. The monarch finds personal glory in the wealth of his or her nation: and he or she only has a finite circle of friends on which to bestow gifts, so there will be a limit to how much wealth will be diverted from the public good (EL II.v.5–7; DC vi.13, x.6–7, x.18; L xix.4–9).

Hobbes's speculative claims regarding how a sovereign is likely to behave are not entirely unconvincing. But once we weight them against alternative speculative scenarios, they are shown to be implausible. I will now consider briefly two sources from the history of political thought that provide alternative accounts of the pressures on the absolute sovereign. First, in Machiavelli's analysis, the structure of immediate incentives in absolutist political orders is not argued to be conducive to peace; it is neutral. But this leaves wide open the path of imprudent and destabilising rule by an incautious sovereign. Machiavelli concedes that perhaps there are some people of proven, strikingly just dispositions who might exercise restraint in their role as absolute sovereign. However, he observes that in absolute monarchy, rule tends to become hereditary, and as soon as it does so, the privileged and pampered upbringing of the younger generation in their personal affairs leads to sumptuous and lascivious rule, with new monarchs spending money to satisfy their own luxurious tastes as well

as showering it on their friends and family. Unlike Hobbes, Machiavelli refuses to downplay the political seriousness of diverting public wealth to private purposes. It is not simply a question of the volume of resources transferred but also a question of the political effects if people view the transfer as unfair. Luxury and favouritism provoke hatred from the populace and hostility from upstanding citizens. Such degenerate rulers become fearful, and in order to suppress the risk of an uprising of the upstanding citizens and the people, they rule tyrannically. This generates longer-term instability and weakness, because it increases the hatred of the populace even as it temporarily prevents that hatred being expressed; yet rulers act this way for the sake of short term maintenance of their preferred way of life (Machiavelli 1996, I.2.3).

Where Machiavelli views the structure of incentives on sovereigns to be neutral, in Spinoza's view it is frankly perverse. It is not merely that there are problematic passions which are not suppressed by the absolutist structure, but worse, the absolutist model of politics predictably and positively produces pressure on the monarchical sovereign to rule badly. The tasks of ruling are too great for an individual to bear alone; consequently an absolute monarchy will always actually have a covert power structure of advisors and confidants. The absolute monarch always lives in fear of usurpation by one of these advisors, and in fear of the public shifting its allegiance. To stave off elite usurpation and to disempower any popular uprising, the absolute monarch rules corruptly and tyrannically, oppressing the weak and appeasing the strong. To stave off being deposed by their own progeny, monarchs deliberately raise them to be politically inept. However, just as in the Machiavellian scenario, this envisaged conduct cuts against the sovereign's self-preservation in the longer term. Those who are appeased recognise the fragility of their privilege at the hands of an unconstrained monarch, and remain ready at any moment to usurp him or her; those who are oppressed do not develop deep commitment to or love for the monarch, so they will shift their allegiance the moment it becomes strategic to do so (Spinoza 2000, 5.7, 6.5–6.7, 7.1, 7.14).

In conclusion, even if Hobbes's doctrine of the sovereign's right is granted, the material plausibility of his model of politics fails to withstand the Machiavellian and Spinozist critiques. Rulers find themselves under pressure to garner short-term political support by buying off potential usurpers and repressing the populace. Such a logic steers rulers away from rational rule. Hobbes may achieve a plausible

mechanical model of the behaviour of political subjects, but the model fails the plausibility test with respect to the conduct of the sovereign.

§3

In this final section, I imagine and address an objection to my argument, and I demonstrate how the Spinozist model of politics provides an alternative to the Hobbesian absolutist model which avoids the problems that I have raised with the latter, and at the same time should satisfy my imagined objector.

My argument to this point has taken a strong view of what is required for an adequate science of politics. I have demanded that the behaviour of each element in the model of politics, including the sovereign, should be consistent with the predictions of the materialist science of human beings. But I can imagine someone objecting that the demand I have levelled at the science of politics is incoherent. A science which seeks to meet the Hobbesian practicality requirement needs some space for action that is different from what people are just plausibly likely to do. Insofar as the science of politics seeks not merely pure understanding but also to make a practical difference, it must be possible for someone to change existing reality in accord with its recommendations – and this means someone who acts differently from how people usually act. My demand for a strict compatibility with the materialist account of usual sovereign behaviour would result in a political model that is a mere description or analysis of the *status quo* rather than anything which might serve and advance human purposes. We can have any hope at all for practical political change because human beings are potentially amenable to reason regarding their interests, and occasionally, unexpectedly far-sighted rational behaviour will occur despite contrary incentives. And so, with respect to the artificer or actor of a political order, it is necessary to relax the materialism requirement.

