



## Normativity and Projection in Hobbes's Leviathan

Stephen Darwall

*The Philosophical Review*, Vol. 109, No. 3 (Jul., 2000), 313-347.

Stable URL:

<http://links.jstor.org/sici?sici=0031-8108%28200007%29109%3A3%3C313%3ANAPIHL%3E2.0.CO%3B2-Z>

*The Philosophical Review* is currently published by Cornell University.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/sageschool.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

## Normativity and Projection in Hobbes's *Leviathan*

Stephen Darwall

A perennial problem in interpreting Hobbes's moral and political thought in *Leviathan* has been to square the apparently irreducible (or, at any rate, unreduced) normativity of central Hobbesian concepts and premises with his materialism and empiricism. Thus, Hobbes defines a "law of nature" as a "precept or general rule, found out by reason, by which a man is forbidden to do that which is destructive of his life" (14.3)<sup>1</sup> and the "right of nature" as "the liberty each man hath to use his own power, as he will himself, for the preservation of his own nature" (14.1). Both are plainly normative. A law of nature is a *precept* that tells us what we ought not, or must not, do. And the right of nature tells us what we may do.

But how does Hobbes intend these normative claims to be understood or, better, to be accounted for? "Words are wise men's counters" (4.13). According to Hobbes's materialist theory of meaning, however, a word can "enter into . . . an account," only if it refers to: (i) some matter or body, (ii) "some accident or quality" of body such as "being moved" or having a certain length, (iii) the "properties of our own bodies" when we have "fancies" or sensory appearances, for example, as of color or sound, or (iv) other names, as, Hobbes thinks, words like 'general' and 'universal' refer to. Where, in this framework, can a normative claim fit? In what properties of bodies might laws or the right of nature consist? Indeed, since Hobbes believes that only terms with these referents can have meaning, what saves his central normative claims from

---

I am indebted to audiences at Brown University, Central Michigan University, Emory University, the Midwest Conference in History of Philosophy, University College, Cork, University of California, San Diego, University of Iowa, University of Pennsylvania, University of Pittsburgh, University of Saskatchewan, University of Southern California, and West Virginia University for helpful discussion of earlier versions. I am grateful to, among others, Richard Arneson, David Brink, David Gauthier, Barbara Herman, Paul Hurley, Louis Loeb, Sharon Lloyd, and, especially, to an anonymous reader for the *Philosophical Review* for many helpful suggestions.

<sup>1</sup>References will be to chapter and paragraph number in Edwin Curley's edition (Indianapolis: Hackett, 1994).

being “insignificant speech” of the kind he disdains among the schools (1.5, 4.20)?

We can put this puzzle another way. Hobbes writes within the tradition of natural law, but he is a modern rather than a classical natural lawyer. On the classical theory deriving from Aquinas, the normativity of natural law is grounded in teleological metaphysics. Implicit in the nature of every natural being is what it is *to be*, a teleological archetype that determines its perfection and good.<sup>2</sup> This is what Aquinas calls the “eternal law.” Since human beings can know their end through reason, eternal law is also a “natural law,” that is, a law rooted in our nature that we can follow or flout. As with all natural beings, however, the normativity of this law derives from metaphysical teleology, from our having a *telos*: an end we are *to seek* built into our nature.

“Words are wise men’s counters.” But, Hobbes adds, they are but “the money of fools,” for those who “value them by the authority of an *Aristotle*, a *Cicero*, or a *Thomas*” (4.13). Hobbes holds that ends are unintelligible except for creatures with “sense and will.”<sup>3</sup> And for them, he re-identifies final causes as efficient causes, namely, as the “endeavours” that cause all voluntary action.

Thus, Hobbes confronted directly the problem that all modern natural lawyers faced in the seventeenth century. If the metaphysical teleology of classical natural law is incompatible with the worldview of the emerging empirical sciences, how is the normativity of natural law to be accounted for? What place is there for oughts in a world of empirical fact? Hobbes’s definition of ethics as a “science” that draws out “consequences from the passions of men” only compounds the difficulty (9, table). How can anything normative follow from propositions of psychology?

Faced with this puzzle, commentators have tended to pursue one of two strategies. Some have taken Hobbes’s normative claims at face value, as largely independent of his apparent empiricism and materialism. These interpreters argue that Hobbes’s moral and political arguments are to be understood within the framework of a non-empiricist moral realism or as some form of theological voluntarism.<sup>4</sup> By far the more usual tack, however, has been to un-

---

<sup>2</sup>I discuss this further in *The British Moralists and the Internal ‘Ought’: 1640–1740* (Cambridge: Cambridge University Press, 1995), 4–6.

<sup>3</sup>Thomas Hobbes, *English Works*, 5.1, 132.

<sup>4</sup>For the former, see A. E. Taylor, “The Ethical Doctrine of Hobbes,” in

derstand Hobbes's empiricism and naturalism as significantly informing his moral philosophy.<sup>5</sup> Writers who follow this line take their cue from Hobbes's remark that the "dictates of reason" he terms natural "laws" are only "improperly" so called. Really, they are "conclusions, or theorems concerning what conduceth to the conservation and defence" of oneself (15.41). Reading this together with Hobbes's characterization of ethics as a science that draws conclusions "from the passions of men," these writers frequently infer that what Hobbes must believe is that normative claims, such as that everyone should seek peace or "keep their covenants made," can be reduced to propositions of empirical fact, such as that each person unavoidably seeks self-preservation and that seeking peace and keeping covenant are necessary means to achieving this end.

These writers also claim support from Hobbes's theory of value, which they interpret as some form of metaethical subjectivism.<sup>6</sup> Thus, when Hobbes writes that "*good* and *evil* are names that signify our appetites" (15.40), and that whatever a person desires he "calleth good" (6.7), he is generally read as saying that 'good' means something like "desired by me." Added to the premises that I desire self-preservation, that keeping covenant is necessary to achieve this end, and that whatever is necessary to realize something good is likewise good, this yields the conclusion that my keeping covenant would be good. And, assuming that one should do what would be good, it follows that I should keep covenant.

Interpreting Hobbes as a reductionist, however, gives him the burden of explaining how facts about the agent's psychology can

Keith Brown, *Hobbes Studies*, ed. Keith Brown (Oxford: Blackwell, 1965). For the latter, see, somewhat problematically, Howard Warrender, *The Political Philosophy of Hobbes* (Oxford: Oxford University Press, Clarendon Press, 1957) and, more straightforwardly, A. P. Martinich, *The Two Gods of Leviathan: Thomas Hobbes on Religion and Politics* (Cambridge: Cambridge University Press, 1992).

<sup>5</sup>See, for example, David Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes* (Oxford: Oxford University Press, Clarendon Press, 1967), and "Thomas Hobbes: Moral Theorist," *Journal of Philosophy* 76 (1979): 547–59; Jean Hampton, *Hobbes and the Social Contract Tradition* (Cambridge: Cambridge University Press, 1986); and Gregory S. Kavka, *Hobbesian Moral and Political Theory* (Princeton: Princeton University Press, 1986).

<sup>6</sup>See the references to Gauthier, Hampton, and Kavka in notes 25–27 below.

have normative practical force. Why does the fact that I happen to desire something create a reason *for me* to seek it? That I actually desire something seems to be one thing, that it is desirable, something I *should* desire or seek, another. Facts about my desires seem to be psychological descriptions of my situation as they might be viewed from an observer's standpoint. Claims about what I should do, however, concern what there are (normative) reasons for me to do as these might be viewed from the perspective of an agent deliberating about what to do.<sup>7</sup> How does an agent get from the fact that some means are necessary to an end she actually seeks to the conclusion that she ought to take those means? Maybe the end is one she ought not seek.<sup>8</sup> What gets her from the apparently non-normative, psychological claim to the normative practical conclusion? Reductionism seems to change the subject from ethics to psychology. Interpreting Hobbes in this way, therefore, makes him subject to the problem with which we began.

There is a further problem with subjectivism that must also infect a subjectivist interpretation of Hobbes. When two people, *A* and *B*, say of something, *X*, that it is good and that it is not good, respectively, there seems to be a disagreement between them about some objective matter, concerning which both cannot be correct. If subjectivism is true, however, what *A* really says is that she, *A*, desires *X*, and what *B* says is that he, *B*, does not or, perhaps, that he is averse, to *X*. But if this is so, then there is no objective issue between *A* and *B* concerning which no more than one of them

---

<sup>7</sup>For criticism of the idea that desire provides a normative reason for action (as opposed, say, to being a condition in which the desirer *takes* something about her desire's *object* as a normative reason), see E. J. Bond, *Reason and Value* (Cambridge: Cambridge University Press, 1983), 1–41; Stephen Darwall, *Impartial Reason* (Ithaca: Cornell University Press, 1983), 25–82; Philip Pettit and Michael Smith, "Backgrounding Desire," *Philosophical Review* 99 (1990): 565–92; and T. M. Scanlon, *What We Owe to Each Other* (Cambridge: Harvard University Press, 1998), 33–55.

