Georgia State University
Digital Archive @ GSU

Philosophy Theses | Department of Philosophy

7-11-2012

Hume on the Nature of Moral Freedom

Getty L. Lustila
glustila1@student.gsu.edu

Follow this and additional works at: http://digitalarchive.gsu.edu/philosophy_theses

Recommended Citation
Lustila, Getty L., "Hume on the Nature of Moral Freedom" (2012). *Philosophy Theses*. Paper 118.

This Thesis is brought to you for free and open access by the Department of Philosophy at Digital Archive @ GSU. It has been accepted for inclusion in Philosophy Theses by an authorized administrator of Digital Archive @ GSU. For more information, please contact digitalarchive@gsu.edu.

HUME ON THE NATURE OF MORAL FREEDOM

by

GETTY L. LUSTILA

Under the Direction of Eric E. Wilson

ABSTRACT

Paul Russell argues that the interpretation of Hume as a *classical compatibilist* is misguided. Russell defends a *naturalistic* reading of Humean freedom and moral responsibility. On this account, Hume holds two theses: that moral responsibility is a product of our moral sentiments, and that our concept of moral freedom is derived from our considerations of moral responsibility. Russell claims that Hume's theory of the passions is non-cognitivist, and thus that his account of moral judgment fails to distinguish between *voluntary* and *involuntary* actions or qualities of mind. He concludes that Hume's account of moral responsibility is inadequate. I argue that Hume has a cognitivist account of the passions. For Hume, our character is judged to be a proper object of praise or censure on account of our ability to partake in a moral community with our fellows. I conclude that Hume does not *naturalize* freedom and moral responsibility, but *socializes* it.

INDEX WORDS: David Hume, Free Will, Moral Responsibility, Causation, Sentiments, Naturalism

HUME ON THE NATURE OF MORAL FREEDOM

by

GETTY L. LUSTILA

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2012

Copyright by
Getty Lustila
2012

HUME ON THE NATURE OF MORAL FREEDOM

by

GETTY L. LUSTILA

Committee Chair: Eric Wilson

Committee: Sebastian Rand

Andrew Altman

Jessica Berry

Election Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2012

DEDICATION

For Joel E. Mann, who made this journey possible; and for Eric E. Wilson, who saw me through.

ACKNOWLEDGMENTS

Thank you Eric Wilson: for your supervision on my thesis. I came to Georgia State having much to learn about being a philosopher. My success in this program is largely a result of your constant vigilance, patience, and wisdom. Thank you Sebastian Rand: for the opportunity to study the works of G.W.F. Hegel. The experience of doing so will stay with me throughout my philosophical career. Thank you Jessica Berry, for your helpful comments on my thesis, and for the interest you showed in my work on Nietzsche. Thank you Andrew Altman: for your commentary, and for your generally encyclopedic knowledge of all things philosophy.

To all my friends and colleagues at Georgia State University: I cannot express the joy I have felt over the last two years. To Alex Economides, Sam Richards, Shawn Murphy, James DiGiovanni, Nick Roberson, Rebecca Harrison, Sam Sims, and the rest of my cohort: for your friendship and support over the last two years. I could not have asked for a better home away from home. And to for all the new friends I made this year: I wish you the best of luck. Also, thanks to my family. To my mother, Stacey Lustila-Siats, for always pushing me; and to my father, Ward Siats, for never letting me pass the buck. More than being a good philosopher, I have always aspired to be a good son, and to be a good role model for my brothers and sisters. I can only hope I have done an adequate job at both. Lastly, thank you Lauren Papiernik for your love and patience: I could not have done it without you.

TABLE OF CONTENTS

ACKNOWLEDGMENTS…....……....v

1. INTRODUCTION……..……..1

2. THE FREE WILL PROBLEM……..……3

1.1 The Bramhall and Hobbes Debate……3

1.2 Hume on the Will, Necessity, and Liberty…….……8

1.3 Classical Compatibilism and Hume…….12

3. NATURALIZING THE PROBLEM: PAUL RUSSELL ON HUME…..……...……16

2.1 The Regularity Theory of Causation……16

2.2 Feeling Naturalism……..…….20

2.3 Scientific Naturalism……....……26

2.4 Does Hume Solve the Free Will Problem?……...31

4. SOCIALIZING THE PROBLEM: HUMEAN MORAL FREEDOM…..…….35

3.1 Socializing or Naturalism?……..……...…35

3.2 The Normativity of Causal Reasoning……..……...39

3.3 The Sentiments and their Reasons……...……44

3.4 Freedom and Moral Community……..50

5. CONCLUSION……..57

REFERENCES……..58

1. INTRODUCTION

The subject of this paper is David Hume's account of moral freedom.[1] Hume is traditionally considered, alongside Thomas Hobbes, to be a classical compatibilist. In, *Freedom and Moral Sentiment: Hume's Way of Naturalizing Responsibility*, Paul Russell argues that the classical compatibilist interpretation of Hume is misguided. In its place, Russell defends a *naturalistic* reading of Hume's account of freedom and moral responsibility. Two claims are central to Russell's naturalist interpretation: first, Hume regards moral responsibility as product of the moral sentiments, not reason; and second, Hume's account of moral freedom, the liberty of spontaneity, is derived from his concept of moral responsibility, and not the other way around (as the classical compatibilist interpretation suggests). Russell gives the label "feeling naturalism" to the first claim, and "scientific naturalism" to the second one. On both of these points, Russell finds parallels between Hume and P.F. Strawson.[2]

With his naturalistic interpretation, Russell claims that he can breathe new life into Hume's philosophy and bolster the larger naturalist research project in contemporary philosophy. Nevertheless, Russell argues that Hume's account of freedom and moral responsibility is flawed. Insofar as Hume claims that moral responsibility is a product of the moral sentiments, Russell argues that he must allow for cognitive-evaluative content in his account of passions. Otherwise, attributions of responsibility could be the result of mere prejudice. Russell also argues that because,

---

[1] I use the terms "free will," "freedom," and "moral freedom" interchangeably. For my purposes, it is important to distinguish these three terms only from "political freedom," which I do not consider in this paper.

[2] In his essay "Freedom and Resentment," Strawson famously argued that our concept of moral responsibility is composed of our reactive attitudes (gratitude, resentment, forgiveness, etc.). See P.F. Strawson, "Freedom and Resentment," in *Freedom and Resentment and Other Essays* (New York: Routledge Press, 2008). Because Russell spends a great deal of time considering the similarities between Strawson and Hume in his book—and with good reason—I do so as well throughout this paper.

for Hume, moral judgments are made about character traits, these judgments fail to track the crucial distinction between *voluntary* and *involuntary* actions or qualities of mind.

While Russell's concerns are reasonable, the force of his criticisms derives from three matters of questionable interpretation. First, Russell construes Humean necessity as a *regularity* theory, and argues that it commits him to the "dual ontology of existence" thesis, or a dualistic metaphysics. Second, Russell takes Hume's account of the passions to be purely *descriptive* and *mechanistic*. Third, Russell's claims that, for Hume, attributions of moral responsibility function independently from considerations of moral freedom.

I argue that Hume can provide adequate responses to Russell. Following Annette Baier, I propose that Hume's concept of causal necessity is *normative*, being based on the correct functioning of our rational capacities.[3] This normative construal of causal necessity allows for a tight connection between reason and the sentiments; which, combined with the insight that the sentiments are already socially reflexive, yields an account of the passions that is both *social* and *infused with reason*. From this, we learn that our character is judged to be a proper object of social praise and censure on account of our moral competency, or our ability to partake in a moral community with our fellows.

On account of the latter insights, I argue that Hume does not *naturalize* freedom and moral responsibility, but *socializes* it. My interpretation is not a radical departure from Russell's naturalistic reading, but a furthering of his project. Russell recognized that Hume's account of moral responsibility was both *social* and *sentimental*, but he failed to reevaluate Hume's theory of moral freedom in light of this fact. My object in this paper is to pick up where Russell left off; modifying his view where it is necessary to do so.

[3] Annette C. Baier, *A Progress of Sentiments: Reflections on Hume's Treatise* (Massachusetts: Harvard University Press, 1991).

2. THE FREE WILL PROBLEM

2.1 The Bramhall and Hobbes Debate

In 1645, amidst the calamity of the English Civil War, two prominent intellectuals exiled from their native soil convened in Paris to discuss the highest of metaphysical matters: the freedom of the will. The first, Bishop John Bramhall, was a theologian equally well-known for his apologetics and his unwavering support of the English monarchy. The second, Thomas Hobbes, was a scientist and philosopher, whose recently published political treatise, *De Cive*, had greatly stirred the European *Respublica literaria.*

Hobbes held a *mechanistic* world-view, influenced equally by Galileo and Vesalius, which stated that both nature and human action are governed solely by the laws of motion.[4] As he notes in the *Leviathan*, "life is but a motion of the limbs, the beginning whereof is some principle within."[5] The principle that Hobbes is alluding to is appetite and fear; for it is our appetites (desires, pleasures, etc.) and fears (aversion, pains, etc.) that serve as the springs of all human action.[6] In addition, Hobbes was a *necessitarian*, as he judged that all events (natural and human) arose necessarily out of their preceding conditions: "the sum of all those things which, being now existent, conduce and concur to the production of that action hereafter, whereof if any one thing now were wanting, the effect could not be produced. This concourse of causes, whereof everyone is determined to be such

---

[4] Galileo Galilei (1564-1642): Italian physicist, astronomer, mathematician, philosopher, and author of *Dialogue Concerning the Two Chief World Systems*. Andreas Vesalius (1514-1564): Dutch physician, anatomist, and author of *On the Fabric of the Human Body*. Galileo and Vesalius are, in many ways, the fathers of the *mechanist* philosophy.

[5] Thomas Hobbes, *Leviathan*, ed. C.B. Macpherson (New York: Penguin Classics, 1968), 81.

[6] "The motion, in which consisteth pleasure or pain, is also a solicitation or provocation either to draw near to the thing that pleaseth, or to retire from the thing that displeaseth; and this solicitation is the endeavour or internal beginning of animal motion, which when the object delighteth, is called *appetite*; when it displeaserth, it is called *aversion*." Thomas Hobbes, "Human Nature," in British Moralists 1650-1800, vol. 1, ed. D.D. Raphael (Indianapolis: Hackett Publishing Company, 1991), 4.

as it is by a like concourse of former causes, may well be called…the decree of God."[7] The concurrence of appetites felt by the human being is sufficient for all subsequent behavior; nothing else is requited, as behavior flows mechanistically from the experience of desire or aversion.

Bramhall rejected Hobbes' mechanism and necessitarianism. As a staunch *libertarian*, Bramhall denied Hobbes' conjecture that human action is governed solely by mechanistic laws. Instead, human action is exercised through "the will," the rational faculty of the mind that gives an agent the "power of election" when she is confronted with potential courses of action.[8] When God offered to grant to King Solomon wisdom in exchange for the possibility of immeasurable riches, Bramhall notes that "it was in his [King Solomon] own power to give it, and it was in his own power to retain it. Yet if he did give it, he could not retain it and if he did retain it, he could not give it. Therefore we may do what we do not, and we do not what we might do. That is, we have true liberty from necessity."[9] Bramhall argues that the reality of human *choice* secures our liberty from the threat of Hobbes' mechanistic philosophy.

The disagreement between Hobbes and Bramhall centers on the controversy regarding the relation between *liberty* and *necessity*. Bramhall deems the task of reconciling liberty with necessity to be to be woefully misguided because "necessity consists in an antecedent determination to one…true liberty consists in the elective power of the rational will."[10] He reasons that "that which is necessitated may agree well enough with my fancy or desires, and obtain my subsequent consent; but that which is determined without my concurrence or consent cannot be the object of my election."[11] Bramhall argues that genuine liberty involves the ability to "elect" actions under conditions where the agent has alternative possibilities, making liberty and necessity logically *incompatible* concepts. This

---

[7] Vere Chappell, ed., *Hobbes and Bramhall on Liberty and Necessity* (Massachusetts: Cambridge University Press, 1999), 20.
[8] *Hobbes and Bramhall on Liberty and Necessity*, 1
[9] *Hobbes and Bramhall on Liberty and Necessity*, 2
[10] *Hobbes and Bramhall on Liberty and Necessity*, 43
[11] *Hobbes and Bramhall on Liberty and Necessity*, 43

seems like an intuitive position. If I could not—under the compulsion of the laws of nature—do other than throw a rock through my neighbor's window, it would surely be absurd to say that I was *free* to throw the rock through her window. As Bramhall points out, "if the supposition be not in the agent's power, nor depend on anything in his power; if there be an exterior antecedent cause which does necessitate the effect, to call this free is to be mad with reason."[12]

Hobbes denies the antagonism between liberty and necessity. He argues that the concepts of liberty and necessity, properly construed, are *compatible* with one another because all actions (voluntary or otherwise) are the product of necessary causes: "whatsoever is produced is produced necessarily, for whatsoever is produced has a sufficient cause to produce it, or else it has not been; and therefore also voluntary actions are necessitated."[13] Hobbes' account of liberty, being thus annexed to necessity, concerns the ability of an agent to do as she *chooses* to do, and not on her ability to *do otherwise*. As he points out, "liberty is [simply] the absence of all the impediments to action that are not contained in the nature and intrinsic quality of the agent."[14] An agent is at liberty to learn how to play the violin if she both has the desire to learn how to play the violin, and the ability to do so. Notice that this does not require that the agent elected the violin playing, for her desire to play the violin may be determined. Nevertheless, on Hobbes' view, we could rightfully say that her learning to the play the violin was voluntary.

