

PART I

APPLICATION TO GOVERNMENTAL POWERS  
AND THEIR LIMITS

## I

## Getting Past Hobbes

Joshua Cohen

In *Leviathan*, Hobbes defends absolute political authority – authority both institutionally undivided and normatively unlimited. He argues that such authority is preferred to alternative forms of political authority by all rational individuals (insofar as they are rational). The principal advantage of absolute political authority – whether monarchical, aristocratic, or democratic in form – lies in its greater capacity to ensure lasting internal peace. Absolute authority can keep the state from “perishing by internal diseases.”<sup>1,2</sup>

However, a political order subject to absolute authority – what I shall call an “authoritative order” – carries, as Locke famously noted, serious risks. Drawing on premises that Hobbes and Locke plausibly shared, I argue that these risks outweigh the benefits of an authoritative order: Hobbesian foundations do not demand Hobbes’s absolutist conclusions. Moreover, an authoritative order sharply limits self-government. I suggest that a Rousseauian variant of the sense of self-worth so central to Hobbesian psychology and politics makes a normative order – a political order defined by shared public norms – a real alternative to an authoritative order. Such an order can achieve the reconciliation of self-government and authority that Hobbes thinks is humanly unavailable.

Taking the Lockean and Rousseauian points together, we should not be surprised that many people who find great insight in Hobbes’s core assumptions nevertheless reject his political conclusions.

This essay draws from my *Protection for Obedience* (manuscript on file with the author, available on request). I am very grateful to Sharon Lloyd for her wonderfully generous editorial work in extracting this essay from the larger manuscript. I am also grateful to the many graduate students at MIT and Stanford who have been in classes in which I presented the larger argument on which the essay draws.

<sup>1</sup> In the footnotes, references to *Leviathan* are given by chapter and paragraph number, followed by page number(s) in Thomas Hobbes, ed. Richard Tuck (Cambridge: Cambridge University Press, 1991).

<sup>2</sup> *Leviathan*, XXIX.1, p. 221.

## PROTECTION FOR OBEDIENCE

Hobbes's political theory is founded on an account of human nature – the “known natural inclinations of mankind”<sup>3</sup> – and the natural facts of human interdependence. He draws two important conclusions from those foundations. First, it is rational for individuals to cooperate peacefully with others – to follow the laws of nature, which together provide a code for the peaceful cooperation of a multitude of individuals – *on condition that* they expect peaceful cooperation from others. Second, despite this conditional rationality of peaceful cooperation, conflict may arise from three sources: *ignorance* of the laws of nature and of the conditional rationality of compliance with those laws; the *temptations* to irrationally short-sighted conduct associated with myopic passions such as pride, envy, covetousness, and jealousy; and lack of *assurance* that others will comply with the laws of nature, despite the conditional rationality of such compliance.

These three sources of conflict – problems of ignorance, temptation, and assurance – are exacerbated by competition for scarce means (what Hobbes generically calls “powers”) for satisfying desires. Further pressure to conflict arises from by rationally unresolvable disagreements about which particular specification of the abstract requirements of natural law is best. Because the desire for self-preservation is so fundamental, neither *competition* for scarce means nor the need for *coordination* around a particular specification of the natural laws will by itself generate violent conflict. But they will encourage the passions that do.

Although cooperation is conditionally rational, then, actual cooperation faces serious hurdles. Still, the most pressing human desires are for self-preservation and felicity or happiness (the satisfaction of desires over the course of a whole life). Because the failure of peaceful cooperation presents calamitous threats to both preservation and felicity, these desires give each person strong reasons to want the three sources of conflict resolved. But the only resolution, Hobbes argues, demands a troubling sacrifice of autonomy or self-government. Troubling, because each person “naturally love[s] liberty and dominion over others”<sup>4</sup> Indeed, “there are very few so foolish, that had not rather govern themselves, than be governed by others.”<sup>5</sup> However, preferring preservation and felicity above all, thus aiming to secure “their own preservation and a more contented life,” all individuals have good reason to support the introduction of “restraint upon themselves.”<sup>6</sup> Such (self-)restraint is achieved by subordinating one's own individual will and judgment to the will and judgment of an absolute authority, thus submitting to the constraints on

<sup>3</sup> Ibid., Review and Conclusion, p. 489.

<sup>4</sup> Ibid., XVII.1, p. 117.

<sup>5</sup> Ibid., XV.21, p. 107.