<sup>8</sup>For a critique of the idea that the fact that something is a (necessary) means to the agent's end is a normative reason for the agent to take it, as opposed, say, to making it irrational for the agent to both maintain the end *and* fail to take the means, see R. M. Hare, "Wanting: Some Pitfalls," in *Agent, Action, and Reason*, ed. Robert Binkley, Richard Bronaugh, and Ausonio Marras (Toronto: University of Toronto Press, 1971); Patricia Greenspan, "Conditional Oughts and Hypothetical Imperatives," *Journal of Philosophy* 72 (1975): 259–76; Stephen Darwall, *Impartial Reason*, 43–50; and John Broome, "Normative Requirements," *Ratio* 12 (1999): 398–419.

can be correct. *A* will not be affirming what *B* denies, nor vice versa. And subjectivism is not simply problematic in its own right. It is especially dubious as an interpretation of Hobbes since, as we shall see, Hobbes says that disagreements about good and evil can be serious enough to lead to war.

I have argued elsewhere that Hobbes can be interpreted as initiating an *internalist* tradition in British moral thought, which held that normative force should be understood as motive force for an agent deliberating about what to do.<sup>9</sup> This tradition came in two versions, *empirical naturalist internalism* and *autonomist internalism*, both of which held that practical normativity—an agent's *oughting* to do something—consists in her being moved so to act were she to reason properly. I claimed that Hobbes initiated naturalist internalism, a tradition that also included Cumberland, Locke, Hutcheson, and Hume. According to naturalist internalists, practical reasoning consists entirely in empirically based *theoretical* reasoning about practice.<sup>10</sup> An agent ought to do something if, and only if, she would be moved to do it were her empirical beliefs about her practical situation error-free. *Autonomist* internalists, like Cudworth, Shaftesbury, Butler, and, in certain moods, Locke, on the other hand, held that only *free* agents can be subject to oughts. They believed that obligation derives from a distinctive source of motivation available through self-determining, autonomous agency—a uniquely *practical* reason.<sup>11</sup> According to autonomist internalists, what a person ought to do is what he would be moved to do were he correctly to exercise this capacity for self-determining deliberation and choice.

I argued that Hobbes was a naturalist internalist who held that all practical reasoning is instrumental. Theoretical reasoning—for example, that concerning the relation between keeping covenant and self-preservation—acquires a normative force in the practical reasoning of an agent who has self-preservation as an end. But this brings us back squarely to the problem I raised earlier: *How* can it do this? Reducing normative propositions to propositions about our ends or desires and the means to satisfying them seems to change the subject from the normative practical question facing a

---

<sup>9</sup>In chapter 3 of *The British Moralists and the Internal 'Ought'*.

<sup>10</sup>With some caveats for Hume and Hutcheson.

<sup>11</sup>Again, with caveats, this time for Locke.

deliberating agent—What should I do?—to a question of empirical inquiry—What occurrences would cause the satisfaction of my desire?<sup>2</sup>

What I failed to see sufficiently clearly was that the instrumental reasonings with which Hobbes was concerned are not practical because they proceed from a psychological premise about practice, namely, that self-preservation is one's end. Rather, according to Hobbes, we reason *from the end* of self-preservation itself, not from the fact that it is our end. In having self-preservation as end, we accept a normative premise that differs from any description of our psychology. We see our survival as good or as something that ought to be achieved. And so we conclude that we ought to keep covenant when we see doing so as a necessary means to staying alive. But how exactly is this supposed to work, according to Hobbes? How can Hobbes provide an account of instrumental practical reasoning that both respects his empiricism and materialism and does justice to normative practical force as it presents itself to agents deliberating about what to do?

In what follows, I argue that *Leviathan* is best interpreted as providing a *projectivist* theory of this normative premise and, indeed, of normative language and thought more generally. There are many passages, I shall argue, where Hobbes is saying, not only that the proper place to locate ethical thought is in practical reasoning, from the agent's point of view, but also that ethical or normative thought and discourse are expressions or projections of what, in Hobbes's view, makes us deliberating agents in the first place, our *desires*. In desiring our survival, Hobbes holds, we ascribe to it a property, that of being good (something we ought to achieve), that it does not literally have.<sup>12</sup> Moreover, I shall argue, Hobbes explic-

---

<sup>12</sup>Although one might wish to distinguish the evaluative from the normative for various purposes, I will take it that the evaluative propositions with which Hobbes is here concerned are normative, if only implicitly, since they are neither good-of-a-kind judgments nor hedged in some other way. On my interpretation, Hobbes holds that the (evaluative) thought that something is good implies the (normative) thought that there is some (normative) reason for the agent whose thought it is to bring it about and, thus, that she ought to, other things being equal. (For the distinction between normative and motivating reasons (earlier termed justifying and explaining or "agents'" reasons, respectively), see, for example, Michael Smith, *The Moral Problem* (Oxford: Blackwell, 1994), 94–98; Stephen Darwall, *Impartial Reason*, 28–32; Kurt Baier, *The Moral Point of View* (Ithaca: Cornell University Press, 1958), 148–56.)

itly analogizes ethical judgments in this way to color judgments, of which he suggests a similarly projectivist theory.<sup>13</sup>

This interpretation may seem to face significant problems of its own, however. If values and oughts are, like color, strictly illusory, then why doesn't knowledge of their projective character undermine ethics? Hobbes doesn't consider this question directly, but I shall argue that we can infer his answer, however unsatisfactory it may be. Desire—and deliberation under its influence—are, Hobbes believes, what give ethical thought its point and function. So long as we are alive, we desire. And so long as we desire, we find ethical thought and, perforce, deliberation, unavoidable. Hobbes addresses *Leviathan's* central normative claims, the laws of nature, to his readers as deliberating agents. As theorems about what leads to self-preservation, these provide lemmas that can be combined in each agent's practical reasoning with a normative thought each finds unavoidable under the influence of a desire he cannot shake.

Before proceeding, I should make it clear what I am claiming and what I am not. I am not saying that Hobbes himself had any clear notion of the metaethical differences between projectivism and subjectivism, nor that he ever explicitly rejects subjectivism. These are our distinctions, and it would be anachronistic to read them back into Hobbes in any detailed way.<sup>14</sup> Nevertheless, I do claim, against the main tenor of Hobbes scholarship, that projectivism provides a significantly better interpretation of what Hobbes says than subjectivism does. *Leviathan* thus provides one of the first, if not the very first, expressions of a projectivist metaethics, for similar reasons, indeed, to those that underlie projectivism today.

### Color

We should begin with color. Analogies to color and to secondary qualities more generally run through metaethical writing of the

---

<sup>13</sup>Strictly speaking, this does not simply extend the story I told about Hobbes in *The British Moralists and the Internalist*. It amends it. I there placed Hobbes at the head of a naturalist tradition in early modern British metaethics. I now think, however, that while Hobbes is a methodological and metaphysical naturalist, the most plausible metaethical category to place him in is not naturalism but projectivism. But note the next paragraph but one in the text.

<sup>14</sup>For a discussion of these distinctions, see Stephen Darwall, Peter Railton, and Allan Gibbard, "Toward *Fin de Siècle* Ethics," *Philosophical Review* 101 (1992): 115–89.



last twenty years. This perhaps began with J. L. Mackie's presentation of what he called the "error theory" and, in this connection, with his discussion of projective "objectification" in Hume.<sup>15</sup> According to Mackie and Mackie's Hume, value judgments are like color judgments in that both are rooted in "objectifying" states of mind that project objective qualities onto objects that do not actually have them. In color experience we see objects *as colored*, as though they had an objective, categorical color property. But they actually have no such property. There are many related properties they do have, for example, the disposition to cause certain experiences (as of color) in normal observers in standard viewing conditions. But this is not how they seem to us in color experience. Our experience is as of an intrinsic or categorical color property in the object.

Mackie argues that Hume holds that ethical properties should be understood similarly. Thus, Hume famously compares vice and virtue "to sounds, colours, heat and cold, which, according to modern philosophy, are not qualities in objects, but perceptions in the mind."<sup>16</sup> Although both ethical and secondary qualities *are* nothing but "perceptions," it is important to Hume that both appear to us *as* anything but items in our own minds. "[T]he mind has a great propensity to spread itself on external objects, and to conjoin with them any internal impressions, which they occasion."<sup>17</sup> We "naturally imagine a conjunction, even in place, betwixt the objects and qualities"—we experience color and virtue as really being in objects and characters, respectively. Still, these qualities "really exist no where." In the case of vice and virtue, Hume attributes the projection to sentiment or taste, which he contrasts with reason. Reason "discovers objects as they really stand in nature; without addition or diminution," whereas taste "has a productive faculty,

---

<sup>15</sup>J. L. Mackie, *Ethics: Inventing Right and Wrong* (Harmondsworth: Penguin Books, 1977) and *Hume's Moral Theory* (London: Routledge and Kegan Paul, 1980). It should be noted that Mackie usually discusses objectification in relation to Hume, he does cite Hobbes as an example of the view that value is a projection of desire (*Hume's Moral Theory*, 43). However, Mackie does not work this out in any detail, nor does he relate it to what Hobbes says about color.

<sup>16</sup>David Hume, *A Treatise of Human Nature*, 2d ed., ed. L. A. Selby-Bigge, 2d. ed., with text revised and variant readings by P. H. Nidditch (Oxford: Oxford University Press, 1978), 469.