I now identify a kernel of truth in this imagined objection, but I show that the appropriate response does not lie in Hobbes's absolutism. I start by clarifying my agreement with three of the imagined objection's premises. First, I grant that there needs to be some space in a practically-oriented science of politics to call for behaviour or intervention which is active and would not simply otherwise occur. Second, I grant that people on occasion can have moments of clear-sightedness which allow them to act rationally in the Hobbesian sense (acting in view of long-term self-

interest), and I even grant that a few individuals may consistently have this capacity. And third, I grant that the actions of such individuals can have lasting positive political ramifications. The point which separates my view from the view of my imagined objector is the problem of reliably maintaining and expressing that rationality. Recalling Machiavelli's and Spinoza's analysis from the previous section, the difficulty with absolutism was not simply that it required the rationality of the sovereign, but that it relied on that rationality being sustained over time, without or contrary to institutional pressures. It was the ongoing imbrication of the sovereign's action in the everyday functioning of the political order that forced us to consider the sovereign as an integral part of the materiality of the political model and not rely on a materially unsupported rationality in her or his conduct.

If this analysis is correct, the appropriate response to my imagined objector becomes clear. The sovereign should be treated as part of the model, due to its ongoing role in its functioning. By contrast, the artificer of the political order should not be a role occupied by the sovereign; but rather it should be a catalytic role external to the model of politics which puts it into play in the first place. My imagined objector can then be satisfied that the science of politics features an agent of change, without the model being implausible in its ongoing material functioning. I'll first lay out the alternative Spinozist model of politics which features the sovereign as an integral part; then I'll address the question of agency.

Despite adhering closely to a Hobbesian model of political right (Spinoza 2000 3.2–3.3, 4.1–4.2), Spinoza diverges from Hobbes in his material analysis and institutional proposals. He sees it as essential that the ongoing function of the model of politics should not rest on any individual's rationality and self-control, and observes explicitly that this applies also to the sovereign ruler.

Now if human nature were so constituted that men desired most of all what was most to their advantage, no special skill would be needed to secure harmony and trust. But since, admittedly, human nature is far otherwise constituted, the state must necessarily be so established that all men, *both rulers and ruled*, whether they will or no, will do what is in the interests of their common welfare; that is, either voluntarily or constrained by necessity, they will all live as reason prescribes. (Spinoza 2000, 6.3; emphasis added)

Instead, the sovereign is thought of as human matter, internal to the political system. Spinoza refuses to leave political decision to the institutionally unconstrained judgement of the sovereign.

So if the safety of a state is dependent on some man's good faith, and its affairs cannot be properly administered unless those responsible for them are willing to act in good faith, that state will lack all stability. If it is to endure, its government must be so organised that its ministers cannot be induced to betray their trust or to act basely, whether they are guided by reason or by passion. (Spinoza 2000, 1.6)

He provides detailed recommendations for the case of monarchical sovereign rule, organised around the principle that a ruler's transient personal passions and desires need to be constrained by stabilising institutions. The monarch should have a governing council constituted representatively from the population, and all governmental decisions and functions should be performed by way of the council (Spinoza 2000, 6.15–6.25). Even though technically the sovereign's will is the source of laws, this should not be interpreted in such a way as to give any authority to the monarch's whims and passions: 'For the fundamental laws of the state should be regarded as the king's eternal decrees, so that his ministers are entirely obedient in refusing to execute his orders if he commands something that is opposed to the state'. (Spinoza 2000, 7.1)

Spinoza claims that the strength of a political order arises not from the sovereign's moral qualities but from the fundamental structures of the commonwealth which durably constrain it. But to return to the crux of the imagined objection: if we do not already possess these structures, who puts them in place? The question is not treated directly or at great length by Spinoza, but drawing upon the historical examples offered in his and Machiavelli's writings, we can reconstruct an answer that should satisfy my imagined objector. Spinoza discusses the case of the Aragonese people. Having cast off the yoke of foreign oppression, they were in a position to deliberate explicitly regarding their preferred constitutional structure and they chose an elective monarchy subjected to rule of law (Spinoza 2000, 7.30). Machiavelli for his part considers examples from the ancient world. Many ancient polities had lawgivers: he offers Lycurgus as a laudable example and Solon as a more flawed one

(Machiavelli 1996, I.2.5–6). These lawgivers did not rule but rather designed systems of laws within which rulers of the polity could govern for the sake of the common good. Machiavelli also takes great interest in the case of Rome, in which there was no lawgiver but rather a gradual exogenous evolution of social forms that gave rise to a division of power between plebs and senate, despite no-one's having planned that outcome (Machiavelli 1996, I.2.7).

