THE UNIVERSITY OF CHICAGO

IN AND OUT OF CHARACTER: MORAL ACTION IN THE EIGHTEENTH CENTURY

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE HUMANITIES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF ENGLISH LANGUAGE AND LITERATURE

BY

SUZANNE TAYLOR

CHICAGO, ILLINOIS

MARCH 2017

Table of Contents

Acknowledgements

This project has been nurtured by many hands, but my greatest debt must be to my committee. My adviser, Jim Chandler, asks the most imaginative and most hair-tugging of questions (always plural) at the most unexpected moments in the most charming way. Without his intellectual guidance and unflagging support over many years, this dissertation would be a lesser thing. Heather Keenleyside has dissected, with a keenness of vision scarcely to be believed, this body of work more times than I can count. Her brains, patience, and sympathy have helped me through my roughest patches. Frances Ferguson's combination of calm and strength has been a constant reminder of what I most admire in our profession. A few words from her have often led to giant upheavals of thought. To all three of my committee members I owe a long arrear of thanks.

I am also grateful to Corrinne Harol and Jayne Lewis for their whip-smart conversation and unfaltering good cheer—to be in their company is to mistake work for pleasure. They have been the shadow members of my committee and each has left an indelible impression on this project. Many faculty at the University of Chicago deserve mention. Very early on, Arnold Davidson taught me to think with abandon; many of my philosophical preoccupations have grown out of conversations with him. Robin Valenza first (properly) introduced me to Hume, and I have been living with him ever since. Richard Strier showed me what great teaching is and Becky Chandler reinforced the lessons. Over the years, Bill Brown, Tim Campbell, and Elaine Hadley all helped me to navigate this landscape in one way or another.

I have many close friends who have doubled as colleagues, and my very affectionate thanks go to Darrel Chia, Trent Leipert, Erin Nerstad, and Carolyn Purnell, all of whom read parts of this project and, even more heroically, listened to my ramblings about it (and other

things). Christine Pallette (W!) did all of that, but to the nth power (and at the eleventh hour). One day I hope to give my above flat-mates the medals they so richly deserve. I have been lucky enough to belong to a sympathetic cohort and I preach solidarity with them: Jim Bassett, Michael Hansen, Jonathan Schroeder, that means you. Daniel Harris, Irene Hsiao, Michael Meeuwis, and Adan Olmedo will always be Hyde Park's Finest. Further afield, Maxime Philippe, Angela Picot, Mary Reid, Elizabeth Steinhauer, and my colleagues at the University of King's College have kept me on my toes. Likewise, the members of the Eighteenth- and Nineteenth-Century Cultures Workshop and the Franke Fellows Workshop at the University of Chicago have been an invaluable source of insight. I found a home at *Critical Inquiry,* and I thank Jay Williams for that. My students at Chicago and King's have been a constant source of inspiration, particularly the legendary no. 12 of 2014/15. I am also grateful to the Whiting Foundation, the English Department and the Humanities Division at the University of Chicago, the University of King's College, Marcia Tillotson, the Jack Miller Center, and the Chawton House Library for generous financial support over the years.

The influence of my friends and family is beyond measure. My Tua Koo, Uncle Tony, and Auntie Cynthia championed my cause from the first. My friendships with Rebecca Donnelly, Claire Hurtig, Juliette Juillerat, Taica Replansky, and Kaitlin Winchester have been a wellspring of happiness for two decades—these must be bonds of kinship by now. Finally, thanks to Koshiro Saito, who has given me many priceless things, Yuki and Bayani among them.

This dissertation is dedicated to my parents. They remain the most Humean people I know: clever, unorthodox, and exceedingly clubbable. I could not have finished—or started—without them.