<sup>17</sup>*Treatise*, 167.

and gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises in a manner a new creation."<sup>18</sup>

Recent invocations of an analogy between ethical and secondary qualities go well beyond Mackie. Michael Smith models a nonreductive dispositionalist theory of value quite explicitly on a similar account of color, using the latter to exhibit how an analysis might be nonreductive but still informative.<sup>19</sup> And other reductive dispositionalists like David Lewis and Peter Railton exploit the analogy as well.<sup>20</sup> Sensibility theorists, like John McDowell and David Wiggins, provide yet another example of how the analogy can be deployed, this time within a philosophical framework that, unlike that of Smith, Lewis, and Railton, stresses discontinuities between ethics and science.<sup>21</sup>

### Hobbes on Color

While contemporary discussion looks to Hume for the analogy between color and value, especially for a projectivist approach to both, I hope to show that these elements were also present in *Leviathan* (almost a century before the *Treatise*). In particular, I shall argue, *Leviathan* exhibits an application to the case of desire, deliberation, and value of a projectivist approach to color that Hobbes took from Galileo.<sup>22</sup>

---

<sup>18</sup>David Hume, *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, 3d ed., ed. L. A. Selby-Bigge, with text revised and notes by P. H. Nidditch (Oxford: Oxford University Press, Clarendon Press, 1985), 294.

<sup>19</sup>Michael Smith, *The Moral Problem* (Oxford: Blackwell, 1995).

<sup>20</sup>David Lewis, "Dispositional Theories of Value," *Proceedings of the Aristotelian Society*, supp. vol. 63 (1989): 113–37; Peter Railton, "Moral Realism," *Philosophical Review* 95 (1986): 163–207.

<sup>21</sup>John McDowell, "Values and Secondary Qualities," in *Morality and Objectivity*, ed. T. Honderich (London: Routledge and Kegan Paul, 1985); David Wiggins, "A Sensible Subjectivism?" in *Needs, Values, Truth: Essays in the Philosophy of Value* (Oxford: Blackwell, 1987).

<sup>22</sup>Recent scholarly discoveries have revealed how Hobbes's thinking underwent a fundamental shift in the 1630s from Renaissance humanism to modern natural philosophy (Quentin Skinner, "Bringing Back a New Hobbes," *New York Review of Books* (April 4, 1996)). New evidence of the former has come in the republication of *Three Discourses*. This was initially published anonymously in 1620, but computer analysis has recently suggested that it was written by Hobbes (Thomas Hobbes, *Three Discourses*, ed. Noel B. Reynolds and Arlene W. Saxonhouse (Chicago: University of Chi-

We should begin with Galileo. Colors, Galileo says, are like tastes and odors in residing, as a matter of “objective existence,” “in our sensitive body,” so that “if the perceiving creature were removed, all of those qualities would be annihilated and removed from existence.” However, “just because we have given special names to these qualities, different from the names we have given to the primary and real properties, we are tempted into believing that the former really and truly exist as well as the latter.”<sup>23</sup> Galileo diagnoses the cause of this projective error not to be our experience of color, but our linguistic habits. It is giving special names to colors as we do to “primary and real properties,” that has tempted us to believe that they also have objective existence.

Compare this with Hobbes:

[T]his *seeming* or *fancy*, is that which men call *sense*, and consisteth, as to the eye, in a *light* or *colour figured*. . . . All which qualities, called *sensible*, are in the object, that causeth them, but so many several motions of the matter, by which it presseth our organs diversely. Neither in us that are pressed, are they any thing else, but divers motions; for motion produceth nothing but motion. But their appearance to us is fancy. . . . [T]hough . . . the real and very object seem invested with the fancy it begets in us; yet still the object is one thing, the image or fancy is another. (1.4)

According to Hobbes, it is color experience that leads us to think, mistakenly, that objects have objective, categorical color properties. The “seeming” or “fancy” is “as to the eye, in a light or colour figured.” The appearance is *as of color* as an objective property or thing. (What Hobbes actually says here is that it is to the eye as if color were a shaped stuff or substance.) But there

---

cago Press, 1995)). One of the *Discourses*, the *Discourse on Law*, is particularly interesting for our purposes, since it is replete with the rhetoric of classical Thomist natural law that Hobbes would come later to scorn. The last two years have also seen the publication of Hobbes’s correspondence, which shows the emergence during the 1630s of a very different Hobbes, someone intensively engaged in experiments with light and optics and fascinated by Galileo’s theory of color (Thomas Hobbes, *The Correspondence*, 2 vols., ed. Noel Malcolm (Oxford: Oxford University Press, Clarendon Press, 1994)). As we know, Hobbes pursued scientific and geometrical research for the rest of life, and his philosophical writings from then on bear the marks, not least, of course, *Leviathan*.

<sup>23</sup>From *The Assayer* (*Il Saggiatore*, 1623). This passage, translated by Arthur Danto, is quoted in Danto and Sidney Morgenbesser, eds. *Philosophy of Science* (Cleveland and New York: Meridien Books, 1960).

really is no such property or thing in nature. While “the real and very object seem invested with the fancy,” object and fancy are quite distinct.

There are, then, the following elements in Hobbes’s projective account of color:

1. We experience color *as* an objective, categorical quality in the object (the object “seem[s] invested with the fancy”).
2. We do not experience it *as* something in us, or as a disposition in the object to cause experiences in us, or creatures like us, etc.
3. There actually is no such objective, categorical, color quality in objects.
4. All there is are the material motions in the object, in us, and in between.

### Hobbes on Desire and Voluntary Motion

Hobbes titles chapter 6 of *Leviathan*, “Of the Interior Beginnings of Voluntary Motions; Commonly Called the Passions; and the Speeches By Which They Are Expressed.” In it, he provides an account of the “fancy” or “appareance” of good and evil that is both structurally analogous to his projectivist account of color and explicitly linked to it.<sup>24</sup> Whereas color qualities are what we projectively attribute to the objects we are viewing when we have certain sensory experiences, Hobbes claims that good and evil are qualities we attribute to objects when we desire or are averse to them, respectively. When a person calls something good, she thus “express[es]” the “passion” that Hobbes calls desire, and contrariwise for aversion.

Hobbes’s thinking is intricate here, so we need to lay it out in some detail. *Animal* or *voluntary motion* differs from *vital motion*, such as is involved in the movement of blood and normal breathing. Since “voluntary motions . . . depend always upon a precedent thought of *whither, which way, and what*, it is evident that the imag-

---

<sup>24</sup>So far as I know, the only interpreter to notice this is Richard Tuck (see his *Hobbes* (Oxford and New York: Oxford University Press, 1989), 52–57). However, Tuck apparently does not see the difference between interpreting Hobbes as a projectivist and seeing him as a subjectivist (relativist), nor the way in which Hobbes’s projectivist treatment of value and normativity fits into his account of deliberation and the will.

ination is the first internal beginning of all voluntary motion” (6.1). Hobbes calls these “small beginnings of motion “*endeavour*.” When an endeavor is “toward” some object, it is an “appetite” or “desire.” And when it is “fromward something,” it is an “aversion” (6.2).

Desire is psychologically identical with *love*—the “same thing,” Hobbes says, except that desire’s object is always something we take to be absent, and contrariwise for love (6.3). And similarly, with appropriate changes, for aversion and hatred. We hate only what we think present, and are averse when we have the same attitude to something we suppose absent. Except for their objects’ mode of presentation, then, desire and love are identical.

To get the next pieces of the puzzle onto the table, we must consider in some depth a passage in which Hobbes makes an explicit analogy with the sensory experience of color.

As, in sense, that which is really within us, is, as I have said before, only motion, caused by the action of external objects (but in appearance, to the sight, light and colour, . . . ), so when the action of the same object is continued from the eyes, ears, and other organs to the heart, the real effect there is nothing but motion or endeavour, which consisteth in appetite or aversion, to or from the object moving. But the appearance or sense of that motion, is that we either call DE-LIGHT, or TROUBLE OF MIND. (6.9)

What sense experience really is, Hobbes thinks, is simply some motion in my sense organs, caused by the sensed object, which motion “is continued inward to the brain” (1.4). Nonetheless, this motion has an “appearance,” which is “as to the eye, in a light or colour figured” (1.4), or, as Hobbes says here, “to the sight, light and colour” (6.9). The appearance is, in fact, a manifestation of the material motions in the sense organs, brain, and object, in the sense that these are what causally underlie the appearance. But they are no part of the appearance’s content. They are not what the appearance is *as of*. The appearance is as of an object with a categorical, objective color property. To mark this distinction, let us call what actually underlies the appearance, the *appearance substratum*. In this case, the appearance substratum is the complex of material motions that are causing the object to look colored. And we can call what the object is appearing *as*, the *appearance content*. In this case, the appearance content is color as an objective, categorical property of an object or “a colour figured.” Finally, we can

refer to the phenomenal “fancy” as the *appearance itself*. So we have three things: the sensory experience or appearance itself, what the object is appearing as (the appearance content), and what actually underlies the appearance, the material motions that are its substratum.

Now follow the motion farther, from the sense organs and the brain on to the “heart.” When the motion extends to the heart, there is frequently a responsive effect, some motion to or from some object, perhaps the sensed object itself. These “small beginnings of motion” are endeavors—desires or aversions, or loves or hatreds, depending on whether they are toward or fromward their objects and whether their objects are conceived as present or absent.