<sup>6</sup> Ibid., XVII.1, p. 117.

liberty imposed by that authority and acknowledging its judgment as taking precedence over one's own. Because our nature and circumstances put us so deeply at odds with each other, because reason only instructs in the pursuit of our separate aims, and because our passions get in the way of rational conduct, we can overcome tendencies to conflict and the attendant misery only by establishing an authority premised on our common subordination: only if I "give up my right of governing myself."<sup>7</sup> Thus the "mutual relation between protection and obedience": a promise of obedience for an assurance of protection.

Hobbes's idea about the need for subordination – more particularly, subordination to an authority that is unified and unlimited – can be understood in terms of the interconnected roles of power and authority in addressing the three sources of conflict. Peace requires an agent with power because power is required to tame or "bridle" the passions and so to resolve the problems of temptation<sup>8</sup> and assurance.<sup>9</sup> Because the laws of nature are "contrary to

<sup>7</sup> Ibid., XVII.13, p. 120.

<sup>8</sup> The theory of natural law in *Leviathan*, XIV–XV tells us that natural laws are elements in a code of peace and that peaceful cooperation is conditionally rational. But those laws run "contrary to our natural passions, that carry us to partiality, pride, revenge, and the like." See *Leviathan*, XVII.2, p. 117. So conflict may arise because (at least some) people fail to act rationally. Even if they expect others to act peacefully, these people are led by such passions as pride (and the anger it characteristically produces), hate, lust, ambition, and covetousness into distorted (short-term and partial) estimates of the consequences of a course of action. Thus blinded by "some sudden force of the passions," agents may (irrationally) fail to cooperate even under conditions in which they expect others to comply, and, therefore, in which rationality dictates their own compliance. The passion of pride is a particularly important source of conflict. Like all passions, pride leads individuals to act with insufficiently prudent concern for their own long-term advantage – and that means without sufficient concern for the long-term personal advantage of peaceful cooperation. Moreover, it leads more immediately to efforts to dominate others and appropriate their powers. For the prideful person has a nonderivative desire to dominate others, experiences joy in such domination, and feels anger from wounded pride at the failure to achieve it. The best response by rational agents in a population that includes agents moved by pride is to take preventive and preemptive measures of self-protection.

<sup>9</sup> Even among fully rational agents, peaceful cooperation may also fail for want of assurance, reflecting mistrust or "diffidence." See *Leviathan*, XIII.4, p. 87. Suppose that everyone recognizes that cooperation is rational if they can count on the cooperation of (most) others. Suppose as well that everyone is in fact fully rational. Still, in the face of uncertainty about the rationality of others (or about their beliefs about the rationality of others) and in the absence of guarantees that others will in fact behave cooperatively, it may be rational to refrain from cooperative behavior oneself, and, anticipating their attacks, to protect oneself by whatever means are available (including preemption). The need for assurance does not depend on the actual irrationality of others. I may think that others are irrational, or that there is some chance that they are irrational (though in fact they are not). If I think they are (or might be), then I will need assurance of my own protection, and if it is not provided by a third party, I will provide it through preemption. Or suppose I think (correctly) that others are rational; and suppose they think I am rational, but I think that they think that I am not rational (whether or not they do makes no difference here). Then I should expect preemptive belligerence from them because I think they expect attacks from me. But if I have that expectation, then it is

our natural passions, that carry us to partiality, pride, revenge, and the like”<sup>10</sup> power is required “to keep them [people] in awe, and tie them by fear of punishment to the performance of their covenants and observation of those laws of nature ...”<sup>11</sup> But keeping people in awe requires a lasting power of considerable magnitude. God has such power naturally, and His overwhelming, awe-inspiring power is the source of His authority: “The right of nature, whereby God reigneth over men, and punisheth those that break his laws, is to be derived, not from his creating them as if he required obedience, as of gratitude for His benefits; but from his irresistible power”<sup>12</sup> But the fact of human equality – our equal vulnerability to injury from others – means that no human being naturally has power of this magnitude. So we have to construct an artificial power of sufficient magnitude to address the sources of conflict. And we construct that awe-inspiring power by creating an authority with command over our powers – by each of us submitting our control over our own powers to the will and judgment of an agent. What must this authority be like, then, if, by creating it, we ensure a power of sufficient magnitude to solve the problems of ignorance, temptation, and assurance?

#### AUTHORITATIVE AND NORMATIVE ORDERS

Consider three ways to construct the power needed to address the sources of conflict.