As the example of Rome shows, political agency is not necessary to establish a flourishing polity: sometimes chance puts in place good institutions, and the science of politics simply identifies and analyses this good functioning retrospectively. However, the other examples show that there is a possibility of political agency. To draw these examples together into a theory of agency in political change, what is required is a political agent willing to serve as lawgiver – or in other words, willing to establish a political order to which they will subsequently be subjected – plus a fertile political juncture or constitutional moment at which this agent can act. In the historical examples given, there is a difference between the lawgiver (be it an individual or the wider populace) and the sovereign thereby established. But this strict separation is inessential: the crucial element is the kind of constraint which comes to bear on the sovereign under the new political order. Lawgivers could subsequently become sovereigns, so long as their exceptional moral qualities are only relied upon in the initial founding acts; their subsequent ongoing conduct needs to be institutionally restrained. As Machiavelli explains, good lawgivers who go on to become rulers are careful to make the 'maintenance' of the common good rest with many, not just in their own hands (Machiavelli 1996, I.9.2).

In the early modern period, Hobbes was amongst the most strident promoters of the idea of understanding human action through a materialistic and deterministic lens. Nonetheless, his own science of politics devotes much of its energy to a different, more juridical investigation of human behaviour, centred on rationality and duty. I have shown that the material view has more profound political ramifications than Hobbes realises. Once irrationality is viewed not merely as a moral failing but instead as a feature of material bodies that needs to be investigated in its mechanical and material causes and effects, the case for a political order that is institutionally non-absolute becomes more persuasive.

References

- Anstey, Peter R. 2005. 'Experimental versus Speculative Natural Philosophy'. In *The Science of Nature in the Seventeenth Century: Patterns of Change in Early Modern Natural Philosophy*, edited by Peter R. Anstey and John A. Schuster, 219–242. Dordrecht: Springer.
- Baumgold, Deborah. 1988. *Hobbes's Political Theory*. Cambridge: Cambridge University Press.
- Curley, Edwin. 1994. 'Introduction to Hobbes' *Leviathan*'. In *Leviathan: With Selected Variants From the Latin Edition of 1668*, edited by Edwin Curley, viii-lxxvii. Indianapolis: Hackett.
- Frost, Samantha. 2008. *Lessons From a Materialist Thinker: Hobbesian Reflections on Ethics and Politics*. Stanford: Stanford University Press.
- Hobbes, Thomas. 1656. *Elements of Philosophy, The First Section, Concerning Body*. London: Printed by R. & W. Leybourn, for Andrew Crocke, at the Green Dragon in Pauls Church-yard. [=DCo]
- Hobbes, Thomas. 1969. *The Elements of Law, Natural and Politic*. Edited by Ferdinand Tönnies. London: Cass. [=EL]
- Hobbes, Thomas. 1972. 'On Man'. In *Man and Citizen*, edited by Bernard Gert, 3–87. Garden City, N.Y.: Anchor Books. [=DH]
- Hobbes, Thomas. 1994. *Leviathan: With Selected Variants From the Latin Edition of 1668*. Edited by Edwin Curley. Indianapolis: Hackett. [=L]
- Hobbes, Thomas. 2006. *On the Citizen*. Translated and edited by Richard Tuck and Michael Silverthorne. Cambridge: Cambridge University Press. [=DC]
- Johnston, David. 1986. *The Rhetoric of Leviathan: Thomas Hobbes and the Politics of Cultural Transformation*. Princeton: Princeton University Press.
- Kavka, Gregory S. 1986. *Hobbesian Moral and Political Theory*. Princeton: Princeton University Press.
- Machiavelli, Niccolò. 1996. *Discourses on Livy*. Translated by Harvey C. Mansfield and Nathan Tarcov. Chicago: University of Chicago Press.

- Matheron, Alexandre. 1986. 'Spinoza et la décomposition de la politique Thomiste: Machiavélisme et utopie'. In *Anthropologie et Politique au XVIIe Siècle: Études sur Spinoza*, 49–79. Paris: J. Vrin.
- Nagel, Thomas. 1959. 'Hobbes's Concept of Obligation'. *Philosophical Review* 68 (1): 68–83.
- Pettit, Philip. 1999. *Republicanism: A Theory of Freedom and Government*. Oxford: Oxford University Press.
- Pettit, Philip. 2008. *Made with Words: Hobbes on Language, Mind, and Politics*. Princeton: Princeton University Press.
- Shapin, Steven, Simon Schaffer, and Thomas Hobbes. 1985. *Leviathan and the Air-pump: Hobbes, Boyle, and the Experimental Life*. Princeton: Princeton University Press.
- Spinoza, Benedictus de. 2000. *Political Treatise*. Translated by Samuel Shirley. Indianapolis: Hackett.
- Tarleton, Charles D. 1978. 'The Creation and Maintenance of Government: a Neglected Dimension of Hobbes's *Leviathan*'. *Political Studies* 26 (3), 307–327.
- Tuck, Richard. 2006. 'Introduction'. In *On the Citizen*, edited by Richard Tuck and Michael Silverthorne, viii-lij. Cambridge: Cambridge University Press.
- Warrender, Howard. 1957. *The Political Philosophy of Hobbes: His Theory of Obligation*. Oxford: Clarendon Press.
- Watkins, John W. N. 1989. *Hobbes's System of Ideas; A Study in the Political Significance of Philosophical Theories*. Aldershot: Gower.