Bramhall defines the will as the legislative faculty of the human mind. It serves two functions: *imperatus* (execution of action) and *actus elicitus* (election of action).[15] In the first case, the will dictates commands to the body. Bramhall gives the example of opening and closing one's eyelids. The will compels the body to open or shut the eyelids, causing direct movement in the

---

[12] *Hobbes and Bramhall on Liberty and Necessity*, 9
[13] *Hobbes and Bramhall on Liberty and Necessity,* 38
[14] *Hobbes and Bramhall on Liberty and Necessity,* 38
[15] *Hobbes and Bramhall on Liberty and Necessity*, 9

muscle, which then opens or closes the eyelids.[16] The second function of the will is to elect possible courses of action. Once an action is selected by the will, the relevant lower faculties are determined to comply. Compulsion derives from the will, but the will itself is not compelled; it is the *causa sui* of all voluntary human action.[17] This notion of the will "electing different courses of action" entails that there truly *are* different courses of action available to him. Take the example of a heroin addict named Bob, whose addiction drives him to take drastic actions to secure the drug. Because of his addiction, Bob has the desire to rob a bank, and succeeds in doing so. Bob's end state—acquiring the funds for heroin—was realized as a result of his desire to rob the bank, which resulted from his original desire to inject heroin. For Hobbes, Bob is *spontaneously* free to both shoot the heroin and rob the bank since in both cases he has "done what he chose to do." Bramhall would find the conclusion that Bob *freely* chose to shoot the heroin to be dubious because his original desire to inject the heroin was *necessitated* by his addiction to it. Because Bob did not elect for his desire to shoot the heroin, it would be incorrect to hold him responsible for any action that necessarily followed from the desire.[18] For Bob to be genuinely free, he needs to enjoy a liberty of *indifference*, so that his will may be *undetermined* in its choices.[19]

Hobbes agrees with Bramhall that the will is the cause of all voluntary human action. However, because all actions are caused by the appetites for Hobbes, the will cannot be anything but a type of appetite—in particular, an appetite that compels the agent to act. Hobbes gives a five-step

---

[16] This, as Bramhall notes, is "in truth the act of some inferior faculty subject to the command of the will." He also refers to this act of the will as more "remote." *Hobbes and Bramhall on Liberty and Necessity*, 9.

[17] Bramhall notes that "This [an *actus elicitus*] may be stopped or hindered by the intervening impediment of the understanding, as a stone lying on a table is kept from its natural motion; otherwise the will should have some kind of omnipotence. But the will cannot be compelled to act repugnant to its inclination, as when a stone is thrown upwards into the air; that is both to incline and not to incline to the same object at the same time, which implies a contradiction." *Hobbes and Bramhall on Liberty and Necessity*, 9.

[18] Bramhall notes that, in this case "I do it [the action] not out of any voluntary election, but out of an inevitable necessity," *Hobbes and Bramhall on Liberty and Necessity,* 1

[19] The "liberty of indifference" is a scholastic term used to describe a form of liberty that was absent all causal necessitation. It is commonly contrasted with the "liberty of spontaneity," which refers to a notion of liberty that is consistent with causal necessitation, and holds that all actions are free insofar as they are *uncompelled*.

mechanistic account of how this works. First, we encounter an external object. Second, we experience a mix of aversion and appetite towards the object. Third, this creates the impetus for deliberation, which takes places in an agent through an "alternate succession of appetite and fear."[20] Fourth, this deliberation eventually ceases due to the potency of a certain desire. Finally, this desire results in a course of action. Controversially, Hobbes takes this last link in this chain of deliberation simply to *be* the will: the passions…*proceed not from, but are the will*."[21] It doesn't make sense on Hobbes' account to say that a willed action is not free. All actions that derive from internal desires are free *by definition*. Hobbes' only criterion for a free action is that it has "its beginning in the will."[22] An agent fails to be free only when she is subject to external constraints on her actions. Bramhall ridicules this "madness of reason"; but for Hobbes, freedom of *spontaneity* is the only freedom that we could have.

The debate between Thomas Hobbes and Bishop John Bramhall represents what is at stake in the controversy of the freedom of the will. Hobbes' mechanism allows him to evade questions about the *metaphysical* nature of the will.[23] As a result, his account of the "freedom of spontaneity" is reasonably straightforward. Nonetheless, many of Bramhall's criticisms of Hobbes are convincing. Hobbes does seem to overlook a "deeper sense" of freedom that most people consider necessary for moral responsibility.[24] Also, it seems that we can rationally deliberate about courses of action

---

[20] See Thomas Hobbes, *British Moralists 1650-1800* Vol. 1, ed. D.D. Raphael (Indiana: Hackett Publishing Company, 1991), 15.
[21] Hobbes, *British Moralists,* 16
[22] Hobbes, *British Moralists*, 15
[23] Contemporary philosophers of a more *naturalist* persuasion may be drawn to Hobbes' account of freedom because of his sole reliance on empirical observation in his analysis of voluntary action. Contrary to this, Bramhall seems to give us a reconstruction of voluntary action that appeals to a rationalist faculty psychology—a branch of metaphysics that has few defenders in contemporary philosophy.
[24] Robert Kane captures this intuition perfectly: "It [an account of freedom like Hobbes] does seem to capture the *surface freedom*…These everyday freedoms do seems to amount to (1) the power or ability to do what we want (and the power to do otherwise, *if* we had wanted) and (2) doing so without any constraints or impediments getting in our way…[but] does it also capture the 'deeper' freedom of the *will*?" Robert Kane, *A Contemporary Introduction to Free Will* (New York: Oxford University Press, 2005), 14.

separate from our desires—something for which Hobbes' conception of the will doesn't have room. In what follows, I will consider David Hume's attempt to put this controversy to rest.

2.2 Hume on the Will, Necessity, and Liberty

In 1739, David Hume published the first two books of his *A Treatise of Human Nature*. Nestled in Book Two of the *Treatise* is Hume's discussion of liberty and necessity—a subject he revisited in 1748, with the publication of his *Enquiry Concerning Human Understanding*. Hume is commonly understood as being a Hobbesian about free will. Before further discussing how to interpret Hume on these matters, it is important to get clearer on the definitions he employs of "the will," "necessity," and "liberty" in his treatment of free will.

Hume expresses doubt about giving a concrete definition of "the will": noting that "this impression…'tis impossible to define, and needless to describe any farther; for which reason we shall cut off all those definition and distinctions, with which philosophers are wont to perplex rather than clear up this question" (T, 2.3.1.2). But, like Hobbes, he denies that it is the legislative faculty of the mind. Instead, Hume treats the will as the most salient of the "direct passions"—those impressions "which arise immediately from good or evil, from pain or pleasure" (T, 2.3.1.1). Other direct passions that Hume considers are desire and aversion. These passions are *direct* because they are fundamental experiences of pleasure and pain, the conjunction of which form "the chief spring or actuating principle of the human mind" (T, 3.3.1.2). One can see why Hume would say this. When one feels pleasure—say, upon listening to a Black Flag record—one experiences the desire to listen to Black Flag records. Likewise, when one feels pain—say, upon listening to the newest Yanni release—one experiences an aversion to listening to Yanni. But, the latter considerations make Hume's classification of "the will" as a direct passion puzzling. Does the will *derive* from our affections of pain and pleasure? If so, this classification seems implausible because it would mean we become aware of our will though purely *passive* means; by being affected.

I take it that there is an alternative explanation. Hume defines the will as "the internal impression we feel and are conscious of, when we knowingly give rise to any new motion of our body, or new perception of our mind" (T, 2.3.1.2). This definition leads us to believe that the will is not derived from our feelings of pleasure and pain, but forms the *precondition* of said feelings.[25] Prior to experiencing the pleasure of listening to a Black Flag record, one experiences opening the record sleeve, taking the record out, putting it on the turntable, and placing the needle on the record. Along the way, one is conscious of one's bodily movement, and one's intent to listen to the record. The affection that accompanies one's playing the record will depend on the *content* of playing the record. If it is a Black Flag record that one is playing, one will feel pleasure; if it is a Yanni record, one will feel pain. Either way, prior to the experience of desire or aversion, one is conscious of the basic physical and mental activities that compose one's "playing the record"—that is to say, one is conscious of one's own "will."

In some ways, Hume's account of the will resembles that of Hobbes. After all, Hume does not consider the the will to be, in Bramhall's words, "the root, the fountain" of reason.[26] Say that you deliberately decided to cheat on a test. Bramhall would explain your decision to cheat as follows: your rational will elected your desire to cheat on the test, and then dictated to the relevant faculties that they would enable you fulfill this desire. Hobbes considers this explanation to be nonsense. There is no *election* of the desire to cheat on the test that can be identified as "the will," for the will is simply the appetite that moved you (necessarily) to cheat on the test. Nevertheless, Hobbes and Bramhall both agree that, under the above description, you have willed to cheat on the test. This is not the case for Hume. The will is not something that we *have*, properly speaking, as it is both

[25] This interpretation is strengthened by Hume's comment that "Of all the immediate effects of pain and pleasure, there is none more remarkable than the WILL; and *tho', properly speaking, it be not comprehended among the passions, yet as the full understanding of its nature and properties, is necessary to the explanation of them*, we shall here make it the subject of our enquiry" (T, 2.3.1.2) [emphasis mine].
[26] *Hobbes and Bramhall on Liberty and Necessity,* 43.

constituted by and experienced through our various impressions of thought and action. My will is only transparent to me through its *activity* (in both acting and thinking), making the impression of willing, in some sense, epiphenomenal to acting or thinking itself. Hume remains agnostic about whether particular actions or thoughts are voluntary, as his solution to this problem presented in experience is to appeal to the concept of causal necessity.[27]

Hume drastically parts from both Hobbes and Bramhall in claiming that we do not *discover* necessity, but *experience* it as a vivid impression received from a repeated course of events. Hume points out that the idea of necessity "arises entirely from the uniformity observable in the operations of nature, where similar objects are constantly conjoined together, and the mind is determined by custom to infer the one from the appearance to the other" (EHU, 64). The idea of causal necessity is composed of two parts: the experience of constant conjunction between perceptions, and an inference from one perception to the other. I notice that every time I drop a glass on my hardwood floor, the glass breaks. The dropping of the glass and the glass breaking are two separate perceptions, but they are constantly conjoined with one another: there is never a dropped glass perception without a broken glass perception. But constant conjunction is not sufficient for the idea of causal necessity: "besides that experience…we may establish the relation of priority by a kind of inference or reasoning" (T, 1.3.2.7). It is our experience of temporal priority of the "glass dropping" to the "glass breaking" that warrants the inference that the former caused the latter. The *necessity* of the causal inference derives from the repeated instances of the constant conjunction. We feel that the connection between the glass dropping and it breaking is necessary because the perceptions are

[27] Hume expands on this topic in a footnote in the first Enquiry. He points out that, in performing actions we "have a false sensation or seeming experience…of liberty or indifference." At the same time "we may observe, that, though, in reflecting on human actions, we seldom feel such a looseness, or indifference, but are commonly able to infer them with considerable certainty from their motives" (EHU, 72, note 1). Our *feeling* of willing is inconclusive evidence for our *voluntary* willing. It appears that, for Hume, we need a stronger criterion than phenomenal experience to decide whether or not a particular action or thought is voluntary. This latter point sets his treatment of "the will" at odds with the accounts that Bramhall and Hobbes give.

constantly conjoined. As Hume points out, "this [necessary] connexion, therefore, which we feel in the mind, the customary transition of the imagination from one object to its usual attendant, is the sentiment or impression from which we form the idea of power or necessary connexion" (EHU, 59).

It is easy to see how Hume applies the concept of causal necessity to natural objects. His more divisive claim is that the idea of causal necessity applies to human action as well. For Hume, we feel a necessary connexion between our beliefs and actions: "it appears not only that the conjunction between motives and voluntary actions is as regular and uniform as any part of nature, but also that this regular conjunction has been universally acknowledged among mankind" (EHU 88). When I see my friend go to get a drink from the water fountain, I infer that she is thirsty. I automatically make this inference because I have seen others do so, and have done the same myself. Never do I find myself in a state of perplexity at the situation, for "the union betwixt motives and actions has the same constancy, as that in any natural operations" (T. 2.3.1.14). This transition from natural to human objects may seem to be unjustified, but remember that Hume provides us a psychological account of causal necessity. But because we make use of the same perceptual faculties to cognize both natural and human objects, Humean necessity (constant conjunction and causal inference) will apply equally to both. So what is it that enables human agents to enjoy liberty?

Hume offers his canonical statement of liberty in the *Enquiry Concerning Human Understanding*: "by liberty, then we can only mean *a power of acting or not acting, according to the determinations of the will*; that is, if we choose to remain at rest, we may; if we choose to move, we also may. Now this hypothetical liberty is universally allowed to belong to everyone who is not a prisoner and in chains" (EHU 73). Liberty requires that our actions be representative of our desires. If we enjoy liberty, then we have the ability to act in accordance with out desires. Hume's account of liberty does not exclude *causation*, as the preconditions for liberty require that the action be caused by the agent's desires; but

it does exclude *compulsion*, where this means impediments on the agent's action separate from the agent's desire. The conception of liberty Hume expresses here simply *is* the "liberty of spontaneity," as discussed in the previous section. Like Hobbes, Hume argues that the liberty of spontaneity is the only plausible way of conceiving liberty. Bramhall's resistance to this idea lies in his conflation of causation with compulsion: whereas the latter refers to *external* impediments to an agent exercising her will, the former refers to the *internal* mechanisms that make it the case that she acts at all. Rather than constraining us, the necessary connection between our actions and motives gives us (as well as other people) the ability to understand our actions as genuinely *ours*. Hume notes that "liberty [of indifference], by removing necessity, removes also causes, and is the very same thing with chance" (T, 2.3.1.18). Our practices of holding each other responsible depend on the idea that our actions are necessarily linked with our motives; for if we cannot link motivation to action, we can never consider someone to be genuinely responsible for their actions.