This new motion, the endeavor, also manifests itself in a fancy or appearance. Like the material motions that appear in sense experience, it too is an appearance substratum. And Hobbes here gives us his name for the appearance in which the endeavor manifests itself. He calls the appearance itself “delight” or “trouble of mind.” When I want something, my wanting is manifested in a delightful appearance as I contemplate something I take to be absent (although not, of course, its absence). And when I am averse to something, my aversion manifests itself in a troubled appearance as I contemplate something I take to be absent (again, of course, not its absence). Thus, delight (or “pleasure”) and trouble of mind (or “molestation” or “displeasure”) are the appearances themselves that are associated with the respective *appearance substrata* of desire and aversion.

But what are desire’s and aversion’s respective *appearance contents*? What are their appearances *as of*? To get to Hobbes’s punch line, consider the following example. You’ve just finished a long run on a hot day and have a powerful thirst. You spy a bottle of cold Gatorade in the fridge. First, the color story: motion in the bottle begins a causal chain leading first to your eyes and then to your brain, causing certain consequent motions there: the color appearance substrata. These motions manifest themselves in a phenomenal “fancy”: the color appearance itself. And this appearance seems to be as of a lime green color, an objective quality of the liquid in the bottle: the color appearance content.

The motion now runs from eyes and brain to your heart (going through the material underpinnings of your thirst). Here it causes

a motion towards the possible state of affairs of your drinking the Gatorade. This new motion is an endeavor, and because it is toward an object you regard as absent, it is a desire. This desire manifests itself in a second appearance, in addition to the one you had when you saw the Gatorade as green. You experience pleasure or delight when you contemplate actually drinking the Gatorade. Your pleasure is the second appearance itself. But, again, what is the *content* of this second appearance? What is it *as of*?

Here is Hobbes's answer: "*Pleasure, . . . or delight, is the appearance, or sense, of good; and molestation or displeasure, the appearance, or sense, of evil*" (6.10). The content of desire's appearance is *value*, and the content of aversion's appearance is *disvalue*. Every desire and love, Hobbes says, "is accompanied with some delight more or less," and every aversion or hatred "with more or less displeasure" (6.11). And pleasure and displeasure are appearances as of good and evil, respectively. So every desire and love manifests itself in an appearance as of a good quality of something. And every aversion or hatred manifests itself in an appearance as of a bad or evil quality of something. You want to drink the Gatorade, and in so wanting you have an appearance as of that's being good. Or you are averse to remaining in your current thirsty state, and in so being, you have an appearance as of that's being bad.

### Projecting Value

Now that we understand how Hobbes relates desire and aversion to appearances as of good and evil in explicit analogy to the relation between the motions involved in color sensation and appearances as of color, we are in a position to understand Hobbes's "definitions" of good and evil. What emerges is quite different from the standard interpretations. Let us get the passage clearly before us:

[W]hatsoever is the object of any man's appetite or desire, that is it which he for his part calleth *good*: and the object of his hate and aversion, *evil*. . . . For these words of good, evil, and contemptible, are ever used with relation to the person that useth them; there being nothing simply and absolutely so; nor any common rule of good and evil, to be taken from the nature of the objects themselves; but from the person of the man. (6.7)

Commentators who wish to square Hobbes's moral and political

philosophy with his empiricist naturalism frequently interpret this passage as saying that being good is the same thing as being desired. Thus, Gauthier says that Hobbes is saying here that “‘This is good’ means ‘this is an object of desire’.”<sup>25</sup> And Hampton calls it a “baldly subjectivist ethical understanding of ‘good’,” claiming also that Hobbes here defines ‘good’ as “what we desire,” and ‘bad’ as “what we are averse to.”<sup>26</sup> Kavka concurs, referring to it as Hobbes’s “subjectivist definition of good and evil.”<sup>27</sup>

But Hobbes does not say in this passage either that ‘good’ and ‘evil’ mean the same as “what we desire” or “what we are averse to” or that being good or evil are the same thing as being the object of desire or aversion. He says that what we desire we *call good*. Recall his title for chapter 6: “Of the Interiour Beginnings of Voluntary Motions, Commonly Called the Passions; and the *Speeches by Which They Are Expressed*” (emphasis added). When someone calls something good, he *expresses* his desire. He does not say that he desires it. As an analogy consider the difference between expressing a belief and self-attributing it. When I assert that it is raining, I express my belief that it is raining. I do not say thereby that I believe this even if, in asserting that it is raining, I imply that I do.<sup>28</sup> Similarly, when I say that something is good, I express my desire for it. I do not say thereby that I desire it even if I may be taken, at least in some situations, to imply that I do.

Weighing only these considerations, it should be uncontroversial that Hobbes holds that when we call something good, we express rather than assert our desire. And the addition of Hobbes’s explicit analogy between the projection of value in desire and the *illusory* projection of color points fairly directly to a projectivist rather than a subjectivist interpretation. Thus, when Hobbes says that nothing is “simply and absolutely” good or evil and that there is no “common rule of good and evil, to be taken from the nature of objects

---

<sup>25</sup>David Gauthier, *The Logic of Leviathan*, 7. In “Thomas Hobbes: Moral Theorist,” 548, Gauthier again takes this passage as evidence that Hobbes is a value subjectivist, albeit one who differs from contemporary subjectivists in taking desire rather than preference as a measure of value.

<sup>26</sup>Jean Hampton, *Hobbes and the Social Contract Tradition*, 29.

<sup>27</sup>Gregory S. Kavka, *Hobbesian Moral and Political Theory*, 47.

<sup>28</sup>G. E. Moore, “A Reply to My Critics,” in *The Philosophy of G. E. Moore*, ed. Paul Arthur Schilpp (La Salle, Ill.: Open Court, 1942), 561; H. P. Grice, “The Causal Theory of Perception,” in *Perceiving, Sensing, and Knowing*, ed. Robert Swartz (Garden City, N.Y.: Anchor Books, 1965), 444–51.



themselves,” it seems reasonable to interpret him as saying that value is in this respect just like color. Color experience is analogous to desire in that both involve appearances that are as of categorical subject-independent features (color and value, respectively). Sensory experience is as of color, rather than our own states of mind, as the fancies we have when desiring are as of something’s goodness, rather than our state of desiring. There are, however, no such color properties. We only think there are because we “invest” the object “with the fancy.” It is only natural to conclude, therefore, that Hobbes is here making the analogous points about desire and the projective appearance of value. Nothing is “simply and absolutely” good, just as nothing is objectively colored.

But this is not all that Hobbes says. He says also that ‘good’ and ‘evil’ are “ever used with relation to the person that useth them,” and that although there is no rule of good inherent in “the nature of objects themselves,” a rule can be given by “the person of the man.” In the same vein, he says that “*good*, and *evil*, are names that signify our appetites and aversions,” and that “appetite is the measure of good” (15.40). Moreover, Hobbes frequently talks about good and evil in an explicitly relativized way. “The voluntary acts of every man,” he says, invariably aim at “some good to himself” (14.8). And Hobbes defines benevolence as “desire of good to another” (6.22). Don’t all these remarks point in the direction of subjectivism rather than projectivism?

Not necessarily. If Hobbes believes, as he clearly does, that speakers use ‘good’ and ‘evil’ to express their desires, then there is an obvious sense in which he must think they use them “with relation to” themselves even if they don’t *assert* some relation to themselves (as subjectivism requires). As for Hobbes’s other apparently subjectivist remarks, these are best interpreted by bearing in mind two things. First, as with color, the only thing that can enter into an “account” for the projected appearance content is the appearance substratum. So just as, in this sense, color signifies the material motions in our sense organs, so likewise do good and evil signify the material motions of desire and aversion. With color, Hobbes says, “we bring into account the properties of our own bodies, whereby we make such distinction” (4.17). So, similarly, does good signify our appetites in this sense. But this is no more evidence against a projectivist interpretation of Hobbes’s remarks about value than it would be against interpreting him as a projectivist about

color. Or, to put the point the other way round, if one is persuaded by this talk to reject a projectivist interpretation of Hobbes on value, then one should also reject a projectivist interpretation of Hobbes on color.

Second, although Hobbes frequently speaks of things being good *to* people, this can easily be interpreted as referring to what people *think* good or what appears good to them. So when Hobbes says that voluntary action invariably aims at some good to the agent, this can readily be seen to be consistent with projectivism if we understand him as saying that voluntary action invariably aims at something the agent thinks good or that seems good to her. On a projectivist reading, this will be true whenever the agent desires the thing, so we could as easily say that voluntary action invariably aims at something the agent desires. This interpretation has the great merit of following directly from Hobbes's definitions of "voluntary motion" and "desire." As opposed to "vital motion," "animal" or "voluntary motion" invariably involves a "fancy" or "imagination," a "precedent thought of *whither, which way, and what*" (6.1). This is the appearance itself of the "small beginnings of motion, within the body of man," that are the fancy's substratum: "endeavour" or desire (6.2). Based on our earlier inquiry, we know what the content of this fancy must be, namely, that an object of action is good or something that ought to be brought about. Here again, projectivism gives the most natural reading.