First, with a “normative order,” an association with a shared set of public values and norms such that everyone regards him- or herself (and knows that others regard themselves) as having a supreme obligation to act on those values and norms on condition that others act on them as well. For example, the norms might be that the society should be fair; that everyone should have certain basic liberties; that there should be equal protection under the laws; and that it is impermissible for the government to make laws that conflict with

fully rational for me to act belligerently; and if I do, I confirm others’ view about me. There are infinitely many cases here; the general point is that even if all agents are fully rational, there may be conflict rather than cooperation, and therefore a need for assurance. Notice, however, that while cooperation among rational agents fails in many cases, pursuing conflict is not a “dominant strategy”: Conflict, that is, is not the best response to the actions of others whatever those actions may be. The state of nature is not a prisoner’s dilemma. The mutually disadvantageous outcome of conflict does not result from each doing what is best no matter what the other person does. On the contrary, if I am assured of the cooperation of (most) others, then my best response is to comply with the laws of nature, not to defect from them. That is the thrust of Hobbes’s response to the fool, and, in general, of the derivation of the laws of nature. See *ibid.*, XV.4–10, pp. 101–4. The problem of assurance (or diffidence) is that since I am uncertain about the willingness of others to reply to my cooperation with their own, I may need to protect myself by pursuing conflict rather than cooperation.

<sup>10</sup> *Leviathan*, XVII.2, p. 117.

<sup>11</sup> *Ibid.*, XVII.1, p. 117.

<sup>12</sup> *Ibid.*, XXXI.5, pp. 246–7.

the requirements of fairness, liberty, and equality. Alternatively, the shared norms might simply be the laws of nature. Individuals would compact with one another to comply with those laws; they would in effect assign an agent the responsibility to interpret and enforce the laws; and the authority would be acceptable if it remained within the bounds of those laws, reasonably interpreted. In a normative order, it is common knowledge that everyone takes these norms or principles to be the final authority. Political authority in a normative order is understood to be legitimate only if it satisfies these norms. In short, in a normative order, we have limited or conditional authority – limited, because the authority would be legitimate only if it complied with the common understanding of the terms of its proper exercise.

Second, with a political order under divided authority. Intuitively, the idea is not that there is system of agreed norms but an organization of authority – a set of institutions – that is not unified around a single agent. So there might be separate branches of government, each of which is acknowledged by the population to have supreme authority in a specified sphere – for example, one in foreign policy, another in raising revenue, a third in adjudication, and several together in legislation. (Here, I focus on the case of horizontal division, but federal-style vertical division is another possibility.) Policies would then require coordination across these separate spheres of responsibility. For example, foreign policy decisions would require coordination between the policymaking authority and the revenue authority. This might involve constitutionally defined procedures that effectively require the agreement of the different powers before any action is taken.

Third, with Hobbes's authoritative order. By contrast with the normative order, final authority does not lie in a system of norms that each person aims to interpret and follow, but with a determinate agent: not a sovereign scheme of laws, rules, or principles, but a sovereign lawmaker capable of acting. By contrast with the scheme of divided authority, the authoritative order does not have distinct authorities who are understood to be supreme in separate spheres and whose coordination is necessary for state action. Instead, supreme authority is held by an individual (in a monarchy) or group of individuals (in an aristocracy or democracy) who has or have the authority to make the rules, enforce them, and conduct relations with other states (with subordinate officials standing in an agency relation to the supreme authority). That agent is authoritative in that each subject accepts the right of that agent to rule, and so accepts the judgment and will of that one agent as his or her own.

The authoritative order differs from the normative order because it subordinates norms to an agent with authority. It thus rejects the “error of Aristotle’s *Politics*, that in a well-ordered commonwealth, not men should govern but the laws.”<sup>13</sup> The authority itself is unconditional or unlimited in that it

<sup>13</sup> *Ibid.*, XLV.1.36, p. 471.

stands juridically superior to the laws and norms of the society: Will, not law or reason, is the basis of the state. So there are, for example, no constitutional laws defining the legitimate scope and limits of sovereign authority – thus imposing legal limits on political authority – because all the laws are themselves the commands of and so subject to alteration by the will of the sovereign: “The sovereign of a commonwealth ... is not subject to the civil laws. For having the power to make, and repeal laws, he may when he pleaseth, free himself from that subjection, by repealing those laws that trouble him, and making of new; and consequently he was free before.”<sup>14</sup>

Two apparent limits on sovereign authority, as Hobbes describes it, may appear to qualify this description of the authoritative order. First, the “true liberties of subjects” are defined as areas of conduct in which subjects are morally at liberty not to comply.<sup>15</sup> Thus, subjects are not obliged, for example, to kill themselves if the sovereign orders them to, nor is a person required to refrain from taking from others when his or her survival is at stake, nor is there any obligation of self-incrimination. But the true liberties do not limit authority because they are not claim-rights that the sovereign or other subjects are required to respect.