2.3 Classical Compatibilism and Hume

When it comes to the problem of free will, the vast majority of philosophers in the twentieth century have taken Hume to be Hobbesian. The label that philosophers have introduced to capture both of their views is *classical compatibilism*. As we established, a compatibilist like Hobbes or Hume maintains that liberty, properly construed, is compatible with the doctrine of necessity. Classical compatibilism, represented most definitively by Hobbes, defines liberty as the ability to choose *spontaneously*. Contemporary philosophers present Hobbesian liberty as having two parts: a "power or ability requirement" and an "absence of constraint requirement."[28] For the classical compatibilist, S is free to Φ *iff* S has the ability to Φ and S's attempt to Φ is not interfered with. Because classical compatibilism defines "liberty" within the bounds of "necessity," an agent's voluntary action is *always* subject to the laws of causal necessity. What makes an agent's action voluntary is not that it

[28] Robert Kane, *A Contemporary Introduction to Free Will*, 14-15

lies outside the jurisdiction of necessary causal laws, but that the determination of his action flows from *his* will and not from the will of others. Instead of directly answering the incompatibilist challenge to provide a genuine account of action election, classical compatibilism attempts to modify the grounds of the debate by shifting the conversation from voluntary *willing* to voluntary *action*. Because of this, the incompatibilist charges that the compatibilist has missed the point: if one cannot elect one's action then one's action cannot be considered voluntary. In response, the compatibilist argues that the incompatibilist is simply confused about the nature of liberty: to will voluntarily simply *is* the ability to act as one chooses.

How does classical compatibilism apply more specifically to what Hume says about free will? Paul Russell presents three "arguments" that Hume puts forth in both the *Treatise* and the first *Enquiry* to highlight the spirit of the classical compatibilist interpretation. Russell calls the first argument the *necessity argument*—the conclusion of which states that causal necessity is nothing but "the constant conjunction of objects and the inference of the mind from one object to the other."[29] The necessity argument is conjoined with two *liberty arguments*, the first of which Russell refers to as the *spontaneity* argument. The spontaneity argument claims that "voluntary action is to be distinguished from involuntary action not by the *absence* of necessitation (as is suggested by the liberty of indifference), but rather by a different *type* of cause."[30] The second liberty argument is the *antilibertarian* argument, which concludes that the liberty of indifference, if actual, would abolish the possibility of free actions.[31] The three arguments, taken as a whole, constitute Humean classical compatibilism according to Russell. Liberty and necessity are taken to be compatible with one another because liberty, properly construed, consists solely in the ability to *do as one chooses*. My decision to learn how to ride a bike may be causally necessitated by my desire not to be made fun of

[29] Russell, *Freedom and Moral Sentiment*, 11
[30] Russell, *Freedom and Moral Sentiment*, 13
[31] Russell, *Freedom and Moral Sentiment*, 12

by my peers; but insofar as my desire is genuinely *mine*, the fact that this desire necessitated my decision to learn how to ride a bike does not render my action involuntary.[32] Likewise, consider if my desire not to be made fun of by my peers did not causally necessitate my decision to learn how to ride a bike. What if my action could not be causally linked to any of my desires, wishes, intentions, etc? Would we choose to call my action voluntary? Hume does not think we would. A tight causal link between my motives and my action seems, perhaps counterintuitively, to be *required* for liberty.

Russell claims that the classical compatibilist approach to free will is *rationalist* in two senses.[33] First, classical compatibilism is concerned with the *logic* of concepts we make use of in discussion about freedom and moral responsibility, a strategy Russell says involves "a priori reflections about the *meanings* of the *terms* involved."[34] That is to say, the classical compatibilist makes use of conceptual analysis in his treatment of free will, with the aim to clearing up any linguistic confusion that impedes the debate. As Russell points out, the classical compatibilist argues that it is "because we confuse [our concepts of] causation with force and compulsion we mistakenly conclude that freedom (and responsibility) requires the absence of causation and necessity."[35] The second sense in which Russell takes classical compatibilism to be rationalist is in its treatment of the "problem of moral responsibility" as *reducible* to the "problem of free will." For the classical compatibilist, to be morally responsible just *is* to be free. As Russell puts it, advocates of classical

---

[32] As Hobbes points out: "liberty is the absence of all the impediments to action that are not contained in the nature and instrinsical quality of an agent," *Hobbes and Bramhall on Liberty and Necessity*, 38.

[33] Russell, *Freedom and Moral Sentiment*, 171. Here, it is important to not conflate the form of "rationalism" that Russell is concerned with and the "rationalism" that people often speak of in discussions of Early Modern philosophy. Russell is *not* claiming that Hume is a rationalist in the sense that Descartes, Leibniz, and Spinoza are thought to be. Instead, he argues that Hume *qua* classical compatibilist interpretation is a rationalist in the sense that he approaches the free will problem with the tool-kit of something like conceptual analysis. Russell contrasts this latter form of rationalism with *naturalism*, which "relies on empirical observations regarding the nature and circumstances, or moral psychology, involved in actual ascriptions of moral responsibility" in its treatment of free will. See Russell, *Freedom and Moral Sentiment*, 14.

[34] Russell, *Freedom and Moral Sentiment*, 14

[35] Russell, *Freedom and Moral Sentiment*, 4

compatibilism believe that "any adequate interpretation and discussion of moral responsibility must begin with, or develop out of, some relevant and appropriate account of the nature of moral freedom."[36] In order to be responsible, we have to be free; and in order to be free, we have to act voluntarily. Russell claims that Hume parts from classical compatibilism on both of these points: approaching the problems of freedom and moral responsibility from the perspective of *naturalism*, and not *rationalism*. I will discuss Russell's alternate interpretation in the following chapter.

The popularity of classical compatibilism has waned in recent years due to pressure from more sophisticated forms of incompatibilism. And one can see why. Going back to the example I used previously in my discussion of Bramhall: it doesn't seem to be the case that Bob the drug addict is free in his robbing of the bank simply because he both acts from a desire and enjoys non-interference in his action. Bob suffers from an addiction that blocks his ability to *elect* possible desires and actions. Any theory of freedom that considers Bob's spontaneous action to be voluntary must be impoverished. It follows that *if* we take classical compatibilism to be representative of Hume's position on freedom and moral responsibility, we should conclude, along with Russell, that "Hume's thinking on this subject is now dated and somewhat passé."[37]

[36] Russell, *Freedom and Moral Sentiment*, 11
[37] Russell, *Freedom and Moral Sentiment*, 4

3. NATURALIZING THE PROBLEM: PAUL RUSSELL ON HUME

3.1 The Regularity Theory of Causation

In his book, *Freedom and Moral Sentiment: Hume's Ways of Naturalizing Moral Responsibility*, Paul Russell argues that the classical compatibilist interpretation of Hume is flawed on two grounds. First, he claims that Hume's concept of causal necessity (constant conjunction and inference of the mind) *undermines* the classical compatibilist account of liberty (liberty of spontaneity). Second, Russell argues that the classical compatibilist interpretation fails to recognize the central role the passions play in Hume's account of moral responsibility. On account of these criticisms, Russell concludes that Hume eschews the *rationalism* of classical compatibilism for a more *naturalistic* approach to the problems of free will and moral responsibility. In this section, I will only be concerned with Russell's first criticism of the classical compatibilist interpretation.

Russell argues that Hume holds a *regularity* theory of causation, which he differentiates from a *metaphysical* theory of causation—where the latter is understood as giving an account of the causal structure of the world, and the former is taken to give an account of our *perceptions* of that world.[38] One can see why Russell would attribute a regularity theory of causation to him; after all, Hume claims that "the mind has never anything present to it but the perceptions, and cannot possibly reach any experience of their connexion with objects" (EHU, 119). Hume explicitly refrains from making judgments about various "powers" or "forces" that underlie our perception of constant conjunction (EHU, 50). And so, following Russell's reasoning, Hume simply cannot accept a *metaphysical* theory of causation. Hume appears to be much more interested in the *psychology* of causation and causal inference. As Russell points out, for Hume "the connexion we think of as holding between cause and effect turns out to be the same as that which holds between our

[38] Russell, *Freedom and Moral Sentiment*, 47

perceptions whereby one idea naturally introduces the other. Thus, the necessary connexion turns out to be an association between our perceptions."[39] It is an account of how this process of association *functions* that Hume is concerned to give in his treatment of causation (EHU, 59).

Classical compatibilism, according to Russell, presupposes a *metaphysical* theory of the causation. The classical compatibilist account of freedom, the liberty of spontaneity, distinguishes between those causes that are *internal* to the agent (desires, intentions, etc.) and those that are *external* to him (laws of nature, etc.). For a classical compatibilist to consider an action to be *voluntary*, it must be the case that the action was the result of a *necessary*, internal cause. Russell claims that because Hume defines causal necessity in terms of our perceptions, and not objects, our causal inferences are limited to saying that two perceptions are regularly conjoined. For this reason, Russell argues that Hume does not have the resources to establish that a particular action was the result of an agent's motives or desires, and thus he is unable to discover whether or not the action in question was truly voluntary: "the liberty argument presupposes that any adequate theory of responsibility must establish that agents produce or determine their actions, and are thereby, connected with their deeds. The necessity argument suggests that there exist only constant conjunctions between these objects (i.e., willings and actions) and that these constant conjunctions do not reveal or uncover any power or agency in any cause, nor any connexions between cause and effect."[40]

Russell argues that a result of Hume's regularity theory of causation is that it commits him to the "double-ontology of existence" thesis. There are two components to this view: first, that "we naturally suppose that there exists a material world of bodies and that they operate on one another 'independent of our thought and reasoning'"; and second, that "in ordinary life we 'confound perceptions and objects' and thereby naturally 'transfer' those connexions which we 'feel' between

[39] Russell, *Freedom and Moral Sentiment*, 31
[40] Russell, *Freedom and Moral Sentiment*, 55

our perceptions to the bodies (i.e. bodies) themselves."[41] We believe that a world exists separate from our perceptions of it, but we understand that we only have knowledge of this world through our perceptual faculties. Our belief in the external world is central to our understanding, since "we must take it for granted in all of our reasonings" (T, 1.4.2.1). To put it in Strawsonian terms, the belief in an external world is a *commitment* of ours.[42] As a result of this commitment, and because we recognize (at least when we are being careful with our reasoning) that our perceptions fail to tell us anything about the world apart from the functioning of these faculties, we are stuck with two realities.[43]

At times, Hume does seem to assert the double-ontology of existence thesis: "philosophy informs us, that every thing, which appears to the mind, is nothing but a perception, and is interrupted, and dependent on the mind; whereas the vulgar confound perceptions and objects" (T, 1.4.2.14). However, I propose that Hume' distinction between perceptions and their objects is *epistemological* and not metaphysical. Consider his discussion of skepticism in the *Enquiry Concerning Human Understanding*. Hume derides Cartesian skepticism because it eliminates the possibility of human knowledge: "the Cartesian doubt, therefore, were it ever possible to be attained by any human creature (as it plainly is not) would be entirely incurable; and no reasoning could ever bring

---

[41] Russell, *Freedom and Moral Sentiment*, 46

[42] P.F. Strawson speaks of Hume's affirmation of the existence of an external world as an "unavoidable natural conviction, commitment, or prejudice [that is] ineradicably implanted in our minds by Nature." P.F. Strawson, *Skepticism and Naturalism: Some Varieties* (New York: Columbia University Press, 1987), 19. Earlier, in the same work, he notes that "his [Hume's] point is really the very simple one that, whatever arguments may be produced on one side or the other of the question, we simply cannot help believing in the existence of body," 11.

[43] One need not conclude with Russell that there is a *metaphysical* distinction between "objects" and "bodies" for Hume. Instead, one could claim, along with Strawson, "Hume, then, we may say, is ready to accept and to tolerate a distinction between two levels of thought: the level of philosophically critical thinking which can offer us no assurances against skepticism; and the level of everyday empirical thinking, at which the pretension of critical thinking are completely overridden and suppressed by Nature," *Skepticism and Naturalism: Some Varieties*, 13. Later, Strawson refers to these "levels of thought" as standpoints that one can take, neither of which is opposed to one another: "the appearance of a contradiction arises only if we assume the existence of some metaphysically absolute standpoint from which we can judge between the two standpoints," 39. The distinction between the two levels of thought—skepticism and naturalism—is not a *metaphysical* distinction then. I will make another version of this point against Russell in Chapter 3.

us to a state of assurance and conviction upon any subject" (EHU, 116). Hume finds this conclusion distasteful, and proposes a more mitigated form of skepticism—represented by the Ancient Greek schools. The latter form of skepticism urges that "the senses alone are not implicitly to be depended on; but that we must correct their evidence by reason, and by consideration, derived from the nature of the medium, the distance of the object, and the disposition of the organ, in order to render them, within their sphere, the proper *criteria* of truth and falsehood" (EHU, 117). Surely this more ancient form of skepticism recognizes a gap between perceptions and their objects, but it is an epistemological gap, and not a metaphysical one. The skeptical schools are meant to teach us moderation in our causal inferences; not that such inferences are impossible.

The problem with the double-ontology of existence thesis, on a philosophical level, is simply stated: if the distinction between "objects" and "perceptions" is a *metaphysical* one, and not merely an *epistemological* one, then we have no criterion by which to judge the extent to which our perceptions of objects "hook up" with the objects themselves. Further, Russell's interpretation of Hume's account of causation as a regularity theory is equally problematic on *Humean* grounds. If causal inferences lie at the base of all knowledge claims about matter of fact as Hume thinks they do (T, 1.3.2), and our inferences can only pick out "regularities," then we have no rule to demarcate *correct* causal inferences from *incorrect* ones. Yet, Hume assumes that "there be no such thing as chance in the world," and thus commits himself to the view that there are correct causal inferences to be made (EHU, 46). I explore this topic further in Chapter 3, section 2.