Introduction

Character, Action, Passion

The second volume of William Godwin's 1794 novel, *Things as They Are; or, The Adventures of Caleb Williams,* begins with a description of Caleb's emotional "embarrassment."[1] Having just heard an account of the sad series of events that has reduced his master, Mr. Falkland, to "the mere shell" of his former self, Caleb reports that he is overwhelmed by the passions he has "witnessed." He writes that his "feelings were successively interested for the different persons that were brought upon the scene"—as his mind turns from one character to next, he experiences sensations of veneration, approbation, astonishment, sorrow, and love—and it seems that only the "shockingly perverted" Mr. Tyrrel fails to find a place in his impressionable young heart (103). Once this initial psychological turmoil subsides, however, Caleb starts to realize that Mr. Collins's account of Tyrrel's death is not as "distinct and satisfactory" as he first thought. One detail in particular troubles him. He writes: "There was something strange in the character of Hawkins. So firm, so sturdily honest and just, as he appeared at first; all at once to become a murderer!" (104). In other words, Hawkins's crime seems to imply a sudden change in his character that does not ring true to Caleb. This incongruity between action and character leads Caleb to wonder, "Was it possible after all that Mr. Falkland should be the murderer?" (104). It is this question—and Caleb's unshakable

[1] William Godwin, *Caleb Williams* (Oxford: Oxford University Press, 2009), 103. Hereafter cited parenthetically in the text and abbreviated CW where disambiguation is necessary.

determination to answer it—that signals the beginning of Caleb's own tragic history, the beginning of the "theatre of calamity" that will become his whole world (3).

That the decisive plot point of *Caleb Williams*—"one of the literary sensations of the 1790s"[2]—is at once a scene of unrestrained sympathy, critical analysis, and moral inquiry is a useful reminder of just how intertwined these issues were at the tail end of the eighteenth century. The age of reason was, after all, the age of sensibility as well.[3] It was an age marked by efforts to develop a science of man that would lay bare the natural operations of head *and* heart, and in doing so, point the way to "intellectual and moral happiness or pleasure" for both the individual and the species as a whole.[4] *Caleb Williams* exemplifies the eighteenth-century preoccupation with understanding and improving the human condition insofar as it focuses on questions about the springs of human action, the possibility of and means for moral change, and the nature of social bonds. Yet Godwin's novel—like the other literary and philosophical works of this dissertation—does not simply rehearse these questions. Rather, it maps them out in a very particular way. Caleb's irresistible sympathy for the characters of Collins's story leads,

---

[2] Gary Handwerk and A.A. Markley, introduction to *Caleb Williams,* by William Godwin (Peterborough: Broadview Press, 2000), 37.

[3] Literary scholars have long recognized the "messiness" of the eighteenth century, and the tendency is to resist traditional periodizations that draw a firm line between the "Age of Reason" (roughly 1700-1740) and the "Age of Sensibility" (the mid to late eighteenth century).

[4] This is Godwin's turn of phrase in the first chapter of *An Enquiry Concerning Political Justice,* ed. Mark Philp (Oxford: Oxford University Press, 2013), 11. Hereafter cited parenthetically in the text and abbreviated PJ where disambiguation is necessary. Newtonian physics, which reduced the chaos of the universe to a few simple principles of motion, captured the cultural imagination of the eighteenth century, inspiring thinkers to imagine the human world in terms of law-like relationships. See for example David Hume's *A Treatise of Human Nature* (1738)*;* David Hartley's *Observations on Man* (1749); and Thomas Reid's *Inquiry into the Human Mind* (1764). For historical background, see Jacqueline Taylor, "The Idea of a Science of Human Nature," in *The Oxford Handbook of British Philosophy in the Eighteenth Century,* ed. James A. Harris (Oxford: Oxford University Press, 2013), 65-84.

quite naturally, to a curiosity about what kind of characters these characters are.[5] Caleb's mission throughout the narrative is to discover which characters performed what actions and, by extension, to determine who is deserving of sympathy and why. In other words, Caleb's attempt to understand his passions and the characters of the persons who inspire them becomes the driving force behind his actions as well as the action of the novel as a whole. Which is all to say that at its core, Godwin's novel is a study of the relationship between character, action, and passion.