Second, Hobbes presents a set of responsibilities associated with the office of sovereign, which include a responsibility to make good laws, which serve the public benefit.<sup>16</sup> But the sovereign’s authority is not limited to making good laws. Thus a law’s failure to be good has no bearing on its validity as law. And this observation generalizes to all the responsibilities of the sovereign office.

In an authoritative order, then, the test for legal validity looks entirely to a regulation’s source, not at all to its content: Whatever the substance of the requirements it imposes, it is valid law, and falls within the legitimate authority of the sovereign, just in case it issues from the sovereign’s will. Moreover, in the authoritative order, we have a way to identify the sovereign, the sovereign’s will, and the legitimate acts of the sovereign, quite apart from the content of what the sovereign wills. In contrast, the normative order involves

<sup>14</sup> Ibid., XXVI.6, p. 184. On the legal illimitability of sovereignty, see John Austin, *The Province of Jurisprudence Determined*, and the criticisms of legal illimitability in H.L.A. Hart, *The Concept of Law*, chapter 4, and “Sovereignty and Legally Limited Government,” in *Essays on Bentham*, chapter 9. According to Austin, “Supreme power limited by positive law is a flat contradiction in terms.” Like Austin, Hobbes argues against legal limitation by showing that such limitation is inconsistent with the nature of civil law as a system of sovereign commands. But in the end Hobbes is moved less by considerations of analytical jurisprudence than by the substantive thesis that legal limits on sovereign authority would have destructive implications. So Hart’s replacement of commands with rules as the “key to the science of jurisprudence” is much less damaging to Hobbes than to Austin. In any case, my own reconstruction of Hobbes’s case for political absolutism does not appeal at all to considerations about the incoherence of legal limitations.

<sup>15</sup> Ibid., XXI.10–7, pp. 150–2.

<sup>16</sup> Ibid., XXX.1, p. 231.

content constraints on legitimate law: An enactment counts as legitimate only if it meets those constraints (or at least, meeting them counts in favor of its standing as legitimate law).

The rationale for a system of divided authority is reasonably clear (I will put the normative order aside for now).<sup>17</sup> Recall that the idea of a scheme of divided authority is to establish several bodies with authority in different spheres – for example, an elected parliament with the authority to control revenue and an independent executive with the authority to enforce the rules and conduct relations with other states. A system of this kind might, first, provide good governance by vesting different sorts of authority in bodies especially suited to its exercise – foreign policy in the hands of a unified and therefore energetic executive, control of revenue in the hands of a body closely attentive to popular concerns, legislation jointly in the hands of a legislature with information about local interests, and an executive with a sense of the demands of enforceability and prospects of compliance.

Second, we might hope that the division of public authority would limit public power, thus leaving greater scope for individual liberty. Now we all “love liberty” because, *ceteris paribus*, more liberty means more opportunity for achieving our aims. So the scheme of divided authority would be preferred by everyone to a system that is equally likely to keep the peace but less protective of liberty.

The problem, of course, lies in “equally likely.” Preservation and felicity (and therefore peace) are more fundamental goods than liberty. Liberty is desired as a means; preservation and felicity are the ends. So if divided authority is less likely to keep the peace than an order that leaves less scope for liberty, then it would be collectively rejected in favor of that alternative. To be sure, limits on liberty are undesirable. But the “condition of man in this life shall never be without inconveniences.”<sup>18</sup>

The central problem with dividing authority might be understood in two ways. One problem is that such division may be seen as establishing separate fundamental allegiances within a single territory. In effect, we have separate states within a single territory – “not one independent commonwealth, but three independent factions” – comprising those who acknowledge the legislative as supreme, those who acknowledge the executive as supreme, and those who accept the authority of the two operating in concert.<sup>19</sup> Thus, when the predictable conflicts between authorities emerge, when they need to coordinate but disagree, different subjects will ally with different sides in that conflict, and the political society may degenerate into civil war. Alternatively, the

<sup>17</sup> On the distinction between mixed government and a normative order, see Hobbes, *Elements of Law* II.1.13–6. The first two paragraphs describe a normative order; the second two are about mixed government.

<sup>18</sup> *Leviathan*, XX.18, p. 145.

<sup>19</sup> *Ibid.*, XXIX.16, p. 228.



problem might be that when the predictable conflicts emerge, subjects have no fundamental allegiance at all. Either way, civic life threatens to degenerate into civil war.