Finally, the context in which Hobbes speaks of desire as the "measure of good" and says that 'good' and 'evil' are "ever used with relation to the person that useth them" (15.40) is one in which these remarks cannot be given a subjectivist interpretation without doing violence to Hobbes's argument. What Hobbes is discussing at this point is the kind of conflict that is expressed (and furthered) when people apply 'good' and 'evil' to the same thing.

Nay, [even] the same man, in divers times, differs from himself; and one time praiseth, that is, calleth good, what another time he dispraiseth, and calleth evil: from whence arise, disputes, controversies, and at last war. (15.40)

If all a person says in calling something good is that she desires it (now), then there is no conflict at all in what two people say when one calls it good and another calls it evil, or between what one person says in calling something good at one time and bad at an-

other. The one says that she desires it (then) and the other that he is averse to it, and since these can both be true, no conflict is expressed. There is no dispute over the putatively subject-independent issue of whether something is good or evil. If, however, in respectively desiring and being averse to something, two people see that thing as, respectively, good and evil (that is, absolutely, and not just in relation to them), then the way they see things is in conflict. In having conflicting desires, they see the world in conflicting ways, which they express by attributing the mutually incompatible properties of good and evil. All things considered, therefore, a projectivist interpretation enjoys a substantial preponderance of evidence over a subjectivist one.

### **Desire, Deliberation, and Normative Reasons**

Nonetheless, it will still be true on Hobbes's materialist theory of meaning that all that 'good' and 'evil' can actually signify, in the sense of enter into an account for, are appetites and aversions. There really are no such categorical, objective properties as good and evil seem to be to us when we have desires and aversions. All that exist are the material motions that constitute our endeavors. But if that is so, why wouldn't knowledge of the projective character of ethical judgments undermine ethical thought and practice? Shouldn't Hobbes have believed that in putting forward a projective theory of value judgment he was placing a significant obstacle to his readers' accepting the normative propositions he wished to convince them of in *Leviathan*?

Hobbes never faces this question directly. It is clear enough, however, that any answer he could give would have to do with what, on his view, is the essentially practical character of ethical thought. In this section, I will sketch this aspect of his view. My object is not to show that Hobbes's response would be satisfactory, but simply to lay out what its main lines would have to be.

In considering it, we need not worry about our contemporary issues of whether projectivism leads to the conclusion that all ethical judgments are literally false, but perhaps not in a way that undermines ethical practice, or whether projectivism should be understood noncognitively as holding that ethical judgments are, literally speaking, neither true nor false.<sup>29</sup> The important point is

---

<sup>29</sup>Along the first line, J. L. Mackie argues that it is possible to hold a

that Hobbes believes that ethical thought is simply unavoidable for us because we are *agents*, since, so long as we are alive, we have desires. “[L]ife itself is but motion, and can never be without desire” (6.58). Once we are in the state of desire, Hobbes believes, we are *per force* deliberating. We face the practical ethical question of what to do, of what there are normative reasons to do. Deliberation just *is* “the whole sum of desires, aversions, hopes and fears continued till the thing be either done, or thought impossible” (6.49). And an agent’s will is nothing but “the last appetite or aversion immediately adhering to the action, or to the omission thereof” (6.53). There is no possibility, therefore, that believing projectivism could lead us to give up substantive ethical thought.

This reveals a deep irony in Hobbes’s ethics. Hobbes’s moral psychology is frequently criticized for its woefully inadequate theory of agency.<sup>30</sup> Hobbes’s critics, both early modern and contemporary, argue that agency involves far more than thoughts related to a succession of desires; it requires as well the capacity to gain critical distance on desires and so make them, and thus one’s actions, one’s own.<sup>31</sup> But it is important to recognize despite this that Hobbes treats ethics as fundamentally an agent’s phenomenon. Ethical thoughts—concerning good and ought—are those an

projectivist “error theory” at the level of “second-order” metaethics and nonetheless engage sincerely in substantive, “first-order” normative ethical thought and discourse (*Ethics: Inventing Right and Wrong*, 15–49). For a projectivist view that has greater affinities to traditional noncognitivism, see Simon Blackburn, “How to Be an Ethical Antirealist,” *Midwest Studies in Philosophy* 12 (1988): 361–75.

<sup>30</sup>Not least, by the Cambridge Platonist Ralph Cudworth, in his manuscripts on freedom of the will and autonomy. I discuss this aspect of Cudworth’s views in *The British Moralists and the Internal ‘Ought’*, 130–48.

<sup>31</sup>This was a central objection of Cudworth’s. We should note, however, that although Hobbes makes no place for critical reflection in his account of deliberation, his political philosophy relies on this capacity in various ways. Consider, for example: “For all men are by nature provided of notable multiplying glasses, (that is their passions and self-love,) through which, every little payment appeareth a great grievance; but are destitute of those prospective glasses, (namely moral and civil science,) to see afar off the miseries that hang over them, and cannot without such payments be avoided” (18.20). For *Leviathan* to beneficially affect deliberation, Hobbes here says, agents must be able to put on the critically corrected “prospective glasses” of Hobbesian moral science and revise the appearances that momentary passions produce. I discuss in the next paragraph but one how such remarks can be fit within a projectivist framework. I am indebted here to a reader for the *Philosophical Review*.

agent has in *deliberation*, as her desires and aversions are expressed *seriatim*. We should therefore think of *Leviathan* as a deeply practical work, one that is addressed to agents and designed to inform their practical reasoning.

This response may seem obviously unsatisfactory, however. Granted, if Hobbes is right, we can't stop having it appear to us as though some things are good and ought to be brought into existence and others are evil and ought not. But that doesn't mean that believing projectivism won't alienate us from these appearances. We can't avoid having straight sticks appear bent to us in water either, but we nonetheless can (and do) abstract from these appearances in theoretical and practical reasoning. Why should Hobbes have thought that things would be any different with the illusory appearances created by desire?

Compare, however, the case of color. A projectivist about color can hold that although, strictly speaking, all color judgments are false, our practical purposes are nonetheless better served by speaking and thinking as though they weren't, normalizing our color judgments to the experiences of the normally sighted under normal conditions. We can easily imagine a philosophically sophisticated, projectivist interior decorator whose "first-order" thought and speech about color are regimented in this way. Asked whether a swatch of cloth is really sienna red, she might judge that it isn't but only looks that way owing to the light. Its real color, she might judge, is carmine. At the same time, her philosophical opinion could be that the cloth has no real color, in the sense of the categorical color property it appears to have in color experience, in contrast, say, to the literally straight shape of an apparently bent stick. Such a philosophical position apparently poses no obstacle to facility, even expertise, in color judgments for her practical purposes as a decorator.

Similarly, Hobbes might reply to the objection we are considering by saying that the point and function of ethical terms and concepts is practical, even more so, indeed, than those of color. Consequently, if the color thoughts and speech of a projectivist interior decorator are not undermined, there is even less reason to suppose that the ethical convictions of a (necessarily deliberating) human agent would be, whether he accepted projectivism or not.

All deliberation begins with desire. And "from desire, ariseth the

thought of some means we have seen produce the like of that which we aim at; and from the thought of that, the thought of means to that mean; and so continually, till we come to some beginning within our power" (3.4). But what, again, are the contents of the deliberating agent's thoughts? "In deliberation," Hobbes says, "the appetites and aversions are raised by the foresight of good and evil consequences and sequels of the action whereof we deliberate" (6.57). The desires that Hobbes is talking about here, however, are, not those with which we begin to deliberate, but those arising as a result of our deliberation.<sup>32</sup> It is because we see an action as a means to something we deem *good* that a desire to perform that action can be "raised." But for this to happen, we must already think that certain consequences would be good or evil. It follows that if desires begin deliberation, they must provide us with such thoughts. Under the influence of the desires that begin deliberation, we must be disposed to see some consequences of alternative actions as good and others as evil.

Alternatively, Hobbes writes that the "language of desire, and aversion, is *imperative*, as *do this, forbear that*" (6.55). In seeing that something is necessary to achieve our desire, we can see it as something we must do, since it is necessary to achieve what is good, as we think in having the desire.

To sum up: All deliberation begins in an agent's desires, but this does not mean that they begin in a premise about her desires. The deliberating agent reasons from a premise she accepts in having a desire, not from the premise that she has a desire. And this premise is something normative—that something would be good, that she is to or ought to do something. The agent has these normative thoughts because she has desires. They are the "appearances" of her desires. As it happens, there is nothing in the nature of the objects of her desires that answers to the normative properties she attributes to them in having desires. But Hobbes evidently believes that theoretical knowledge of this fact need not undermine deliberation. So long as the agent has appetites and desires, like the desire for self-preservation, that are entirely independent of the-

---

<sup>32</sup>Note, this means that Hobbes must hold that we can acquire new desires by practical reasoning, that is, by reasoning from the thoughts of *apparent good and evil* we have when we have desires and aversions, respectively.

ory, and beliefs about what will accomplish their objects, she will take action in the thought that in so acting she is achieving good and doing as she ought.