3.2 Feeling Naturalism

Russell argues that the classical compatibilist interpretation fails to take into account the central role the moral sentiments play in Hume's account of moral responsibility. The moral sentiments are central to what Russell calls Hume's *naturalization* of the problems of free will and moral responsibility. Russell parses naturalism in two ways: scientific naturalism and feeling naturalism. Whereas classical compatibilism claims that an agent is morally responsible if her action is voluntary, *feeling* naturalism maintains that an agent is morally responsible if she is the object of a sentiment of approbation or disapprobation. Likewise, while classical compatibilism approaches the problem of free will with the tools of conceptual analysis, *scientific* naturalism undertakes an "*a posteriori* project of *describ*[*ing*] the circumstances under which people are felt to be responsible."[44] In this section, I will be concerned with feeling naturalism component of Russell's interpretation. Central to Russell's interpretation of Hume as a "feeling naturalist" is his reading of Hume's account of the passions. Before we get into Russell's interpretation of the Humean passions, it is necessary to lay some groundwork.[45]

Hume distinguishes between two kinds of passions: calm and violent. *Calm* passions concern "beauty and deformity in action, composition, and external objects," while the *violent* passions are those of "love and hatred, grief and joy, pride and humility" (T, 2.1.1.3). The difference between the two is one of *vivacity*. Hume makes a further distinction between direct and indirect passions. The *direct* passions (desire, aversion, grief, joy, etc.) are those which "arise immediately from good or evil, from pain or pleasure," while *indirect* passions (pride, humility, ambition, vanity, etc.) are those that

[44] Russell, *Freedom and Moral Sentiment*, 65

[45] The following account of the passions, though detailed, will become especially important for my treatment of Russell's interpretation of Hume in Chapter 3.

"proceed from the same principles, but by the conjunction of other qualities" (T, 2.1.2.4).[46] The qualities that Hume speaks of here are explained by the principles of association: contiguity, resemblance, cause and effect.[47] One can see why Hume would make this distinction between direct and indirect passions. If I see an apple on a table and remember that I receive pleasure from eating apples, *ceteris paribus*, I may *desire* to grab the apple and eat it. However, I will never experience *guilt* at my eating the apple unless it was the case that the apple belonged to someone else: that is to say, unless my passion of desire can be mediated through another person's feeling of disapprobation at *wrongly* eating the apple (i.e., they did not give me permission to eat it, etc.). Indirect passions are then, at bottom, social passions.[48]

The indirect passions consist of two parts: the *object* of the passion (self, other, natural object), and the *sensation* that accompanies the passion (pain or pleasure). A specific passion is produced by a "double relation of ideas and impressions," whereby the "cause, which excites the passion, is related to the object, which nature has attributed to the passion; the sensation, which the cause separately produces, is related to the sensation of the passion" (T, 2.1.5.5). Say I feel pity for a friend who has just lost her father. The *sensation* (impression) of the passion is pain, while the *object* of

---

[46] Hume also includes "the will" in his account of the direct passions (2.3.1). But, as I argued in the previous chapter, there is reason to believe that he takes the will to be a *necessary component* of one's experience of the direct passions, but not as one of the direct passions themselves.

[47] See *Treatise*, 1.1.4. Hume claims that the relationship *between* passions (say, between malice and pride) functions according to associations of resemblance: "nature has bestow'd a kind of attraction on certain impressions and ideas, by which one of them, upon its appearance, naturally introduces its correlative" (T, 2.1.5.10). That said, all indirect passions "require the existence of some foreign object, and that the organs, which produce it, exert not themselves like the heart and arteries, by an original internal movement" (T, 2.1.5.7). It follows from this that associations of cause and effect are required for the function of our indirect passions, and so the latter are then dependent on the tools of causal inference. One cannot feel pride about the state of one's garden unless it is also true that one believes that (1) one's garden is one's own, and (2) that there is a necessary connexion between the state of one's garden and it being one's own. Otherwise, the passions of pride would be *improper*.

[48] Annette Baier makes the case that direct passions are *also* social. She finds evidence for this in the fact that Hume treats the indirect passions before the direct passions in the *Treatise*, which leads her to believe that Hume treats the direct passions as "abstractions" from our more social indirect passion. This is not to say that Baier thinks direct passions *reduce* to indirect passions; it is just that the former are dependent on the latter for their existence. I will pursue this line of thought further in Chapter 3. See *A Progress of Sentiments*, 133-135. Baier makes a similar point about the relation between impressions and ideas: see 34-35.

the passion is my friend (idea). The *cause* of my pity was her father dying. This cause produced an independent feeling of pain (impression), which was then bolstered by the fact that the death was *of* my friend's father (idea). The relation of these two ideas and impression created the passion "pity for my friend on account of the death of her father" in me. Notice that Hume is claiming that we feel pain on two fronts: the fact of death, and the fact that it was my friend's father who died. This may seem unlikely, until one takes into account the central role that sympathy plays in Hume's account of the indirect passions: "no quality of human nature is more remarkable, both in itself and in its consequences, than that propensity we have to sympathize with others, and to receive by communication their inclinations and sentiments, however different from, or even contrary to our own" (T, 2.1.11.2). Because we naturally sympathize with our fellows, news of the death of someone's father will cause us pain, even if the pain is minimal. However, this pain receives "additional weight and authority" when affixed to the idea of my friend (T, 2.1.12.19).

The *moral sentiments* of approbation and disapprobation—the constituents of Humean moral judgments—are classified as calm, indirect passions.[49] These passions comprise the distinction between virtue and vice: when one's sentiment of approbation is directed at the conduct or character of another agent, we judge that the agent is *virtuous*; when an agent is the object of our disapprobation, we deem him vicious. But, while our moral sentiments may take another agent's *action* as its object—say, an agent's giving to charity—Hume claims that "if any *action* be either virtuous or vicious, 'tis only as a sign of some quality of character. It must depend upon durable principles of the mind, which extend over the whole conduct, and enter in the personal character" (T, 3.3.1.4). That is to say: when we deem someone virtuous on account of her giving to charity, we

[49] There is controversy concerning whether or not Hume can rightfully speak of "moral judgments," as he argues at length that "the approbation of moral qualities most certainly is not deriv'd from reason, or any comparison of ideas; but proceeds entirely from a moral taste, and from certain sentiments of pleasure or disgust" (T, 3.3.1.15). I will speak of moral judgments because this is the language that Russell employs. I also happen to believe, contrary to what *may* seem to be the case, that Hume's account of responsibility attribution is best cast in the language of moral judgment.

are making a judgment about her character; namely, that she is benevolent. If we knew this agent to be a *prideful* person, we may judge her conduct differently; we may judge that she gave to charity out of self-interest. In the latter case, Hume thinks that we would be hesitant to praise her action. This fact shows that our moral sentiments are ultimately directed at an agent's character, and though "actions are, indeed, better indications of a character than words, or even wishes and sentiments; but 'tis only so far as they are such indications, that they are attended with love or hatred, praise or blame" (T, 3.3.1.5).

According to Russell, Hume's account of the passions is a "feeling theory."[50] On this view, the violent, *amoral*, passions (love, hate, etc.) are distinguished from the calm, *moral*, passions based on how they *feel* to the agent.[51] Say I feel love towards myself, and hatred towards another. Russell reads Hume as making the following point: my self-directed love will *feel* differently than my other-directed hatred, and so we can know that love and hatred are different passions. Further insight into the relation between the calm and violent passions will be received from considerations about the kind of effect they have on us, for we may "*infer* them from the behavior they have aroused in us."[52] The latter is centrally important to understanding the *calm* passions because they are "so soft an emotion, as to become, in a manner, imperceptible" (T. 2.1.1.3). It follows that, though we can distinguish between violent and calm passion on account of the *feelings* they produce in us, when it

---

[50] Russell, *Freedom and Moral Sentiment*, 88

[51] Russell notes that "Hume's feeling theory of emotion commits him to the view that the moral sentiments *feel* different from all other passions and that this sui generis constitutes the very nature or essence of moral sentiment." Russell, *Freedom and Moral Sentiment*, 89. I take it that Russell's classification of Hume as a proponent of the "feeling theory" here is misguided because he takes "feeling" to be phenomenological and not somatic. The traditional formulation of the "feeling theory," presented by William James and Carl G. Lange, stated that emotions were to be defined on the basis of certain *effects* they were having on the body: "we feel sorry because we cry, angry because we strike, afraid because we tremble, and [it is] not that we cry, strike, or tremble, because we are sorry, angry, or fearful, as the case may be." See William James, "What is an Emotion?" *Mind*, 9: 188–205. The "feeling theory" maps onto Russell's second formulation of Hume's view: that the passions may be defined on the basis of the effect they have on the body, but not the first formulation: that emotions are *sensations*. Because the latter view is utterly inane, we have reason to believe that Russell is engaged in hyperbole against Hume.

[52] Russell, *Freedom and Moral Sentiment*, 90

comes to *kinds* of calm passions, we can only differentiate between them on the basis of physiological or behavioral evidence.

Likewise, when we are trying to *infer* the existence of a particular passion (either violent or calm) in another agent, we will only have physical evidence to draw our inferences from, as we cannot share in the feelings of another. This is precisely why Hume takes the consideration of conduct to be so centrally important to understanding others: "when I see the effects of passion in the voice and gesture of any person, my mind immediately passes from these effects to their causes, and forms such a lively idea of the passion, as is presently converted into the passion itself" (T, 3.3.1.6). It is the passion that *motivated* the person's conduct, and not the conduct itself, that produces a feeling of approbation or disapprobation in us. If we see someone kick a dog, and infer that they did so in order to get the dog out of the road (lest it get hit by a car), we will feel differently about him than if we infer that he kicked the dog out of malice.

We do not receive praise or censure on the basis of our actions themselves (even our intentional ones). Instead, it is our actions that reveal stable or enduring character traits that subject us to moral sentiment. On Russell's reading of Hume, we are "responsible *for [our] character* in the sense that it is (beliefs about) character that generates moral sentiment."[53] This does not require that we have actively *shaped* our character: we can be held responsible for who we *are* regardless of whether or not the traits that constitute us were voluntarily chosen. Russell finds evidence for this reading in Hume's treatment of *excusing conditions*, or the conditions under which we let someone off the hook, morally speaking. He claims that "Hume holds that excusing conditions must be explained or accounted for with reference to character."[54] So, when we excuse someone for her conduct, we must always do so on basis that the action in question was incongruous with the character that we know the agent to have. For example, if your friend was known to be kind, we may excuse a

[53] Russell, *Freedom and Moral Sentiment*, 124
[54] Russell, *Freedom and Moral Sentiment*, 102

particular outburst he made under pressure because we understand that it did not reflect the kindness we know him to have. Though we could let your friend off for his "out of character" conduct, we could not let him off the hook for his *character*. According to Russell then, Hume is not concerned with the voluntary or involuntary nature of the character in question, but with the "character doing the acting."[55]

Nevertheless, Russell argues that this does not mean that Hume does not concern himself with moral freedom. If a person were forced to shoot the president because an evil neuroscientist was controlling the firing of his neurons, we would consider her action involuntary, and her character untainted. But, as Russell points out, the moral sentiments reach beyond matters of conduct, and take character traits or qualities of mind as their objects as well: "feeling and desires, even though they might never engage the will or direct our intentions, are distinct and vital indications or signs of a person's character and qualities of mind, and, as such, they are quite capable of arousing moral sentiments toward that person."[56] It follows from this that there can be moral judgments about involuntary, or non-willed, aspects of one's character.[57] Russell takes this fact to prove that the liberty of spontaneity is not a *necessary* condition of being taken as object of the sentiments of approbation or disapprobation, and so that "the sphere of responsibility, therefore, extends beyond the sphere of action and liberty of spontaneity."[58]

I do think that Russell's criticism of the classical compatibilist interpretation is right. Clearly, Hume's account of the passions is central to his theory of moral responsibility. However, I disagree with him on two fronts. First, I am skeptical of Russell's "feeling theory" interpretation of the Human passions, which amounts to a *descriptive* story of various reactions, almost chemical in nature,

---

[55] Russell, *Freedom and Moral Sentiment*, 102
[56] Russell, *Freedom and Moral Sentiment*, 119
[57] Russell notes that "on Hume's account, there are other channels, independent of will and intention, through which an agent can reveal her moral character and qualities of mind," *Freedom and Moral Sentiment*, 119.
[58] Russell, *Freedom and Moral Sentiment*, 119

firing in our minds; reactions that have no necessary connections to the objects they are meant to be reactions *of*. I am willing to grant that Hume is a "sentimentalist" about moral responsibility, but not that he holds a "feeling theory" of the passions. Second, I am skeptical of Russell's claim that Hume considers the sphere of moral responsibility to extend beyond considerations of freedom. Perhaps the sphere extends beyond the "liberty of spontaneity," but I argue that it does not extend past freedom itself. I expand on both of these points in Chapter 3.

3.3 Scientific Naturalism

The classical compatibilist, or philosophical rationalist more generally, looks at the problems of free will and moral responsibility and asks two questions: "are the concepts we are working with coherent?" and "do these concepts apply to human beings?" Contrary to rationalist, the scientific naturalist starts not with the concepts, but with the *facts* of the case.[59] The naturalist looks to our social lives and discovers that we naturally hold one another responsible. It is a *fact*, as Russell puts it, that "men regard other people as objects of praise or blame—that is, they hold them responsible."[60] It is also the case that people automatically draw inferences about one another's character by taking their conduct as evidence. When we see someone steal, we naturally search for the motives that pushed him do so; and from these motives, we decide whether or not the stealing was blameworthy or praiseworthy. It is the conjunction of these two facts—that we naturally hold one another responsible, and that we automatically infer a person's motivations from their actions—that explains our concept of "moral responsibility." For the naturalist, our concept of responsibility is no more in need of *justification* than our concept of friendship is. As Russell points out, "naturalists do not ask 'Does the concept of responsibility have any application to human beings?' Rather, they ask, 'Do the attitudes, sentiments, and practices *which are constitutive of responsibility* have any adequate or proper

[59] Russell, *Freedom and Moral Sentiment*, 59
[60] Russell, *Freedom and Moral Sentiment*, 59

justification?"[61] A naturalist philosopher may seek to reform a *particular* conception of responsibility on the grounds that it does not cohere with what we know about ourselves and the world, but he cannot reform the concept of responsibility itself; not without drastically redefining what it is to be *us*.