With favorable circumstances, systems with divided authority may of course (as Hobbes acknowledges) have considerable longevity, providing subjects with conditions of security and felicity. But, as the contention that they are not really one commonwealth but several is meant to suggest, they bear the seeds of their own dissolution into open conflict of the kind experienced in the English civil war. Indeed, the widespread idea that political powers were properly divided “between the King, and the Lords, and the House of Commons” was, Hobbes claims, the principal cause of the English Civil War.<sup>20</sup> Systems with divided powers suffer – as Hobbes thought was evident in England – from a failure to solve the problem of coordinating around a determinate interpretation of the abstract laws of nature. Thus, there are a “diversity of opinions”<sup>21</sup> about the terms of cooperation, with none of the bodies fully having the authority to fix those terms. Moreover, and more fundamentally, because the division of authority (and allegiance) limits public power, problems of temptation and assurance are not adequately resolved. By limiting power and thus reducing the capacity to overawe subjects into obedience, they increase the chances that some – perhaps united by one of the recognized authorities – will be tempted to act against others. Recognizing the increased likelihood of temptations, others lack assurance. Recognizing that lack, others must prepare themselves for conflict.

In short, dividing authority (like limiting authority) limits power. And limiting power diminishes the capacity of the sovereign to address the main sources of conflict.

One final point before proceeding to Hobbes’s solution. I have described the defects in schemes of divided authority (parallel points apply to limited authority). But the mere presence of defects is insufficient ground for rational agents to reject them: life is never without inconveniences. Both normative order and divided authority have some advantages over a state of nature with no political authority at all, even if each might, under certain strains, slide back into a calamitous state of nature. Furthermore, each gives greater scope to the good of self-government than an authoritative order. To show, then, that it would nevertheless be rational to reject them, it needs to be shown that the unchained authority of the authoritative order promises to create public power of a kind that is more likely to motivate compliance, therefore more likely to keep the peace, and so more likely to ensure preservation and felicity than the alternatives. If it is, and certainly if the likelihoods are large

<sup>20</sup> Ibid., XXVIII.16, p. 127. In *Behemoth*, Hobbes presents a different diagnosis, with more emphasis on religious authority. See *Behemoth, Or The Long Parliament* (Chicago: University of Chicago Press, 1990).

<sup>21</sup> Ibid., XVI.17, p. 115.

*Getting Past Hobbes*

11

enough, then the gains in preservation and felicity arguably outweigh the loss in self-government.

In the authoritative order, then, the unbound sovereign has the right (and the responsibility) to fix a determinate interpretation of the laws of nature, thus addressing the problem of coordination. Moreover, the sovereign has the right and responsibility to inform the subjects of the foundations of their duties, thus addressing the problem of ignorance. Furthermore, each person accepts the authority of the sovereign by authorizing the sovereign to provide supreme guidance for the use of his or her powers. Each thus agrees to subject the use of his or her powers to the supreme direction of the sovereign will. The implication of this common authorization is that the sovereign has the sum of the powers of all members at his or her or its disposal, thus as great a power as is possible within the territory. Given this aggregation of all powers, the sovereign has considerably greater power than any subject. The magnitude of that power, created by authorization, should be sufficient power to overawe individual subjects (with fear) into obedience, should they be tempted to depart from the commands of the sovereign. And because individuals will not be tempted, others will be assured of compliance. People will comply from fear of sovereign power, as the passion of fear is pitted against the conflict-inducing passions. And they will comply because reason recommends compliance, on condition that one expects others to comply.

The establishment (through authorization) of absolute authority, then, better addresses the motivational problems of assurance and temptation than either conditional or divided authority. The essential idea is that unconditional and unified authority maximizes sovereign power, by putting the powers of each individual at sovereign command. And that power in turn is most likely to generate the obedience necessary for protection. “For by this authority, given him by every particular man in the commonwealth, he hath the use of so much strength and power conferred on him, that by terror thereof, he is enabled to conform the will of them all, to peace at home, and mutual aid against their enemies abroad.”<sup>22</sup> Limits on or divisions of political authority would not deprive the authority of all power. But lesser power would decrease the level of awe, thus increase the level of temptation, thus weaken the assurance to those who are prepared to obey that their obedience will not be exploited.

To be sure, subjects prefer stronger guarantees that the sovereign will act responsibly, by fostering the public welfare: stronger guarantees that the sovereign will make *good* laws.<sup>23</sup> And they prefer more liberty. But any attempt to impose such guarantees – in the form of basic norms or institutions limiting the authority of the sovereign – would also limit the power of the sovereign. But in limiting that power, they would threaten to undermine its capacity to motivate

<sup>22</sup> Ibid., XVII.13, p. 120.

<sup>23</sup> Ibid., XXX.20–2, pp. 239–40.