This means that, although he never puts the point this way, Hobbes works with a distinction between theoretical and practical standpoints.<sup>33</sup> Thought and discourse about good and evil encode an agent's view of things in deliberation, from the agent's perspective provided by her desires. Thus, although Hobbes calls "moral philosophy . . . the science of what is good, and evil" (15.40), he clearly must mean that it is a science we engage in from a practical point of view, as we discover how the consequences to be drawn from "the passions of men" bear on the ends that drive deliberative thought.<sup>34</sup>

Again, I am not concerned to argue that this gives Hobbes an effective response to the worry that accepting projectivism should undermine ethical thought and practice or render it unstable in some way. Perhaps projectivism would or should have this effect. I claim only that, for the reasons I have cited, it is clear enough that Hobbes thinks that it wouldn't. Those who disagree about this philosophical issue, therefore, have no reason to reject a projectivist interpretation of Hobbes.

At the same time, we can't infer that Hobbes thought his readers' ethics would be entirely unaffected by accepting his analogy between color and value. To the contrary, I believe that Hobbes thought they *would* be affected and that the effects could be expected to be salutary. Hobbes was a notorious critic, for example, of religious "superstition" and its intellectual expression in the metaphysics of the schools.<sup>35</sup> According to projectivism, when superstitious believers expressed their ethical convictions, they were

---

<sup>33</sup>By "practical" here, I mean the standpoint of agency and deliberation, in contrast with the (theoretical) standpoint we take up when we consider what to believe concerning how things are. We can distinguish a second sense of "practical" within this broadly theoretical standpoint, namely, one involved in ordinary, everyday judgments (color judgments, for example) as opposed to propositions of theory (say, projectivism about color).

<sup>34</sup>More precisely, that part of "moral" science that consists, not in the "apt imposing of names," but in proceeding from these "to assertions made by connexion of one of them to another," must be inherently practical in this way (5.17).

<sup>35</sup>'Superstition' is usually code for Roman Catholicism in Hobbes.

simply giving voice to desires, albeit desires that were highly disciplined by religious ritual and practice. This is not how the faithful saw it, of course. Some, at least, believed their ethical convictions to be grounded in metaphysically real objective values, “final causes,” for example. By Hobbes’s lights, however, this was nothing but “insignificant speech.” Nothing stood behind their ethics but the desires of which their thought was an appearance and their discourse an expression. But that does not mean that all of their desires were on a par in being equally independent of (as Hobbes saw it) “superstitious” theory. Hobbes might well have believed that their ethics were, at least partly, “ideological” in the sense that some of the desires they expressed wouldn’t have existed but for religious rituals that were themselves based upon confused metaphysical doctrines. Convincing his readers of projectivism, he might have thought, could begin a process that, in time, would lead them to see the ideological character of such ethical thought and discourse and tend to undermine it as a consequence.<sup>36</sup>

### Natural Law and Normativity

This brings us to the normativity of the laws of nature. Recall that Hobbes says we only “improperly” call these laws. Really, they are “conclusions or theorems concerning what conduceth to . . . conservation” and self-defense. As we noted at the outset, commentators who stress Hobbes’s naturalism frequently combine this claim with his thesis that human beings unavoidably desire self-preservation (“by a certain impulsion of nature, no lesse than that whereby a Stone moves downward”)<sup>37</sup> to yield a metaethical naturalist interpretation, namely, that laws of nature tell us what we must do to achieve what we (unavoidably) desire.

In a sense, a projectivist will agree. As we saw, however, metaethical naturalism faces a problem in accounting for the normativity

---

<sup>36</sup>Necessary here is a distinction between desires, like thirst (or, as Hobbes views it, the desire for self-preservation), that are relatively impervious to changes in belief, and desires that are not, either because they are based on belief, or because they are conditioned by causal processes (for example, religious rituals) that are sensitive themselves to changes in belief.

<sup>37</sup>*De cive, the English Version, Entitled in the First Edition, Philosophicall Rudiments Concerning Government and Society*, ed. Howard Warrender (Oxford: Oxford University Press, Clarendon Press, 1983), 1.7.



of natural laws, that is, in explaining how they can “dictate” conduct or are necessarily seen to do so. And while Hobbes scruples at “law” he doesn’t at “dictate.”<sup>38</sup> As we put the problem before, the fact that an action will achieve something one desires, even unavoidably, is not necessarily a normative reason for acting—it has no *inherent* bearing on the question of what an agent should do. A drug addict may have a desire for heroin that is as good as unavoidable in her circumstances, but she need not (either actually or rationally) take the fact of her desire as creating a reason to take it.<sup>39</sup> That one desires something is one thing, that it is desirable or worth pursuing, another. And while the latter is intrinsically normative, the former is not.

If, however, Hobbes is a projectivist rather than a metaethical naturalist, we can see why he could think that someone who accepts a proposition about what conduces to self-preservation will take it to have normative force. If Hobbes can assume that his readers desire self-preservation, then he can take for granted both that each will accept that his own preservation is something good, to be sought, *and* that each is already disposed to reason practically from that premise. Consequently, if laws of nature tell us what we must do to preserve ourselves, they inherit the normative force of this end—this is nothing that is entailed by the fact that we have this end, but the normativity we *attribute* to self-preservation in *having* it as end. The thought is not, however, that this makes it true that the agent ought to act as the laws of nature dictate, but that the agent will *take* it that she should insofar as her reasoning is instrumentally rational.<sup>40</sup>

---

<sup>38</sup>Hobbes’s scruples about law relate to his definition of law, which relates it analytically to authoritative command. As with his definition of ‘obligation’ (the result of transferring a right), however, showing that something is a law in this sense is logically independent of establishing its normative force, that is, its power to dictate. On this point see note 45. The point here is that Hobbes *does* think that the “laws of nature” are properly seen as providing normative reasons for acting, even if they are improperly called “laws.”

<sup>39</sup>*In* desiring heroin, she will see it as good in the sense that that is how heroin will seem to her under the influence of her desire. She may, however, not just reject that the fact that she desires heroin is a reason for her to seek it, but also that the appearance that it is good is a reason.

<sup>40</sup>Of course, a heroin addict would similarly (unavoidably) see the means to satisfying her desire for heroin as something she ought to do also. Does a projectivist interpretation of Hobbes’s thought give him any way of distinguishing these cases? As we noted above in considering how

On this interpretation, Hobbes doesn't need to convince his readers that they have this end or that it is unavoidable. As deliberating agents, his readers reason not from the premise that they desire to preserve themselves, but from normative premises they accept in so desiring. If Hobbes is right that we unavoidably desire self-preservation, then, we will regard "theorems concerning what conduceth to . . . conservation" as telling us what we ought to do, as dictating action, whether we believe that we desire self-preservation or not. Thus, Hobbes reasons from the facts of human diffidence, vainglory, and so on to conclusions about what courses of action are necessary for self-conservation and defense. And these conclusions are seen as normative by *Leviathan's* readers, they seem to them to provide normative reasons for acting, because, in having the desire for self-preservation, they see that end, and what is necessary to achieve it, as good and to be done.

This gives us a preliminary gloss on the normativity of laws of nature as Hobbes defines them: "a precept or general rule, found out by reason, by which a man is forbidden to do that which is destructive of his life or taketh away the means of preserving the same" (14.3). In desiring self-preservation, we see what will lead to our destruction as something we should not do. For example, in seeing that the consequences of breaking covenant can be expected to be mortal, we can be brought to accept the third law of nature, "*that men perform their covenants made*" (15.1).<sup>41</sup>

We should stress that this gives Hobbes, as a projectivist, not

---

Hobbes could respond to the objection that projectivism undermines substantive ethical thought, Hobbes might say here that, for practical purposes, we normalize our ethical judgments in various ways. He certainly does say this for the case of judgments in the commonwealth, as we shall consider in the final section. Whether this reply would be satisfactory would, of course, be another matter. It is worth noting, however, that a version of the same problem is faced by a subjectivist interpretation.

<sup>41</sup>But only, based on anything said so far, as holding "other things being equal." For laws of nature to have the sort of "all things considered" force that moral norms are frequently supposed to have, Hobbes would need to show that they counsel necessary means, not just to what (the agent unavoidably judges) is good, but to what (the agent unavoidably judges) is best. We should regard Hobbes as taking steps in this direction when he argues that we must keep covenant, not just to avoid death, but to avoid "the danger of violent death," "continual fear," and the lack of all of the following: "industry," "culture of the earth," "navigation," "use of commodities that may be imported by sea," "commodious building," "arts," "letters," "society," and so on. I am indebted here to Sharon Lloyd.

strictly an account of the normativity of the laws of nature, but of why, on seeing that they spell out what we must do to preserve our lives, we *accept* that we ought to act as they dictate. It is an account of our normative thought and judgment, not of normativity itself. There can be no account of the latter, because Hobbes evidently believes that there really is no such thing, just as there really is no such thing as color. When Hobbes puts forward the laws of nature as normative claims, therefore, he is not arguing from any account of normativity. Nor, indeed, is he arguing *from* his theory of normative judgment. Rather, his projectivism underlies his confidence that his readers can be brought to accept these normative claims by reasoning instrumentally from the thought that preserving their lives is good.

### Remaining Issues

However, this may only provide a first approximation of Hobbes's views on the laws of nature. In this concluding section, I briefly consider a more elaborate, arguably more accurate, version that is in some tension with Hobbes's projectivism, although not, I shall argue, with a projectivist interpretation of Hobbes. I consider, also, how a projectivist interpretation might deal with judgments about what *others* should, or have reason, to do. Finally, I consider what Hobbes says about the normalizing of ethical judgments in a commonwealth and how this is not only consistent with a projectivist interpretation, but actually supports it.