This project explores how eighteenth-century literary and philosophical works imagine and represent this constellation of terms. I focus on a small but influential group of British and French writers who framed their ideas about character, action, and passion in terms of a specific question, namely: what does it mean to act out of character? This is the question that underlies Caleb's suspicion that Falkland is guilty, for it is Caleb's inability to imagine how a good person can do a bad thing that convinces him that Hawkins must be innocent. By the end of the novel Caleb will, of course, realize that it is "possible to love a murderer… the worst of murderers" because the best of characters may yet commit the worst of crimes (126). Falkland is no less virtuous than Hawkins, and his "one act of momentary vice" transforms his character only superficially; his "noble nature" remains untainted by his transgression (301). Although Godwin does not explicitly address it here, this truth—that persons can and do act out of character—raises an important ethical problem. If Caleb is right, then it seems that there are situations in

[5] Adela Pinch could easily be talking about this scene from *Caleb Williams* when she notes that Joanna Baillie's "gothic language suggests that the drive to know passions converts its object into a compelling mystery." For many eighteenth-century writers, the desire to know persons and the desire to know passions amount to the same thing. See Adela Pinch, *Strange Fits of Passion: Epistemologies of Emotion, Hume to Austen* (Stanford: Stanford University Press, 1996), 3.

which persons should not be held accountable for their actions. It is this dilemma that motivates the literary and philosophical texts of this dissertation. In one way or another, they are all attempts to formulate theories of character that can explain how a virtuous person might do a vicious thing (or vice versa) without forfeiting notions of moral responsibility. As entities that can potentially shore up or undermine the relationship between character and action, the passions—variously figured as motives for action, connections between persons, and forms of moral evaluation—will play a crucial role in these theories. If the idea of moral accountability hinges on the connection between character and action, the passions matter precisely because they have the power to modulate that connection.

One of the chief premises of this project is that eighteenth-century concepts of character were developed within the context of the debate—ubiquitous in the period—about free will. Within the terms of one version of this debate, persons were either autonomous psychological beings or faceless functions in a law-bound universe. While the proponents of liberty insisted on the importance of desire and intention, their opponents believed that a necessary connection between character and action was the only possible foundation for moral judgment. I argue that authors like Godwin recognized that acting out of character presented a difficulty for both determinists and libertarians. Neither camp could account for why or how a person might suddenly act against type—in a way that is totally out of keeping with her previous conduct—without jeopardizing its ideological position. Deviant behavior was inconsistent with the notion of a unified and temporally continuous psychological being, and downright impossible in a clockwork, law-bound universe. In other words, the concept of acting out of character

was unintelligible within existing epistemological paradigms and, as a result, particularly fertile ground for reconsidering them. Thus, Godwin and others used the idea of acting out of character as a heuristic for rethinking existing models of identity, agency, and accountability. The desire of these writers to make one-off actions morally intelligible led them to configure the relationship between character, action, and passion in new and often surprising ways. For them, in other words, thinking about the idea of acting out of character was a way of better understanding what it means to act *in* character.

In and Out of Character begins (and ends) with Godwin in part because there can be little doubt concerning the philosophical preoccupations of his fiction. A year before writing *Caleb Williams,* Godwin had worked out a theory of motivation in a vast philosophical tract which would come to be regarded as his magnum opus, *An Enquiry Concerning Political Justice, and its Influence on General Virtue and Happiness* (1793). The key chapter of this work, for my purposes here, is the section entitled "Of Free Will and Necessity," to which I will turn below. I want to start, however, by briefly looking at the widely acknowledged source for Godwin's chapter, the similarly titled "Of Liberty and Necessity" from David Hume's *A Treatise of Human Nature* (1739-40). Hume's account of human agency is usually described as compatibilist. For him, necessity is simply the idea that the actions of human beings have causes. Thus, he declares that "no union can be more constant and certain, than that of some actions with some motives and characters."[6] Hume's position, put crudely, is that we have motives for what we do, and these motives are necessarily motivating. So although Hume is no fatalist, he does believe that certain actions *must* follow from certain causes. It is this necessary connection between action

[6] David Hume, *A Treatise of Human Nature*, ed. L.A. Selby-Bigge (Oxford: Clarendon Press, 1978), 404. Hereafter cited parenthetically in the text and abbreviated THN where disambiguation is necessary.

and "motive and character" that makes human behavior uniform, predictable, and punishable (as it evidently is).