So how does the scientific naturalism approach the problem of free will? As we saw in the previous section, to hold someone responsible for Hume is to "regard them as the object of a certain kind of passion—namely, a moral sentiment."[62] It is a fact that we do hold one another responsible. It also appears that, when we hold someone responsible, our moral sentiments are directed at their motivation or character, and not their action itself. We *feel* differently when an agent maliciously runs someone over with their vehicle, as opposed to when she does so *accidently*. We may accuse the latter person of gross negligence, while we will consider the former character to be vicious. As Russell points out, if we did not have the ability to attribute motivations to people, we would not be able to come to a decision about whether the person should be held responsible: "without knowledge of anyone's character, no sentiment of approbation or blame would be aroused in us. Therefore, without inference, no one would be an object of praise or blame—that is to say, no one would be regarded as responsible for her actions."[63] This fact also requires that an agent's motives and her actions be *fixed* together in some way. If there were no regular connection between what an agent *did*, and what she *intended* to do, she could not rightfully be held responsible for her conduct. I draw to conclusions from this. First, Humean necessity (constant conjunction and mental inference) is "psychologically essential to ascriptions of responsibility, because in the absence of the relevant regularities and inferences, the regular mechanism which produces our moral sentiments

[61] Russell, *Freedom and Moral Sentiment*, 174
[62] Russell, *Freedom and Moral Sentiment*, 58
[63] Russell, *Freedom and Moral Sentiment*, 63-64

would simply fail to function."[64] And second, "liberty of spontaneity is (psychologically) essential to responsibility because it is only in these circumstances (i.e., in which we discover constant conjunctions between motives and actions) that it is possible for us to draw the specific kinds of inferences required to generate the moral sentiments."[65] Russell sees a modern proponent of the scientific naturalism in P.F. Strawson.

In his 1962 article "Freedom and Resentment," P.F. Strawson claims that the two interlocutors in the free will debate—whom he labels the optimists and the pessimists—have reached an impasse. The pessimist, a card-carrying incompatibilist, asserts that determinism is true and that, as a result, we are *not* free. It follows from this conclusion, the pessimist argues, that "our concepts of moral obligation and responsibility really have no application, and the practices of praising and blaming, of expressing moral condemnation and approval, are really unjustified."[66] Conversely, the optimist, a classical compatibilist of sorts, agrees with the pessimist that determinism is true, but she claims that our concepts of freedom and moral responsibility "in no way lose their *raison d'être*" because of the *utility* they serve in making moral judgments and doling out punishments for harmful acts.[67]

Strawson argues that we should bracket, at least for the time being, the question of determinism and its impact on free will. Instead, we should focus our attention on what he calls the *reactive attitudes*: those "non-detached attitudes and reactions of people directly involved in transaction with each other; of the attitudes and reactions of offended parties and beneficiaries; of such things as gratitude, resentment, forgiveness, love, and hurt feelings."[68] It is these attitudes, Strawson argues, and not our philosophical concepts of "freedom" and "moral responsibility" that

---

[64] Russell, *Freedom and Moral Sentiment*, 67
[65] Russell, *Freedom and Moral Sentiment*, 67
[66] P.F. Strawson, "Freedom and Resentment," in *Freedom and Resentment and Other Essays* (New York: Routledge, 2008), 1.
[67] Strawson, "Freedom and Resentment," 1
[68] Strawson, "Freedom and Resentment," 5

give form to our actual practices of *taking* one another as free and *holding* one another morally responsible. If my friend steals my bike, I may feel resentment towards him. For Strawson, my resentment is constitutive of my holding him responsible for stealing the bike. There is no need for me to reference my theory of free will to decide whether or not my friend was really responsible; my resentment has already done this. But what is the *justification* of my resentment? According to Strawson, there need not be any *external* justification for my reactive attitude. It is simply a fact about our nature that we attach a great importance to the "attitudes and intention [directed] towards us of those who stand in these relations to us and the great extent to which our personal feelings and reactions depend upon, or involve, our beliefs about these attitudes and intentions."[69] In other words: in virtue of our humanity, we are *committed* to these reactive attitudes insofar as they give form to our social lives.

Of course, someone may claim that we ought *not to be* committed to our reactive attitudes. Perhaps I spent the afternoon after my friend stole my bike reading Spinoza, and I came across the following passage: "in nature there is nothing contingent, but all things have been determined from the necessity of the divine nature to exist and produce an effect in a certain way."[70] I think to myself: "I should not feel resentment towards my friend because, after all, he was determined to steal the bike." In thinking this, I have decided to take the theoretical point of view towards my friend, treating him as an object of analytic study, and not the proper object of reactive attitudes. But in order to take this viewpoint seriously, I realize that it must be *universalized*, for insofar as my friend is exempt from the reactive attitudes on determinist grounds, he shares this feature with *all* people, including myself. Should I take my "theoretical point of view" to its logical conclusion and exempt all people from the reactive attitudes?

---

[69] Strawson, "Freedom and Resentment," 6
[70] Baruch Spinoza, *Ethics*, Ip29.

Strawson rejects this course of action on two grounds. First, he finds it psychologically inconceivable: "we cannot, as we are, seriously envisage ourselves adopting a thoroughgoing objectivity of attitude to others as a result of theoretical conviction of the truth of determinism."[71] Although we may have the capability to take the "theoretical stance" in particular cases, we simply cannot maintain it with regard to the whole of humanity because our interpersonal relationships, which provide ultimate meaning to our lives, are predicated on the employment of the reactive attitudes. Second, Strawson claims that when "we do in fact adopt such an attitude in a particular case, our doing so is not the consequence of a theoretical conviction which might be expressed as 'determinism in this case', but is a consequence of our abandoning, for different reasons in different cases, the ordinary inter-personal attitudes."[72] Take the example of a man named Bob who has damage to his amygdala. Bob is inclined to take abnormally large risks, which leads him to gamble most of his money away at the roulette table. We do not consider Bob to be the proper object of our reactive attitudes, but not because he is *determined*; instead, we judge him thusly because he is *deficient* in some way.

For Russell, the key point of scientific naturalism lies in the fact that "both these thinkers [Hume and Strawson], in different ways, shift emphasis and attention from problems of freedom to problems of responsibility. Instead of arguing that we interpret responsibility in terms of freedom, it is suggested that we try to understand the conditions of freedom in terms of an empirically better informed theory of responsibility."[73] The concept of moral freedom that Hume defends—the liberty of spontaneity—does not provide a precondition for moral responsibility. Instead, this notion of freedom is something that we arrive at through an analysis of the functioning our moral sentiments. It turns out that in order to make sense of moral responsibility, we need to presuppose something

[71] Strawson, "Freedom and Resentment," 14
[72] Strawson, "Freedom and Resentment," 14
[73] Russell, *Freedom and Moral Sentiment*, 81

like a non-metaphysical version of the liberty of spontaneity. This *naturalistic* line of reasoning takes Hume far from the previously considered classical compatibilist interpretation.

Unlike the construal of Hume as a *feeling* naturalist, I tend to agree with Russell's construal of him as a *scientific* naturalist. That said, I argue that the particular kind of scientific naturalism that Russell saddles Hume with is overly reductive, being focused more on the internal bare *form*, or mechanical functioning of the moral sentiments than with their social *content*, or the ways in which they are shaped by our interactions with one another. In this way, I think Hume is even closer to Strawson than Russell even realizes.

3.4 Does Hume Solve the Free Will Problem?

Russell takes Hume's commitment to naturalism (in both its *feeling* and *scientific* senses) to be laudable in a number of respects. He praises Hume for recognizing that our concepts of freedom and moral responsibility are not the products of metaphysical inquiry, but are a *given* of human nature itself. Russell also finds Hume's naturalism to be delightfully impervious to a form of moral skepticism.[74] Nevertheless, he argues that Hume's specific, naturalist account of freedom and moral responsibility is deeply flawed because he is unable to give a *justificatory* account of the moral sentiments.

Hume's account of the passions is central to his concept of moral responsibility. On Russell's reading, the account that Hume provides includes descriptive analysis of *how* our sentiments function, but leaves out a proper *justification* of their functioning. But, for Russell, "the philosopher's concern should not be with the mechanics which generate peculiar, atomistic sensation. It is rather, the cognitive-evaluative features involved in ascribing responsibility which

[74] Russell, *Freedom and Moral Sentiment*, 171-173

should be the philosopher's principal concern."[75] Because Hume fails to account for any cognitive-evaluative content of our passions, he is not able to explain how we can *properly* attribute responsibility to ourselves and others by way of the passions. Russell claims that it follows from this inability to distinguish between *proper* and *improper* functioning of the passions that Hume's account of moral responsibility is not satisfactory.

Russell argues that Hume's appeal to the "calm" passions will not rescue his account of moral responsibility. For, as with the "violent" passions, the calm passion is determined either by reference to the *feeling* that arises in us upon experiencing it, or by the *effect* that it has on us (physiologically or behaviorally speaking). The former is not helpful for discerning the nature of a calm passion because, as Hume notes, "tho' they [the calm passions] be real passions, [they] produce little emotion in the mind, and are more known by their effects than by their immediate feeling or sensation" (T, 2.3.4.9). We are left with the latter option. But, because different effects could result from *like* passions—a passion of disapprobation may lead one agent to censure an offender, while it may lead another agent to assault him—Russell argues,

> We need to know more about the beliefs and evaluations that give rise to the hostile behavior and conduct. Furthermore, it seems clear that the behavior and attitudes appropriate to the moral sentiment must be interpreted as expressions of beliefs and evaluations rather than simply as expressions of sui generis sensations. In this way, in order to infer the presence of (calm) moral sentiments—either in our own case or in the minds of others—we need to be able to identify the beliefs and evaluations which are expressed by the conduct and behavior of the persons who maintain these sentiments.[76]

Our ability to identify the various beliefs and evaluations that are expressed by an agent's conduct presupposes that we have an adequate account of her character. That way we are able to recognize a

[75] Russell, *Freedom and Moral Sentiment*, 89
[76] Russell, *Freedom and Moral Sentiment*, 90

particular performance as somehow *representative* of an enduring quality or character trait she possesses.

Russell argues that this is exactly what Hume fails to provide: an adequate account of moral character. One's "character" is defined in terms of the various "traits" one possesses. These character traits are defined as an enduring set of passions. But, as Russell points out, "an understanding of a person's character in these terms generally involves identifying the ways in which the various passions are related to each other and shape the person's will."[77] It is not enough to be able to identify the various character traits that one has, for to give a full account of one's "character" is to understand *how* these traits fit together and inform one's actions. Despite the fact that we always infer a person's character from his actions, Russell stresses we cannot *only* count on action description to grasp a person's character. An agent may have various character traits that do not intentionally manifest themselves. Consider someone that is a naturally short-tempered person, but is gifted with the quality of "strength of mind." She is able to *control* her temper, but she is never able to rid herself of the trait: that trait would still be considered *part* of her character, even if it was *also* true to say about her that she is strong of mind. Because all character traits create feelings of pleasure and pain in others, her character could still be an object of disapprobation. It follows, Russell argues, that people are ultimately held accountable just on account of how others react to their actions. The result is that, "quite apart from anything else, this approach [Hume's "feeling naturalism"] leaves us entirely unable to say why some people are not appropriate objects of moral sentiment. Hume provides us with no adequate account of the nature of moral capacities required for a person (or creature) to be deemed an appropriate object of moral sentiment."[78]

Hume needs a way to distinguish between *appropriate* and *inappropriate* functioning of our moral sentiments. His account of moral freedom, the liberty of spontaneity, may provide the

[77] Russell, *Freedom and Moral Sentiment*, 96
[78] Russell, *Freedom and Moral Sentiment*, 91

necessary criterion for making such a distinction. But, as we have seen, the liberty of spontaneity suffers from major problems. All that the liberty of spontaneity provides us with is the ability to distinguish between *voluntary* and *involuntary* actions on account of the kind of connection that exists between an agent's action and his character. But what if an agent's character is involuntary? What if it comes to pass that one can have little influence over the nature of one's character? If so, then a theory of moral freedom that fails to take this fact into account is far too impoverished to accept.

Throughout this chapter, I have focused on my interpretative disagreements with Russell. If Russell is right about his interpretation of Hume's account of causal necessity and the passions, then he *is* right to judge that Hume fails to give a proper *justification* of the moral sentiments. In the next chapter, I construct a modified interpretation of Humean freedom and moral responsibility that, I hope, can answer Russell's criticism.

4. SOCIALIZING THE PROBLEM: HUMEAN MORAL FREEDOM

4.1 Socializing or Naturalizing?

There is good reason to reject Russell's reading of Hume as a "feeling naturalist." And insofar as Russell takes the project of scientific naturalism consist *solely* in the attempt to provide a descriptive, mechanical account of our sentiments, the term is not sufficiently broad to cover the phenomena that Hume is interested in with his "science of human nature." In light of this observation, I argue that Hume does not seek to *naturalize* the problems of freedom and moral responsibility, but *socialize* them. That said: Hume *is* undoubtedly a naturalist, as he takes the realm of the social, or "artificial," to be a modification of the natural (T, 3.1.2).