The need to refine the interpretation of the laws of nature presented in the last section is occasioned by Hobbes's reply to the fool.<sup>42</sup> The fool questions the third law of nature, saying "that . . . there could be no reason, why every man might not do what he thought conduced [to his own "conservation and contentment"]; and therefore also to make, or not to make; keep, or not keep covenants, was not against reason, when it conduced to one's benefit" (15.4). The fool makes two claims: one should always do what will be for one's benefit, or what one believes to be so; and keeping covenant is not always for one's greatest benefit.

---

<sup>42</sup>Here I am indebted to Greg Kavka. See his *Hobbesian Moral and Political Theory*, 137–56, and "The Rationality of Rule-Following: Hobbes' Dispute with the Foole," *Law and Philosophy* 14 (1995): 5–34.

Significantly, Hobbes does not actually deny this second claim. He doesn't deny that it may sometimes turn out that breaking covenant makes one better off. What he denies is that this fact, notwithstanding anything that "can be foreseen and reckoned on," ever makes the breaking of a covenant "reasonably or wisely done." This suggests a distinction between subjective and objective rightness or reasonableness that the fool ought to be able to accept. The fool's position might then be that an action is *objectively right or reasonable* if it is actually for the agent's greatest benefit (including her survival), and an action is *subjectively right or reasonable* if it is what she would reasonably believe likely to be so, in the sense, say, of maximizing her expected benefit (her survival included). The issue would then be whether keeping covenant always satisfies this latter condition, with Hobbes maintaining, and the fool denying, that keeping covenant is always subjectively right in these senses.

However, this seems unlikely to be so.<sup>43</sup> Kavka points out that Hobbes's reply to the fool need not rely on this questionable assumption, and Gauthier agrees.<sup>44</sup> For Hobbes may hold that when it comes to covenants, the risks *and* uncertainties are sufficiently great that it never makes sense to rely even on one's best estimates. Kavka argues on this basis that Hobbes is a rule-egoist, and Gauthier maintains that Hobbes's reply invokes a theory of "constrained" rather than "unconstrained" maximization. On these readings, the laws of nature are "precepts" or "general rules" that,

---

<sup>43</sup>Hobbes urges, of course, that the costs of being known to have broken covenant are severe. In the state of nature, no one can expect to survive without the help of confederates who are bound by covenant, and covenant is also the only way out of this nasty and brutish state. Anyone known to violate covenant can therefore expect "no other means of safety, than what can be had from his own single power" (15.5). But these substantial risks notwithstanding, can Hobbes really think that circumstances never arise in which a person may reasonably think he is likeliest to do best by breaching covenant?

<sup>44</sup>Kavka, "The Rationality of Rule-Following." See also, David Gauthier, "Thomas Hobbes: Moral Theorist," *Journal of Philosophy* 76 (1979): 547–59; "Taming Leviathan," *Philosophy and Public Affairs* 76 (1987): 280–98; *Morals By Agreement* (Oxford: Oxford University Press, Clarendon Press, 1986), 157–89. We should note, of course, that Hobbes also believes that threats to life cancel the obligation to keep covenant. I have been assuming, and will continue to assume, that the cases of keeping covenant we are considering are not life-threatening.

if *accepted by the agent*, will best promote his good. What the agent should do is what these rules dictate, not what would be recommended by considerations of his own good directly. If, then, as Hobbes evidently believes, a person does best by accepting a rule requiring the keeping of covenant, regardless of how reasonably one believes one can do better by violation, then she should keep covenant. And this vindicates the third law of nature in the face of the fool's challenge.<sup>45</sup>

This may be Hobbes's position, but if it is, we should note a problem that arises when it is combined with the projectivist account of normative judgment I have attributed to him. Suppose Jones reasonably believes she will do better by breaking a particular covenant. Suppose that she also believes that she will do better in general by accepting a norm that requires her to keep covenants

---

<sup>45</sup>Note how the account of the normativity of the laws of nature I have sketched fits with Hobbes's official view of *obligation*. Hobbes defines obligation as the state one comes to be in by renouncing or transferring a right (14.7), the relevant ones deriving from the "right of nature," which everyone has in the state of nature, of doing whatever "in his own judgment and reason" will promote self-preservation (14.1). A covenant, for Hobbes, is a special form of *contract*, where a contract is a "mutual transferring of right" (14.9). Covenant is a contract in which one person performs his part first, trusting that the other will later perform, as per the contract (14.11).

It simply follows from these definitions that a person is obligated to keep covenant, since 'obligation' just refers to the state resulting from the transference of right in which covenant consists. But, of course, nothing with genuine normative force can follow from definitions alone. So what makes it the case that people *ought* to keep their covenants, that is, act as they are obligated? Obviously, Hobbes recognizes this as a genuine question. Otherwise, there would be no need for the third law of nature, "that men perform their covenants made," or for him to bother with the fool.

There is a neat solution to this problem that is available to Hobbes. Suppose that Hobbes's reply to the fool works. It will then be true that an agent should keep covenant, even if "in his own judgment and reason," he believes, even reasonably, that he would do better by breaking it. It follows that the law of nature now requires that, for this case, he *not* do what "in his own judgment and reason," would most advance his "conservation and contentment." But what that means is that the *right of nature* is *effectively suspended for this case*. And so, by covenanting, the agent will indeed have laid down his right of nature not to violate covenant should he think in his own "judgment and reason" it would benefit him to do so. And since obligation just is the state a person comes to be in by renouncing or transferring a right, he will, by covenanting, have undertaken an obligation to keep his covenant. For an extended discussion of these points, see *The British Moralists and the Internal 'Ought'*, 60–79.

such as the current one. The problem is that it is simply unclear what practical relevance this latter fact has from her standpoint in deliberating about whether to keep or break this particular covenant. For one thing, there seems no reason to think one can come to accept some principle as a norm simply by seeing that one would be better off by doing so. Although this gives her a reason to *want* to accept the third law as a norm, it is not a reason *on whose basis* she can accept it. To put the point in projectivist terms, the desire that the judgment that it would be good to accept the third law as a norm expresses is no desire that is itself part of, or that could directly lead to, accepting it as a norm.

Secondly, although the fact that she would be better off by accepting the third law fails to connect any deliberative alternative before her directly to her desires, considerations that lead her to believe (reasonably) that she would benefit from breaking covenant do. It follows from Hobbes's version of projectivism that she *will* see these as normative reasons for breaking covenant, reasons that, without the desires that actually constitute accepting the third law, will be in no way countered by the fact that she would be better off if she viewed things differently.

That combining projectivism with an indirect or rule-egoist theory of justification is problematic in this way is not, however, evidence against a projectivist interpretation of Hobbes. Hobbes would face the same problem if he were a subjectivist, since, in the kinds of cases we have been considering, it would be true that the agent would (reasonably believe she would) maximize the satisfaction of her desires by violating a norm that it would maximally satisfy her desires to accept. On either metaethical theory, Hobbes would confront the problem that reasons for acting that are grounded in values promoted by being guided by a norm can conflict with reasons the agent must credit insofar as she genuinely accepts the norm.<sup>46</sup> The tension between the projectivist interpretation I have sketched and Kavka's and Gauthier's interpretation of the laws of nature in light of Hobbes's reply to the fool may

---

<sup>46</sup>It is worth noting that Hobbes could remain a projectivist and avoid this particular problem if he held that the agents' views of normative reasons for acting express, not the agent's desires, but her acceptance of norms. For a view of this sort, see Allan Gibbard, *Wise Choices, Apt Feelings* (Cambridge: Harvard University Press, 1990).

reveal a real tension in Hobbes's own thought, not any evidence against interpreting him as a projectivist.

To this point, I have discussed Hobbes's account of normative and evaluative judgment as if these are always made from the agent's standpoint. But what about judgments concerning what *others* ought or have reason to do? How are we to understand these, according to Hobbes?<sup>47</sup> To make the issue vivid, suppose the action is one that would be good for the agent (A) but bad for the observer (O) who is judging what the agent should do. As we have been interpreting Hobbes, if an observer judges that an action would be bad *to* or *for* himself (O), then he judges that it would be something he (O) judges bad *simpliciter*.<sup>48</sup> And, according to projectivism, he makes the latter judgment by a projection of his desires (aversions). Strictly speaking, only the latter is an evaluative or normative judgment. The judgment that something is a bad *to* or *for* him is a judgment about his own psychology, about what he deems bad. Similarly, when O judges that A's act would be good for A, he judges that A deems (or would deem) it good. And strictly speaking this is not an evaluative or normative judgment either, but a judgment about A's psychology.

Now in (projectively) judging that A's act would be (or promote) bad *simpliciter*, the observer judges that he, O, has reason to prevent A's doing it. But this is about O's reasons (for preventing A's action), not A's, as O judges these from the perspective of his desires, that is, as a deliberating agent. But what about a judgment by O that A has reason for or against doing something, that she ought or ought not act? How, according to Hobbes, are we to understand judgments of this kind?