Hume's critics take issue with his claim that certain causes must produce certain actions. To think this way is to put man at the mercy of his surroundings. Under this theory, human actions are nothing more than unavoidable reactions to external stimuli, and as a result, they lose their moral valence. Hume takes pains to show that the reverse is true: it is the doctrine of liberty that renders the concept of moral accountability meaningless. After all, he reasons, if a deed is not *necessarily* attached to its doer, then it is only *accidentally* attached to her, and normally speaking, accidents do not invite blame. To deny necessary connection is, in other words, to allow that a person might perform an action without originating it. Hume illustrates the moral implications of such a position by taking it to its logical limit. In his test case, the individual is an agent rather than an author and, for that reason, there is nothing inherent in him that can anchor the action.[7] The action belongs to nobody and no one is responsible for it, no matter how heinous it is: "Actions are by their very nature temporary and perishing; and where they proceed not from some cause in the characters and disposition of the person, who perform'd them, they infix not themselves upon him, and can neither redound to his honour, if good, nor infamy, if evil" (411). The trouble here, according to Hume, is that if motive is invisible and necessary connection does not exist, we have no way of discovering a person's character and, it follows, no way of determining whether his character is the proper cause of his actions. We must begin from the other end. If the proposition "if an action I perform is not caused by something in my character,

[7]The Oxford English Dictionary defines an agent as: "A person acting on behalf of another: A person who acts as a substitute for another; one who undertakes negotiations or transactions on behalf of a superior, employer, or principal." The distinction can be likened to that between a hit man and his employer. "Agent, n.1 and adj.". OED Online. September 2016. Oxford University Press. http://www.oed.com.proxy.uchicago.edu/view/Entry/3859 (accessed October 26, 2016).

then it does not 'infix itself upon' me" holds true, then we can also say that if an action I perform "infixes itself upon" me, then something in my character must have caused it. What Hume's formulation makes clear is that with the contrapositive comes a shift in emphasis from character to action. Connection occurs where character causes action, but it is the self-infixation of action upon character that makes connection visible and thereby enables us to make moral judgments. Here judging a person's guilt is no longer simply a matter of evaluating the behavior of persons, but rather there is a preliminary step which involves observing the behavior of actions: do they infix themselves or not? It is as if action itself attains the status of an agent.[8] Any connection between agent and action will be revealed thus and it is only once this link has been made, says Hume, that an observer will begin to eye the scene with approbation or disapprobation. Actions that do not infix themselves upon a person are not subject to moral judgment.[9] Taken to its logical limit, the doctrine of liberty completely closes the gap between character and action with an unwelcome result. It transforms action into an (literally) impersonal agent without thought or consciousness that cannot, as a result, be held accountable for what it does.

Not even the staunchest libertarian, however, permits action that degree of independence. Hume notes that while some may claim to hold with the doctrine of free will, when it comes time to making sense of the conduct of their fellow men, they inevitably have recourse to necessary connection. Where we distance agent from action, it is because we recognize that the action was committed "ignorantly or casually." The excusing factor is that the cause of the action is

[8] Indeed the verb "infix" was typically used to describe the claws or teeth of a predator. In the *Dialogues Concerning Natural Religion* (1779), for example, Hume refers to "the innumberable race of insects" that "infix their stings" in stronger animals. See David Hume, *Dialogues Concerning Natural Religion: And Other Writings* (Oxford: Oxford University Press, 2008), 96.
[9] In fact, says Hume, here it makes no sense to speak of a "heinous action" as such. Actions are odious only "by their relation to (a hated) person": an action cannot be hated for its own sake (THN, 411).

"momentary, and terminates in [it] alone" (412), the insinuation being that the action was produced under conditions unlikely to repeat themselves. But then Hume goes on to say something strange: "Men are less blam'd for such evil actions, as they perform hastily and unpremeditatedly,… because a hasty temper, tho' a constant cause in the mind, operates only by intervals, and infects not the whole character" (412). My quick temper is less blamable because I do not always act on it, and therefore although it is part of my character, it does not permeate the whole. It is in case of situations like these that rather than speaking of character as cause, Hume talks about causes in characters. In positing that a person's "whole character" consists of "characters" that contain "causes," some of which only sometimes influence action and are less integrated as a result, Hume seems to be arguing that a person may perform actions which are neither out of, nor in, character. That is, such actions may be partly in and partly out of character. For necessitarians, action is always linked to character and, it follows, always subject to moral judgment; but some actions are less blamable than others because they are less closely linked to the "whole character."