In *Skepticism and Naturalism: Some Varieties*, P.F. Strawson distinguishes between two types of naturalism: *reductive* naturalism and *non*-reductive naturalism.[79] Reductive naturalism, best exemplified for Strawson in the work of W.V.O Quine, approaches human agents and their behaviors "simply as objects and events in nature, natural objects and natural events, to be described, analyzed and causally explained."[80] The reductive naturalist claims that concepts like "freedom" and "moral responsibility" can be fully cashed out by the cognitive and behavioral sciences. The strategy of reductive naturalism is clearly exemplified in one of its predecessors: Thomas Hobbes. Consider Hobbes' account of liberty. He defines "the will" as that desire which, by way of a mechanistic process, gives rise to behavior. So long as this process achieves completion, and is not obstructed by external forces (natural or otherwise), then the particular behavior that results is taken to be voluntary. Hobbes treats moral responsibility in a like manner, reducing it to a consideration of whether or not one's action was voluntary. Many philosophers that are sympathetic to a Hobbesian or Quinean approach to these questions are *skeptical* about the nature of moral responsibility. If all

[79] P.F. Strawson, *Skepticism and Naturalism: Some Varieties* (New York: Columbia University Press, 1987).
[80] Strawson, *Skepticism and Naturalism: Some Varieties,* 42

natural objects can be described in purely causal terms, and we fail to hold mountains, rocks, trees, or dogs responsible; then we have reason to believe that our practices of holding other natural objects responsible (adult human beings) are misguided.

In contrast to reductive naturalism, Strawson defends a form of *non*-reductive naturalism. The central claim of this view is that we are "naturally social beings; given with…a natural commitment to that web or structure of human personal and moral attitudes, feelings, and judgments."[81] For the non-reductive naturalist, our concepts of freedom and moral responsibility cannot be fully articulated by the cognitive and behavioral sciences because our human agency is constituted by our social and moral relations, which are irreducibly complex. According to Strawson, the chief error of the reductive naturalist lies in his attempt to give an account of human action abstracted from our conditions of sociality. Though reductive naturalism succeeds in giving a "detached" or "objective" account of the human agent, it fails to recognize that this is not a perspective that can be inhabited indefinitely: "Our natural disposition to such attitudes and judgments is naturally secured against arguments suggesting that they are in principle unwarranted or unjustified just as our natural disposition to belief in the existence of body is naturally secured against arguments suggesting that it is in principle uncertain."[82] And because the non-reductive naturalist is able to capture a wider variety of relevant phenomena in her account of human agency than the reductive naturalist is able to, Strawson thinks that we can say that the non-reductive naturalist is, in fact, the more complete form of naturalism.

Russell makes much of the comparison between Hume and Strawson, arguing that both figures claim that "we cannot understand the nature and conditions of moral responsibility without reference to the crucial role that moral sentiment plays in this sphere" (feeling naturalism) and that "[we should] try to understand the conditions of freedom in terms of an empirically better informed

[81] Strawson, *Skepticism and Naturalism: Some Varieties,* 41
[82] Strawson, *Skepticism and Naturalism: Some Varieties,* 31

theory of responsibility" (scientific naturalism).[83] Because Russell takes Hume and Strawson to be in agreement on fundamental issues, we can assume that Russell considers Hume, like Strawson, to be a non-reductive naturalist. However, the contents of Russell's naturalistic interpretation put Hume closer to *reductive* naturalism on three major points: first, Russell claims that Hume holds a *regularity* theory of causation (section 2.1); second, Russell interprets Hume's account of the passions as purely *descriptive* and *mechanical* (section 2.2); and third, Russell argues that, for Hume, attributions of moral responsibility are *not* dependent on considerations of moral freedom (section 2.3). The picture of Hume we receive from Russell's naturalistic interpretation is that of a modern-day Hobbesian.

I do *not* think this form of hard-headed reductive naturalism captures the spirit of Hume's philosophy. That said: I *do* think that Russell is right to draw attention to the philosophical parallels between Hume and Strawson. In my estimation, Hume accepts a form of what Strawson calls a *non-*reductive naturalism. For the sake of clarity, though not elegance, I have chosen the alternate term *socialized* to mark this interpretation, lest it be confused with Russell's *naturalized* interpretation. Hume is primarily concerned with us as *moral* beings, and not as mere objects of nature. Hume closes the *Treatise of Human Nature* with a discussion of the anatomist and the painter. He notes that the anatomist can give advice to the painter by providing him with "an exact knowledge of the parts," but that there is something "hideous, or at least minute in the views of thing, he [the anatomist] presents" (T, 3.3.3.6). Hume judges that the anatomist, a Hobbesian of sorts, is necessary to the science of human nature, but that he must ultimately answer to the painter: "the most abstract speculations concerning human nature, however cold and unentertaining, become subservient to *practical morality*; and may render this latter science more correct in its precepts" (T, 3.3.3.6). For Hume, we are *of* nature, but we are not *mere* nature: we possess capacities that are not shared by all others that are equally *of* nature (trees, volcanoes, fish, rocks, etc.). On account of our shared

[83] Russell, *Freedom and Moral Sentiment*, 81

capacities, we form a moral community. It is from one's having these capacities that common sympathies are developed and conventions are forged; and it is within *this* context that questions of freedom and moral responsibility arise.

Russell considers a species of the proposed "socialized" interpretation of Humean moral freedom, but dismisses it on three grounds:[84]

> (1) Hume's claim that "objects of moral sentiment should be understood, simply, as creatures who possess pleasurable or painful mental qualities—is plainly inadequate and unacceptable" to any account of moral freedom.[85]
>
> (2) Although Hume was aware of the importance of language to moral capacity and moral community, what he "has to say in this regard is rather thin and sketchy."[86]
>
> (3) Hume downplays "that we have a capacity to reflect critically on our moral character and can in some measure alter or amend our character on this basis," an essential component of an account of moral freedom.[87]

In the sections that follow, I respond to each of Russell's concerns. By arguing that Hume holds a *normative* theory of causation (section 3.2) and that the moral sentiments, which are socially reflexive in nature, require causal inferences to function properly (section 3.3), I show that Hume does not take just *any* "creatures who possess pleasurable or painful mental qualities" to be the object of the moral judgments. In the last part of the paper (section 3.4), I respond to Russell's second concern by arguing that Hume takes the importance of a common moral language to be *central* to our ability to make moral judgments, and thus to hold one another responsible. For Hume, it is the ability to *use* this common language that ties us to our fellows, making us members of a moral community. At the

---

[84] Russell, *Freedom and Moral Sentiment*, 92-93
[85] Russell, *Freedom and Moral Sentiment*, 93
[86] Russell, *Freedom and Moral Sentiment*, 93
[87] Russell, *Freedom and Moral Sentiment*, 93

same time, I show that, in virtue of being part of a moral community, we are able to engage in relevant sorts of character reflection and correction.

4.2 The Normativity of Causal Reasoning

If Russell is right about Hume having a *regularity* theory of causation, then Hume is committed to an account of causation that provides no criterion for demarcating *correct* causal inferences from *incorrect* ones. And because the functioning of the moral sentiments depends on one's ability to make causal inferences, the regularity theory of causation leaves Hume without the conceptual resources to label certain instances of responsibility attribution as *improper*. If Hume is to have a worthwhile sentimentalist theory of moral responsibility, he needs to be able to account for, as Russell points out, the "cognitive-evaluative" content of the passions. Over the next two sections, I argue that Hume can meet Russell's challenge. In this section, I argue that Hume considers causal inferences to be *normative*. In the following section, I show how construing causal inferences normatively supplies the moral sentiments with the proper cognitive-evaluative content.

What does it mean to have a normative account of causation? A *regularity* theory of causation holds that the judgment "x causes y" ought to be cashed out as "*ceteris paribus*, x is regularly succeeded by y." To use A.J. Ayer's term: on a regularity theory, for x to cause y means that they are "factually correlated" with one another.[88] In contrast, a *normative* theory of causation holds that the judgment "x causes y" is more correctly understood as "there exists a law of nature governing the relation between x and y, such that if x, then [*necessarily*] y." In other words, for x to cause y means that there is, as Hume puts, a "power or necessary connexion" that *necessitates* that if x happens, then y does as well (EHU, 58).

Hume offers two definitions of "cause": treating it as, what he labels, both a *philosophical* relation (relation of concepts) and a *natural* relation (relation of psychological states). As a

[88] AJ Ayer, "Freedom and Necessity," *Philosophical Essays* (London: Macmillan, 1954), 22.

philosophical relation, he defines causation as "an object precedent and contiguous to another, and where all the objects resembling the former are placed in like relations of precedency and contiguity to those objects that resemble the latter" (T, 1.3.14.31). This definition makes it look like Hume is giving a *regularity* theory, for the governing relation between an effect and its cause is *resemblance*, and not necessity. As a natural relation, he defines causation as "an object precedent and contiguous to another, and so united with it, that the idea of the one determined the mind to form the idea of the other, and the impression of the one to form a more lively idea of the other" (T, 1.3.14.31). According to this definition, the governing relation between a cause and effect is psychological *necessity*, and not resemblance. Resemblance and necessity are very different relations for Hume: *resemblance* "admits of many different degrees" (1.3.13.25), while *necessity* fails to admit of chance or probability (1.3.11). If the natural relation of causation is governed by necessity and the philosophical relation by resemblance, then the two definitions do seem to be at odds with one another: for how can a cause be merely *regular*, yet *necessary*?

One should not treat the "philosophical" and "natural" relations of cause as two separate definitions, but as singular *components* of Hume's larger account of causation. After distinguishing the two types of causal relation, Hume states that "all causes are of the same kind" (T, 1.3.14.32), and gives a more general account of the causal relation: "*an object precedent and contiguous to another, and so united with it in imagination, that the idea of the one determines the mind to form the idea of the other, and the impression of the one to form a more livelily idea of the other*" (T, 1.3.15.35). Notice that this definition combines the accounts of natural and philosophical causation. One experiences x and y as constantly conjoined, which, by means of the association of resemblance, causes the mind to consider x and y as standing in a relation of constant conjunction with one another (philosophical relation). On the basis of this constant conjunction, the mind, upon experiencing x, is *determined* to form the idea of y (natural relation). The reason why Hume chooses to distinguish between the philosophical and

natural relations is to explain *why* our minds are susceptible to causal reasoning. He thinks that there must be something about our minds such that we experience things causally at all. The concept of "constant conjunction" is not sufficient to explain this fact about us. One could, in principle, notice a relation of constant conjunction without concluding that the two objects in question are causally related. Hume concludes that it is because the relation of constant conjunction has a particular *psychological* effect on us that we tend to make causal inferences. In other words, it is only because our minds are constituted in the way they are that "after a frequent repetition, I find, that, upon the appearance of one of the object, the mind is *determin'd* by custom to consider its usual attendant, and to consider it in a stronger light upon account of its relation to the first object" (T, 1.3.14.1).

After further reflection on our psychological capacities, Hume notes that we come to two conclusions: that our mind is determined "the same with those of matter (material objects)" (T, 1.3.14.12) and that the nature of the causal relation "depends so much on the [casual] inference" (T, 1.3.14.29). Our minds are determined—in the manner that a ball is determined to fall to the ground when thrown—to *infer* from cause to effect upon experiencing the constant conjunction of two objects. At the same time, it is only *through* the performance of the causal inference that the causal relation, and so the determination of our minds, is established. This line of thought seems circular. If causal relations are dependent on causal inferences, and we are the authors of causal inferences, how can we know *if* our mind is subject to the same determination as matter? Further: how can we know if matter is *actually* determined if it is we, the determiners, that dictate its determinateness? But if we look closer, Hume's reasoning is not circular; it is reflexive. His claim is that our minds are causally structured, and that we come to understand that this is the case through an investigation into the nature of casual inference. For Hume, we can only investigate the structures of the mind only *indirectly*, inferring its structure from experiencing what it *does*. The overlap between Hume and the German Idealists on this point is notable. As Annette Baier points out, with his stress on the

reflexive nature of the natural and philosophical relation of cause, Hume strikes "a balance between the empiricist emphasis on the determination of our minds by nature's constancies, and the idealist emphasis on what only mind can contribute, the modal element, the 'must,' the 'tis necessary.'"[89]

So what is *normative* about Hume's account of causal relation? Hume claims that our minds are naturally constituted so as to expect the future to follow the past. He refers to our tendency to "transfer the past to the future" as a "full and perfect" habit (T, 1.3.12.10). This perfect habit of ours accounts for the impression of mental determination that we *feel* when we engage in causal reasoning. The habit functions as a rule for cognition, informing it of justified causal inferences. Every time I decide to raise my arm, my arm goes up. If I were to ask myself, "If I decide to raise my arm, does my arm go up?" I would immediately answer "Yes, certainly; it always has." It is only when we come across "contrary experiments," or phenomena of which we have no experience, that our perfect habit to transfer the past to the future is "weakened" (T, 1.3.22.12). In these cases, we do not feel an impression of mental determination. So, if it were the case that one day, my arm did not go up when I decided to raise it, I would not feel as confident in my causal claim that "if I decide to raise my arm, then it will go up." In cases of reduced confidence, we feel that our causal claim is merely *probable* since our cognition fails to provide us with a rule for governing relations of probability. Why not? Well, for Hume, probability is not a suitable philosophical relation:

> The vulgar, who take things according to their first appearance, attribute the uncertainty of events to such an uncertainty in the causes…But philosophers observing, that almost in every part of nature there is contain'd a vast variety of springs and principles, which are hid, by reason of their minuteness or remoteness, find that 'tis at least possible the contrariety of events may not proceed from any contingency in the cause, but from the secret operation of contrary causes (T, 1.3.12.5).

So, if my arm failed to go up when I decide to raise it, it must either be the case that (1) some, yet unknown, force impeded my ability to raise my arm in this instance or (2) I was simply wrong about

[89] Baier, *A Progress of Sentiments*, 90

what causes my arm to go up in the first place; perhaps my bodily movement is not tied to any "decision" I make after all. Until I can settle whether (1) or (2) is the correct explanation, I will feel that the connection between "my decision" and "my arm going up" has yet to be established.