If such a judgment were the same as the judgment that it would be bad *simpliciter* for A to do it, then Hobbes would be committed to thinking that O makes the former judgment *also* from the perspective of his (O's) desires, even if O judges as well that A's act would be good *to* or *for* her (A). As we have just seen, however, the former judgment concerns reasons that O has (for preventing A's acting), not reasons that A has for not so acting.

Recall at this point Hobbes's doctrine that reason "dictateth to

---

<sup>47</sup>I am indebted to an anonymous reader for the *Philosophical Review* for pressing this issue.

<sup>48</sup>See the section "Projecting Value," above.

every man his own good" (15.4). This is usually interpreted as asserting that each agent has normative reasons (and so should) do what is good for him or her. So we might ask in the present case, how can O judge that A's acting would be bad *simpliciter* (and therefore, give O reasons to prevent it) and also judge that A has reasons to act nonetheless, that reason "dictates" acting to her, because acting is for her good, that is, because it promotes what she, A, judges good?

The sense in which, according to Hobbes, reason "dictates" A's acting for her own good is that reason dictates this *to A*. On the projectivist picture of deliberative reasoning we have been sketching, reason "dictates" taking necessary means to what the agent desires, hence judges good. The idea, again, is not that the agent takes the fact that she desires something, or that she judges it good (hence, that it is a good *to her*), as the reason to seek it. Rather, in desiring something, she sees it as good *simpliciter*, and takes the fact that an action is a necessary means to realizing this good as a reason for her to perform that action. Reason thus "dictates" to each agent her own good in the sense that each agent always takes the fact that something is necessary to realize something (she judges) good as giving her normative reason to take those means (perhaps, other things being equal).

Suppose that O and A are engaged in violent conflict, each person's continued living being seen as a threat by the other. Since O and A each desire self-preservation, reckoning from their other beliefs leads them to an aversion to the other's preservation. O judges his continued living a good thing, and takes the fact that, as he sees it, killing A is necessary to assure that as a reason to kill A. This is what reason dictates *to O*. At the same time, O can see that what reason dictates *to A*—in the sense of what normative conclusions reckoning would lead A, or someone in A's position, to draw—is just the reverse. Reason dictates *to A* that she do whatever is necessary to preserve her life, including killing O.

So far everything proceeds as one would expect on a projectivist interpretation. Moreover, interpreting Hobbes in this way makes good sense of what he actually says about reason's "dictates" to other agents. There is, however, a remaining problem. For as we are interpreting him, judgments about what reason dictates to other agents are not genuine normative judgments. They are psychological judgments about what other agents will *take* as normative.



And one can, of course, judge that another will take herself to have reason to do something (even that she will unavoidably do so) and nonetheless judge that she has no reason to do it, or even reason not to do it.

So far as I can see, Hobbes has no good account of judgments of this latter kind. Various possibilities suggest themselves. Hobbes might identify the normative judgment (by O) that A ought or has reason to act with O's (projective) judgment that it would be bad *simpliciter* for A to do so, in the way we considered above. However, this would commit Hobbes to the position that all reasons are agent-neutral, and that hardly seems in the spirit of his views. Alternatively, Hobbes might interpret observer's judgments as involving the taking up of A's deliberative standpoint in imagination. On this account, O would simulate A by imagining himself in A's position, (projectively) judge self-preservation (that is, now A's preservation) good, take the fact that an action is a necessary means to that end as a normative reason to perform it, and, therefore, (projectively) judge that A has reason to kill O. Hobbes says nothing like this, however, and, in any case, it would require a richer theory of the imagination than he provides. In the end, it seems, this is a problem to which Hobbes has no good solution.

Finally, it is worth noting how what Hobbes says about normalizing ethical judgments by the sovereign's dictates in a commonwealth, and by the rulings of a judge established by common consent in the state of nature, remarks that may seem to cut against a projectivist interpretation, do not. Quite to the contrary, in fact. When earlier I quoted Hobbes's "definition" of good and evil (6.7), I ended just before a passage that might seem embarrassing for a projectivist interpretation. After saying that there is no "common rule of good and evil to be taken from the nature of objects themselves, but from the person of the man," Hobbes adds that this is so only "where there is no commonwealth" and that when a commonwealth exists, the rule is given by "the person that representeth it." He continues, moreover, by saying that a rule other than private judgment (hence, desire) is also given by "an arbitrator or judge" when one is set up to settle disagreements by mutual consent (6.7). On the face of it, both seem to conflict with projectivism.

If value judgments express desires and aversions of those who make them, what can they have to do with the will of a sovereign

or of an arbitrator? This conflict, however, is more apparent than real. A projectivist interpretation can accommodate both remarks if it can account for the normative claims embodied in Hobbes's laws of nature, specifically, in the third law, "that men perform their covenants made." If we put aside the complications discussed earlier in this section, we can construct a general Hobbesian argument that begins with the premise that the sovereign, or an impartial judge, rules that X is good or rules that we are to do Y, and ends with the conclusion that X is, indeed, good, or that we are, indeed, to do Y, where this conclusion is interpreted along projectivist lines.

Hobbes believes that the commonwealth is established by a mutual covenant that gives the sovereign authority to rule and his subjects the obligation to be ruled by him. Suppose the sovereign rules that we are to Y. If we accept the central argument of *Leviathan*, then we will judge that we should do what the sovereign commands because we should keep our covenant to be ruled by him. In judging our survival good (as we must, projectively, in desiring it), and reasoning instrumentally from that premise together with the premise that keeping covenant is necessary for our survival, we conclude that we should keep covenant. Adding in the premises that we have covenanted to obey the sovereign and that he rules that we are to Y gives us the conclusion that we are to Y. Thus, beginning with an apparently nonnormative, empirical premise about what the sovereign rules, we end up with a normative conclusion concerning what we should do, a conclusion that expresses the desire we acquire through this instrumental practical reasoning.<sup>49</sup>

The reasoning is parallel in the case of a judge set up by mutual consent in the state of nature. Hobbes believes that so long as there is no reasonable suspicion of the other's noncompliance, we ought to keep covenants, even in the state of nature. Suppose that you

---

<sup>49</sup>Things are more complicated if the sovereign rules that X is good. If subjects are obligated to be ruled by this ruling, then, presumably, they are bound to accept that X is good. But from the fact that they ought to accept that X is good, it doesn't follow that X is, indeed, good. In other words, the desire that will be "raised" by the instrumental reasoning that underlies the judgment that we should be ruled by the sovereign in this instance is not the first-order desire for X, but the second-order desire to desire X.

and I agree to be ruled by a judge to settle a dispute. Prior to the ruling, I desire that you do Y, and you desire that I do Y. So *ex ante*, I judge that you are to do Y and you judge that I am to do Y. Suppose that the judge rules that I am to do Y. From this premise and the premises that our (respective) survival is good (which judgments express our respective desires to live), that each of us can live only if we keep covenant, and that we can keep covenant only if we are ruled by the judge in this matter, both of us can conclude that we should be so ruled. As in the case of the reasoning concerning the sovereign, Hobbes believes that this instrumental practical reasoning will “raise” in each of us the desire that I do Y (6.57). And so, *ex post*, both of us will now judge that Y should be done by me (rather than you).

If we ignore the complications created by Hobbes’s reply to the fool, therefore, complications that create problems for subjectivist and projectivist interpretations alike, Hobbes’s remarks about normalizing ethical judgments by the rulings of the sovereign and established judges can be seen to be completely consistent with a projectivist interpretation. Indeed, once we see why this is so, we can see that it provides further evidence for interpreting Hobbes as a projectivist rather than a subjectivist. For one thing, as we noted before, the sort of dispute described in the last paragraph cannot be properly expressed by subjectivism since each “disputant” will simply be expressing his belief that he wants Y to be done by the other, and there is no conflict between these expressed beliefs. For another, to get a form of subjectivist instrumental reasoning going that might “shadow” the reasoning described above, the agents must not simply desire that they preserve their lives, they must know that they do. Third, as I have been arguing from the outset, it is simply not clear why an agent should conclude from the fact that she desires some end, that she ought (or has some reason) to take the means to realizing it, since her end might be one she ought not have.<sup>50</sup>

Finally, fourth, even if the shadow reasoning were to go through, so that whenever one judged that the sovereign or an established judge ruled that Y is to be done one could reason to the conclusion that doing Y is something I desire and, hence, that doing Y is good, what would make this latter claim true, according to subjectivism,

---

<sup>50</sup>See note 8 above.

would still be the fact that I desire it and not that the sovereign rules it. It is a further virtue of a projectivist interpretation that it avoids this consequence. According to projectivism, it is not the fact that I desire something that makes it good, but whatever facts about it make it desirable (as I appreciate in coming to desire it on that basis). If I follow the above (“shadowed”) reasoning, I judge that I should do Y because the sovereign rules that I should. The desire I acquire to do Y in following the reasoning is not my ground for judging that I should do Y, or what, in my view, makes it true that I should. Rather it is a desire that, according to Hobbes’s projectivism, is expressed phenomenally in an “appearance” of Y’s value or “to be doneness” and linguistically in the “speeches” “doing Y would be good” and “I ought to do Y” (6.10, title). In virtually every respect, therefore, a projectivist interpretation of Hobbes’s *Leviathan* is superior to a subjectivist one.

*University of Michigan, Ann Arbor*