The proliferation and displacement of terms in Hume's "Of Liberty and Necessity" suggests an anxiety concerning the relationship between character and action, and indeed, over what "character" itself might be.[10] This anxiety was not Hume's alone.[11] To what extent does

[10] It is worth noting, moreover, that throughout Book II of the *Treatise,* Hume is working with a concept of a person's character as something durable, which does not sit well with his famous formulation of the self as "a bundle of perceptions." See Philippa Foot, "Freewill as Involving Determinism," *Virtues and Vices: And Other Essays in Moral Philosophy* (Oxford: Oxford University Press, 2002), 62-74. For an alternative reading, see Annette C. Baier, *Death and Character: Further Reflections on Hume* (Cambridge: Harvard University Press, 2008).
[11] Indeed, "Of Liberty and Necessity" was a particularly contentious section of the *Treatise.* Spirited rejoinders came thick and fast—from critics who argued that necessary connection undermined our notions of personal identity (Lord Kames) and agency (Joseph Highmore, James Beattie, Lord Monboddo) to those who saw flaws in Hume's logic (John Leland, Thomas Reid).

one's character determine one's actions? Which of his actions, if not all, make up a man's character or "characters"? If character can only be observed indirectly, can we know it with any certainty? Most importantly, what are moral judgments to be based on: character, intention, or action? These questions worried Enlightenment thinkers who sought to ground morality in human nature (reason, sensibility) rather than divine or natural law, a project that necessarily granted "character" a certain pride of place in eighteenth-century philosophical and literary culture. The *Treatise* is typical of its period insofar as its focus on the degree to which a person's character serves as a basis for moral judgment puts the concept of character center stage.

This emphasis on character by eighteenth-century writers presents an interesting contrast to the recent work of literary critics who are challenging Ian Watt's "rise of the individual" account of character and are exploring the limits of "impersonality."[12] Theorists of impersonality

---

See James Fieser, ed., *Early Responses to Hume's Metaphysical and Epistemological Writings I* (Bristol: Thoemmes Press, 2000).

[12] "Impersonality" is the term Sandra Macpherson uses to describe this new school of character studies in *Harm's Way: Tragic Responsibility and the Novel Form* (Baltimore: Johns Hopkins University Press, 2009). Macpherson is hereafter cited parenthetically in the text. At the forefront of this movement, Deirdre Lynch advances a "pragmatics of character," arguing that fictional characters are a coping strategy for readers and that these strategies change as the needs of their readership change, see Deirdre Lynch, *The Economy of Character: Novels, Culture and the Business of Inner Meaning* (Chicago: Chicago University Press, 1998). See also Catherine Gallagher, *Nobody's Story: The Vanishing Acts of Women Writers in the Marketplace,1670-1820* (Berkeley: University of California Press, 1995); Liz Bellamy, *Commerce, Morality and the Eighteenth-Century Novel* (Cambridge: Cambridge University Press, 1998), Margot Finn, *The Character of Credit: Personal Debt in English Culture, 1740-1914* (Cambridge: Cambridge University Press, 2003), and Alex Woloch, *The One Vs. the Many: Minor Characters and the Space of the Protagonist in the Novel* (Princeton: Princeton University Press, 2003). These critics stress the necessity of putting the novel in economic context. Finn's study on credit relations shows that the shift from status to contract was not as seamless as it might first appear. She argues that gift exchange and credit tested the notion of an autonomous, possessive individual. Whereas the unsympathetic "debt law mistook persons for things," gift exchange and credit established a stable social network of mutual obligation (Finn, *The Character of Credit,* 57). Woloch in turn points out that psychological realism went hand-in-hand with social realism, the latter tending toward a representation of characters as types defined in terms of social and economic function rather than interior consciousness. Frances Ferguson offers yet another