In this case, my *habit* of expecting my arm to go up when I make the decision to raise it is confronted with what Annette Baier calls my *meta-habit* to expect the future to follow the past.[90] My meta-habit—which seeks to "form more specific epistemic habits"—confronts my particular habit to expect there to be a necessary connection between my "decision" and my "bodily movement," and dictates the following to it: because in this instance, the future did not follow the past, you can no longer suppose that the connection between the decision to raise your arm and your arm going up is a *necessary* one.[91] My meta-habit acts as a check on my more specific epistemic habit, holding them to the standard that "what was, shall continue to be." Not only do we provide causal structure to our minds by the process of causal inference, but we provide the *rules* by which proper causal inferences can be made by holding our epistemic habits to standards of contiguity.

By reflecting on our ability to engage in *self-correction*, we can see that our causal inferences are rule-governed. Those inferences that do not track necessary connections between objects are demoted to claims of probability, and thus do not have the same "hold" on the mind as those inferences that are treated as authentically *causal*. Through reflection, we come to understand that our ability to engage in causal inference is then not a mere reaction to our environment, but a rational engagement with it.

---

[90] Baier, *A Progress of Sentiments*, 85
[91] Baier, *A Progress of Sentiments*, 85

4.3 The Sentiments and their Reasons

Russell argues both that Hume is a "feeling naturalist" and that he holds a "feeling theory" of the passions. In the previous chapter, I gave some preliminary reasons to reject Russell's interpretation. Hume *does* claim that attributions of moral responsibility are made by our sentiments, but he does *not* hold a feeling theory of the passions, as he does not consider our passions to be *mere* sensations. Therefore, I reject Russell's label of *feeling naturalism*. The moral sentiments require causal inferences—concerning the relation between action and motive, motive and character, etc.—to function properly. Since these causal inferences are normative, our sentiments are not devoid of reason, but *guided* by it. Further, I argue that Hume's account of the passions is not mechanistic because of the central role that sympathy plays in how the indirect passions function. I begin by considering the relation between sympathy and the indirect passions.

Let us consider the passion of "pride." Hume considers pride to be a violent, indirect passion. Pride is a *violent* passion because it is a vivacious, and it is *indirect* because it does not arise immediately from sensations of pleasure and pain. He distinguishes between original and secondary causes of pride. The *original* cause of pride is the "double relation of impressions and ideas": some agent observes a beautiful house (impression) that belongs to her (idea), and she feels a sensation of pleasure on account of the relation she has to the beautiful house *qua* beautiful object (impression) and the beautiful house *qua* beautiful object that belongs to *her* (idea). But the double relation of impressions and ideas is not enough to *sustain* her passion; for this, she needs other people to validate her pride. As Hume notes, "our reputation, our character, our name are consideration of vast weight and importance; and even the other causes of pride; virtue, beauty and riches, have little influence, when not seconded by the opinions and sentiments of others" (T, 2.1.11.1). It is through the secondary cause of the *sympathy* of others that our passions of pride—or any other indirect passion—receive continued existence. If one planted a tremendous garden, one would feel a great

sense of pride upon its completion; nevertheless, if the garden was never viewed by anyone except oneself, one's pride would likely diminish over time on account of neglect. This is, of course, why we speak of the need to *share* experiences, situations, projects, etc., with others.

Sympathy does not overcome the original cause of the passion it is affixed with, but acts *reflexively* with it. At one time, you may have felt proud of the tremendous garden you planted. It has been five years since you planted it; and since then, you have learned much about the art of planting gardens. Because of this, you do not feel the same amount of pride for the garden as you once it did. Still, other people seem insistent on praising you for the beauty of the garden. Their praise fails to reignite the original passion in you because, as Hume notes, "the praises of other never give us much pleasure, unless they concur with our own opinion, and extol us for those qualities, in which we chiefly excel" (2.1.11.13). That is to say: if the *original* cause of the passion is not present, the *secondary* causes will fail to have an effect on you. The only way the sympathy of others could reignite our sense of pride is if it somehow strengthened the relation between us and the object *qua* beautiful object (2.1.11.5). Not only is it required that the sympathy of other people match up with our original passion, but we must ourselves be in sympathy with those who seek to praise us. It is not enough for anyone to praise us, for "we are not only better pleased with the approbation of a wise man than with that of a fool, but receive additional satisfaction from the former, when 'tis obtain'd after a long and intimate acquaintance" (T, 2.1.11.12).[92] Still, the sympathy of our fellows is necessary for the continued existence of our indirect passions; meaning that, in a not so insignificant sense, the functioning of the majority of our passions are dependent on our relations with others: "the principle of sympathy is of so powerful and insinuating a nature, that enters into most of our sentiments and passions" (T, 3.3.2.3).

[92] Though our opinion of those that are in sympathy with our passions effects the pleasure we receive from it, Hume notes that "popular fame may be agreeable even to a man who despises the vulgar; but 'tis because their multitude gives them additional weight and authority" (T, 2.1.11.19).

Pride is not, strictly speaking, a moral sentiment (these are the *calm*, indirect passions). But we can see, by way of Hume's treatment of the necessary role sympathy plays in the functioning of the indirect passions, that Russell's "feeling theory" interpretation is wrongheaded. Russell claims that Hume's account of the passions concerns the "mechanics which generate peculiar, atomistic sensations."[93] Russell's appeal to the "effect" that the passions have on "certain modes of behavior" is unhelpful, it provides an impoverished, atomistic account of the indirect passions.[94] Indirect passions, having two causes, are going to be defined not only by "the internal operations of the mind" (T, 2.1.11.7), but by the relation that the particular passions stand in to the passions of our fellows in sympathy with us. The particular passion is not only going to be defined by the kind of sympathy that it causes in our fellows, and in what circumstances it is apt to do so, but also by the effect that the sympathy of others will have on the original cause of a particular passion.[95] The indirect passions can be said to be *outer*-directed, for it is social context as much as it is their "mechanical" story that gives them content.

One may object by pointing out that the social *is* mechanical for Hume. For, he notes at the close of the *Dissertation on the Passions* that, "it is sufficient for my purpose, if I have made it appear, that, in the production and conduct of the passions, there is a certain regular mechanism, which is susceptible of as accurate a disquisition, as the laws of motion, optics, hydrostatics, or any part of natural philosophy" (DOP, 6.19.19). I do not deny that much of Hume's general philosophical project is to give a *mechanistic* account of the various functions of our mind, but it is important to be careful about how we parse the language in this passage. The laws of natural philosophy are all *causal* laws. And though Hume stresses the primacy of causal relations in his account of association, he

---

[93] Russell, *Freedom and Moral Sentiment*, 89.
[94] Russell, *Freedom and Moral Sentiment*, 90.
[95] In many ways, Hume's attention is *primarily* drawn towards questions about "the effect particular passions have on the sympathy of our fellows" and "the effect the sympathy of our feels has on the original cause of particular passions." For a good discussion of this, see his discussion of malice and envy (T, 2.2.8)—and, more specifically, the effect that judgments of merit and demerit have on each of the passions.

also claims that *resemblance* is central to the functioning of sympathy, and thus the passions: "'tis obvious, that nature has preserv'd a great resemblance among all human creatures, and that we never remark any passion or principle in others, of which, in some degree or other, we may not find a parallel in ourselves" (2.1.11.5). It is also on account of sympathy and resemblance that we experience the wide array of "mixed" passions that Hume considers.[96] Sure, we could give the mixed passions a more mechanistic explanation—Hume ends up appealing to "the whole bent or tendency" of a pain or pleasure to account for their existence (T, 2.2.9.2)—but this would be to miss the point. Even if we could account for all of our passions in terms of various mechanisms in the mind, it is still the case that our passions are not *merely* mechanistic. To make the latter argument, as Russell does, is to judge that our mental states *are* physical states on account of their instantiation in the brain: a claim, to quote Sellars, that is "either very exciting but false, or true but relatively uninteresting."[97]

Now that we have discussed the social nature of the passions, let us consider the integral role that reason plays in the functioning of the moral sentiments. Hume is claimed to have held a "combat" theory of the relation between reason and the passions.[98] Philosophers that have traditionally held this view consider virtue to consist in the victory of reason over the passions. Hume, insofar as we attribute this "combat" view to him, is taken to have turned this more traditional account of virtue on its head: defining virtue instead as any "*mental action or quality that gives to a spectator the pleasing sentiment of approbation*" (EPM, 239). Further considerations of Hume's texts support this reading, as his definitions of "reason" and "the passions" seem to be *conflicting*:

---

[96] See Hume's discussion "Of the mixture of benevolence and anger with compassion and malice" in *Treatise*, 2.2.9.

[97] Wilfrid Sellars, "The Identity Approach to the Mind-Body Problem," *In the Space of Reasons* (Massachusetts: Harvard University Press, 1997).

[98] See Christine Korsgaard, "The Normativity of Instrumental Reason," *The Constitution of Agency: Essays on Practical Reason and Moral Psychology* (New York: Oxford University Press, 2008), 27-68.

> The former [reason] conveys knowledge of truth and falsehood: the latter [the passions] gives the sentiment of beauty and deformity, vice and virtue. The one discovers objects as they really stand in nature, without addition or diminution: the other has a productive faculty, and gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises in a manner a new creation. Reason being cold and disengaged, is no motive to action and directs only the impulse received from appetite or inclination, by showing us the means of attaining happiness or avoiding misery: Taste, as it gives pleasure or pain, and thereby constitutes happiness or misery, becomes a motive to action, and is the first spring or impulse to desire and volition. (EPM, 246)

But while Hume considers the function of our reason and our passions to be *contrary*, he does not take them to be *opposed* to one another. In fact, Hume does not consider it possible that such a conflict between our reason and our passions should ever arise. The role of reason is to "direct our knowledge concerning causes and effects," or give us an understanding of different possible courses of action we can take (T. 2.3.3.2). Meanwhile, "'tis from the prospect of pain or pleasure that the aversion or propensity arises towards any object: And these emotions extend themselves to the causes an effect of that object, as they are pointed out to us by reason and experience" (T. 2.3.3.3). Our reason is seen as *guide* for our passions: feeding us understanding of our surroundings, and telling us what courses of action enable us to best fulfill our desires. That said, it is still our passions that serve as the spring of all our actions, and so it is only *through* our passions that we have the ability to act at all.[99]

When a passion suggests an unwise course of action, Hume points out that the role of reason is to suggest that said course of action undermines the ability of the passion to fulfill itself. He thinks that our passions can be subject to our reason in two ways: "first, when a passion, such as hope or fear, grief or joy, despair or security, is founded on the supposition of the existence of

[99] As Hume notes in the *Enquiry Concerning the Principles of Morals*, "here therefore *reason* instructs us in the several tendencies of actions, and *humanity* makes a distinction in favour of those which are useful and beneficial (EPM, 235).

objects, which really do not exist. Secondly, when in exerting in passion in action, we choose means insufficient for the design'd end, and deceive ourselves in our judgment of causes and effects" (T.2.3.3.6). Say I feel resentment towards my friend for stealing my first edition of Black Flag's "Nervous Breakdown" EP. For obvious reasons, I wish to have this record back. The resentment I have for my friend creates the desire to assault him, which I judge will be the best way to get my record back. How could my resentment, and subsequent action, be "incorrect," on Humean grounds? Well, it could be the case that I have simply misplaced the record; in which case, the resentment I feel towards my friend would be misdirected, and so the action that flowed from it would be equally incorrect. In this case, my resentment was directed at an object I supposed to exist—my friend *qua* record thief—that, upon further investigation, did not. According to this condition, if I had gotten the *facts* right, my resentment would have been properly directed. But even if I did get the facts right, Hume thinks my passion may still be improper. Perhaps the best way to get my record back is not to assault my friend, but to give him back a record of his that I had stolen. In this case, because my end was to get the record back, and because my resentment motivated me to undertake an action that would ensure that my goal was not realized, my resentment was *improper*.

But there are still problems about the relation between the passions and reason: reason may somehow "have a say" in how the passions go about realizing their ends, but this does not mean that the passions themselves are intelligent in any way. However, this is not entirely true. Hume notes that "the moment we perceive the falsehood of any supposition, or the insufficiency of any means our passions yield to our reason without any opposition" (T, 2.3.3.7). So, if there was a passion—such as malice—that dictated a mode of conduct to us that was either unreasonable or misdirected, Hume thinks, quite optimistically, that this passion would cease to influence our conduct. Also, remember that the passions are not mere sensations. Reason discerns the *effect* of certain modes of conduct, and the effects are themselves "presented," so to speak, to our sentiment.

But it is because of our elementary ability to experience pleasure and pain (which makes us susceptible to the influence of forces external to us) and our ability to sympathize with our fellow beings (with whom we are naturally in tune) that we judge the conduct to be either proper or improper. As Annette Baier notes:

> Reason directs, while passion sets the aims. Reason can work out the effects of actions, but the decision as to whether such effects are welcome is made by our capacity to feel pleasure or distress, when this is influenced by our capacity to share other's pleasure and distress…our sense of what is admirable and contemptible and our capacity for sympathy are enough to explain our moral judgments.[100]

Our *already* reflexive passions give us plenty of information about the needs and desires of ourselves and others prior to the influence of reason on them. The fact that our minds are "mirrors to one another" leads to me to naturally consider your well-being (T, 2.2.6.21). Reason will frustrate our conduct only when the mirror has built up a layer of dust from neglect; or worse, when the mirror has cracked.

4.4 Freedom and Moral Community

My claim in this section is simple: Hume takes freedom to consist in an agent's being the *proper* object of the moral sentiments, which requires that she possess certain *capacities*. In light of these capacities, she is *properly* taken to be a member of the moral community. The existence of the moral community is predicated on a common moral language and shared conventions; the agent's membership in the community depends on her ability to take part in them. In what follows, I respond to Russell's last two objections to the *socialized* interpretation of Humean moral freedom—that Hume's conception of moral capacity and language "is rather thin and sketchy," and that he neglects our "capacity to reflect critically on our moral character"—by arguing that Hume's account

[100] Baier, *A Progress of Sentiments*, 47-48

of moral language underwrites the ability to make *proper* responsibility attributions, which itself supplies a *social* basis for character correction.[101]

Attributions of moral responsibility are constituted by the moral sentiments of approbation and disapprobation (which, are themselves taken to be a species of the calm, indirect passions). For one to be considered morally responsible is for one to be taken as an *object* of the moral sentiments. As Hume notes, "an action, or sentiment, or character is virtuous or vicious; why? Because its view causes a pleasure or uneasiness of a particular kind" (T, 3.1.2.3). But it is not *mere* pleasure or uneasiness that invites censure or blame from our fellows, for "'tis easy to observe that it [distinctions of virtue and vice] has also a considerable dependence on the principle of *sympathy* so depended on" (T, 3.3.1.29). Sympathy is required for the functioning of the moral sentiments because *all* indirect passions require sympathy for their continued existence (T, 2.1.11). Even our more violent, indirect passions—like vanity or hatred—will dissipate if they are not reinforced by like minds. Not only are the moral sentiments inherently *social* (as opposed to merely *sensational*), but all instances of moral judgment demand that one reference a "common moral language." As Hume writes:

> 'tis impossible men cou'd ever agree in their sentiments and judgments, unless they chose some common point of view, from which they might survey their object, and which might cause it to appear the same to all of them. Now in judging of character, the only interest or pleasure, which appears the same to every spectator, is that of the person himself, whose character is examin'd; or that of persons, who have connexion to him. And tho' such interests and pleasure touch us more faintly than our own, yet being more constant and universal, they counter-balance the latter even in practice, and are alone admitted in speculation as the standard of virtue and morality. They alone produce that particular feeling or sentiment, on which moral distinctions depend. (T, 3.3.1.30)

This moral language is not explicitly established, but it *emerges* through our natural

[101] Russell, *Freedom and Moral Sentiment*, 93

interactions with one another. For Hume, we "back into" a moral language. All humans share the same biological makeup and have "at least the seeds and first principles" of the moral sentiments (EPM, 180). This natural condition, combined with humankind's predisposition to form groups, leads to the development of communication; first as non-verbal expression and then as a formal language. Initially, an agent uses language to secure his interests. But over time, he learns to use words of "intersubjective significance": "bad," "foul," etc.[102] Through the use of these words, the agent learns to speak from a position that is not, strictly speaking, first-person singular. And for Hume, as soon as we learn to set ourselves beyond ourselves, "we bring our own deportment and conduct frequently in review, and consider how they appear in the eyes of those who approach and regard us" (EPM, 225). But this moment does not mark the beginning of an impartial, abstract rule by which we "correct" the partial sentiments of our fellows. We do not occupy a third-person perspective, external to the moral community, when we use moral language. Instead, this ability to step outside ourselves is conjoined with our natural propensity to feel in sympathy with our fellows.[23] We do not distance ourselves when taking the "general point of view," but we learn to immerse ourselves in the lives of others. The result is not a "view from nowhere," but a view from the "party of human kind" (EPM, 275).

How do the above considerations get us a theory of moral freedom? An agent's ability to make moral judgments requires that she share a common moral language with her fellows, such that she can partake in the "common point of view" with them. If she does *not* share a common moral language with others, she is *not* taken to have the relevant capacities for being considered *within* the realm of the moral sentiments of approbation and disapprobation. This fact is seen in how Hume treats the relation between humans and non-human animals. The latter can only experience violent, direct passions: "they can judge of objects only by their sensible good or evil, which they produce,

[102] I adopt this notion of "intersubjective significance" from David Wiggins. See David Wiggins, *Ethics: Twelve Lectures on the Philosophy of Morality* (Massachusetts: Cambridge University Press, 2006), 45.

and from that must regulate their affections towards them" (T, 2.2.12.3). Because of this, non-human animals are unable to form *moral* sentiments (calm, indirect passions), and thus cannot make *genuine* responsibility attributions. Their inability to make said judgments is reflected in how we treat them. One can feel a sense of delight towards one's pet, but Hume claims that our liking of them "is not caus'd so much by relation, as in our species," but by a more simplistic pleasure we feel in their company (T, 2.2.12.4). In other words: we are unable to feel moral approbation for their conduct. Hume does not consider non-human animals to be *a priori* excluded from our moral community—though he admits that we would be unlikely to sympathize with them, for much of the sympathy we feel for our fellows is based on "a peculiar similarity in our manner, or character, or country, or language" (T, 2.1.11.5). If it were the case that non-human animals *could* form calm, indirect passions, and if we came to see them as more like ourselves, Hume would have no principled reason for stopping them from entering the moral community.

As noted previously, not all members of the moral community are taken to be *equally* responsible. Hume points to three "excusing conditions." The first condition states that an agent that performs an evil act "ignorantly or causally" is *not* subject to blame (T, 2.3.2.7). So, if you accidently take someone's suitcase at the airport because it looks like yours, or you are physically compelled to do so at gunpoint, Hume claims that it would be incorrect to hold you morally blameworthy. The second excusing condition states that an agent that performs an evil act "hastily and unpremeditatedly" (T, 2.3.2.7) is to have *diminished* responsibility. In our legal system, we distinguish between first-degree murder and voluntary manslaughter for particularly this reason. Hume wants to make the claim, and rightfully so, that there is difference between *premeditatedly* murdering your wife's lover and doing so in the "heat of the moment." The last excusing condition states that if an agent repents for her wrongdoing, she is *absolved* of blame (T, 2.3.2.7). So, if you *sincerely* seek atonement for your past life of crime, Hume argues that it would not be right for us to

persist in our considerations of you as a vicious character. In all cases, the agent is excused of some, or all, responsibility because her conduct somehow mismatched the nature of her character: whether it be because the "actions are only momentary, and terminate in them alone," or because the action did not proceed from a "constant cause in the mind" or because the agent's reformation disproves the "just proof" of the necessary connection between the criminal act and her criminal nature (T, 2.3.3.7). Before the agent could be considered a proper candidate for the excusing conditions, it is necessary that she be taken as a *member* of the moral community.

One must be careful to not conflate the following two positions: "an agent is free insofar as she is a member of the moral community" and "an agent is free insofar as she has the relevant capacities to be considered the *proper* object of the moral sentiments." I am arguing that Hume's view is best represented by the latter position, and not the former. Consider the consequences of the former position. If an agent is taken to be free on account of his membership in the moral community, and he is *in fact* a member of said community, then it is impossible for him to be considered *unfree*. Hume's claim that we employ "excusing conditions" in the treatment of our fellow members is rendered conceptually impossible by such a view. The latter position is much more reasonable. Our membership in a moral community is a natural product of our having certain capacities or abilities. It is these capacities, and not the membership *per se*, that renders us free. We come to a better understanding of the nature of the relevant capacities through reflection on the functioning of the moral sentiments, and our regular practices of holding one another responsible.

How does the socialized interpretation of moral freedom answer Russell's concerns about the ability to critically reflect on one's character? In short, Hume is not able to answer Russell's *specific* concern. When Russell levels this criticism against Hume, he has in mind the ability of an agent to "*identify* herself with her actions and feelings insofar as they are what she wants them to

be."[103] Hume never mentions the capacity to identify oneself with one's conduct. However, one should not confuse his silence on this matter with a lack of concern for the nature of moral character. When it comes to questions of morality, Hume is not particularly interested in the first-person perspective; thus his account of "critical reflection" on one's character, norms, society, etc. does not take the shape of an individual deliberating about these matters. Consider the following reflection from Hume: "pride, or an over-weaning conceit of ourselves, must be vicious; since it causes uneasiness in all men, and presents them every moment with a disagreeable comparison" (T, 3.3.2.7). By recognizing the effect that our *undue* sense of pride has on the sentiments of others, we are able to see that our passion is vicious, being contrary to the conduct of a person "of sense and merit" (T, 3.3.2.10). We need not even be subject to *direct* censure from our fellows to reconsider our conduct and character, for we can sympathize with the disapprobation of those we have never met.[104] I may be of a disagreeable nature, and yet sympathize with the Athenian populace's annoyance with Socrates on account of his disagreeable nature. This sense of sympathy I feel with the Athenian populace could be used to temper my own disagreeable nature. From these examples, we can see that Hume is clearly interested in correction of one's character; he simply considers the solution to the problem to be *social*, and not deliberative.

Now that I have explained the *socialized* interpretation of Humean moral freedom, how does it relate to its *classical compatibilist* and *naturalistic* alternatives? The classical compatibilist interpretation correctly points out that Hume takes freedom and moral responsibility to be inextricably linked: an agent can be properly held responsible *only if* his action or character fulfills certain prior conditions (where these conditions are cashed out in terms of capacities, abilities, or conditions of

---

[103] Russell, *Freedom and Moral Sentiment*, 92. For more on this notion of identification, see Harry Frankfurt's essay "Freedom of Will and the Concept of a Person," in *The Importance of What We Care About: Philosophical Essays* (Massachusetts: Cambridge University Press, 1988).

[104] Hume refers to this phenomenon as the *double rebound of sympathy*: "we sympathize with those people in their uneasiness; and as their uneasiness proceeds in part from a sympathy with the person who insults them, we may here observe a double rebound of sympathy" (T, 3.3.3.17).

noninterference). That said: the classical compatibilist fails to grasp both the *sentimentalist* nature of Hume's account of moral responsibility, and the extent to which we gain knowledge of the capacities necessary to our freedom through the articulation of our *practices* of holding one another responsible. The *naturalistic* interpretation accurately recognizes that Hume's account of moral responsibility is based in our passions; and that through an investigation into human nature, we can come to a better understanding of our concepts of freedom and moral responsibility. Still, the naturalistic interpretation falls short by failing to regard attributions of responsibility as dependent on judgments about capacities or abilities that an agent possesses, and by overemphasizing the *physiological* and *psychological* aspects of Hume's account of moral responsibility. As I see it, the proposed socialized interpretation cuts a middle path between the classical compatibilist and naturalistic interpretations: recognizing the sentimentalist nature of Humean moral responsibility, but anchoring it in considerations of one's capacity to participate in a moral community.

5. CONCLUSION

With this project, I set out to provide a modified account of the *naturalistic* interpretation that Paul Russell defends in this work *Freedom and Moral Sentiment*. Russell is right to criticize the overly *rationalist* classical compatibilist interpretation, and to stress the central role the moral sentiments play in Hume's account of moral responsibility. However, Russell fails to consider Annette Baier's observation that "Hume's project all along has been not so much to dethrone reason as to enlarge our conception of it, to make it social and passionate."[105] Because Russell places so much importance on the passions, he tends to underemphasize, or miss, the fact that the moral sentiments depend on judgments about certain capacities or abilities to function properly. It is these capacities, which are *social* in origin, that constitute Hume's account of moral freedom

Though it is a mere sketch of a theory, the *socialized* interpretation of Humean moral freedom I defended in this paper has merit on three fronts. First, Russell admits that a species of the socialized interpretation is conceivable, though implausible. I gave what I took to be an adequate response to each of Russell's arguments, leading me to believe that such a view *is* plausible. Second, Russell is right to claim that Humean moral responsibility is deeply *social*, but his hesitancy to reevaluate Hume's concept of moral freedom leads Russell to attribute an implausible theory of moral responsibility to Hume. The socialized interpretation of Human moral freedom preserves the sociality of the sentiments, while providing ample room for our reason to exert its influence on them. Third, it seems to me that we are lead to attribute the socialized interpretation of moral freedom, or something like it, to Hume on account of the texts themselves: especially considering his account of the passions and causal inference.

[105] Baier, *A Progress of Sentiments*, 278

REFERENCES

Ayer, AJ. "Freedom and Necessity." In *Philosophical Essays*. London: Macmillan Press, 1954.

Baier, Annette C. *A Progress of Sentiments: Reflections of Hume's 'Treatise'*. Massachusetts: Harvard University Press, 1991.

Chappell, Vere, ed. *Hobbes and Bramhall on Liberty and Necessity*. Massachusetts: Cambridge University Press, 1999.

Hobbes, Thomas. *Leviathan*, edited by C.B. Macpherson. New York: Penguin Classics, 1968.

Hume, David. *A Treatise of Human Nature*, edited by David Fate Norton and Mary J. Norton. New York: Oxford University Press, 2000.

_____*A Dissertation on the Passions; The Natural History of Religion*, edited by Tom Beauchamp. New York: Oxford University Press, 2009.

_____*Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, edited by L.A. Sebly-Bigge, and P.H. Nidditch. New York: Oxford University Press, 1975.

Kane, Robert. *A Contemporary Introduction to Free Will*. New York: Oxford University Press, 2005.

Korsgaard, Christine "The Normativity of Instrumental Reason." In *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. New York: Oxford University Press, 2008.

Raphael, D.D., ed. *British Moralists 1650-1800*, Volume I. Indiana: Hackett Publishing, 1991.

Russell, Paul. *Freedom and Moral Sentiment: Hume's Way of Naturalizing Responsibility*. New York: Oxford University Press, 1995.

Sellars, Wilfrid. "The Identity Approach to the Mind-Body Problem," In *In the Space of Reasons*: *Selected Essays of Wilfrid Sellars*. Massachusetts, Harvard University Press, 1997.

Spinoza, Baruch. *Complete Works*. Translated by Samuel Shirley. Indianapolis: Hackett Publishing Company, 2002.

Strawson, P.F. "Freedom and Resentment." In *Freedom and Resentment and Other Essays*. New

York: Routledge Press, 2008.

_____*Skepticism and Naturalism: Some Varieties*. New York: Columbia University Press, 1987.

Wiggins, David. *Ethics: Twelve Lectures on the Philosophy of Morality*. Massachusetts: Cambridge University Press, 2006.