eschew the framework of the liberal individual and consider instead what character means when issues of personhood and interiority have been put to one side. In *Harm's Way: Tragic Responsibility and the Novel Form,* for example, Sandra Macpherson offers an ethics that is consistent with impersonality because it is based on action rather than character or, in other words, justice rather than virtue. She argues that the eighteenth-century novel subscribes to the legal doctrine of strict liability, whereby people are held accountable for the unintended consequences of their behaviors. According to Macpherson, eighteenth-century novels ask readers "to think of responsible persons as causes rather than agents," that is, to accept a total indifference to character. We tend to find this disturbing because, as Macpherson points out, abstraction is equated with dehumanization: "flatness [is] ethically and politically unconscionable" (138). But under Macpherson's reading, flatness opens the door to a new ethics where a victim need not be virtuous and a criminal need not be evil. Macpherson suggests that dispensing with the human leaves room for being humane.

Yet Hume's arguments appear to undermine strict liability. He seems to claim that even under necessary connection, a person is less to blame for a bad action if it stems from a cause peripheral to his character.[13] Lord Shaftesbury is of like mind. How many bad actions must one

---

approach, thinking about character with regard to action rather than personality. Utilitarianism, she says, offers an alternative understanding of the individual in that it creates social structures that consider persons not in relation to property or personal character, but rather as causes of actions that benefit or injure the group. See Ferguson, *Pornography, the Theory: What Utilitarianism Did to Action* (Chicago: Chicago University Press, 2004).

[13] Interestingly, when Hume returns to the question of character and action in Book III of the *Treatise,* he seems to align himself even more closely with the position he attributed to the libertarians earlier in Book II. He explains that "actions themselves, not proceeding from any constant principle, have no influence on love or hatred, pride or humility; and consequently are never consider'd in morality" (575). Yet elsewhere Hume insists that necessary connection *is* a constant principle. Part of the confusion lies in the fact that Hume sometimes ventriloquizes the libertarian position in order to undermine it. Critics have not always been attentive to this rhetorical device.

perform before they "infix themselves upon" one's character and render him unfit for society as a consequence? Shaftesbury's answer is as generous as it is vague: "when the mistakes are either in their nature so gross, or so complicated and frequent, that a creature cannot well live in a natural state nor with due affections compatible with human society and civil life, then is the character of virtue forfeited."[14] Far from divorcing ethics from character, Hume and Shaftesbury link them so closely that the primacy of character threatens to spread to the realm of law (Hume says "punishable" rather than "blamable"). Thus whereas Macpherson sees the novel as adopting and adapting to the law's unbending formalism, moral philosophers like Hume and Shaftesbury seek to reestablish a respect for persons by bringing the idea of character—and acting out of character—into the legal domain.

In grounding this study of character in moral philosophy, I am suggesting that literary characterization is in some sense speaking about "real persons in a world"—that is, characters who are precisely *not* merely forms. Yet this is not to endorse a humanist story that, in Macpherson's words, "aligns justice with what looks like an interiority thesis" (16). Indeed, it should be clear by now that Hume is nothing if not a formalist. For him, both the physical world and the moral world follow an inviolable logic. Part of my concern in this dissertation is to argue for the importance of a Humean account of the passions to the history of the novel. Since Watt, we have associated the eighteenth-century novel with psychological interiority and causal logic. As Macpherson's work neatly shows, these two techniques are in tension insofar as they each point to a different understanding of character and action. Interpretations that focus on psychology think of character and action in terms of states of mind whereas those that focus on

[14] Anthony Ashley Cooper Shaftesbury, *Characteristics of Men, Opinions, Times* (Cambridge: Cambridge University Press, 1999), 175.

plot think of them as causes of events in the world. The novels I study describe character and action in terms of passion, but passion seen from the outside, so to speak. For Hume and his intellectual heirs, to talk about passion is not to enter the realm of subjective experience. Passions operate according to certain laws, and it is only by recognizing those laws that we can properly understand our own attitudes towards moral terms like harm, guilt, and punishment.

In short, this project explores how eighteenth-century thinkers in Britain and France describe and account for persons when they act out of character. I argue that for them, the idea of acting out of character is a starting point for mapping out the relationship between character, action, and passion. Ultimately, their aim is to delimit the person in such a way that there is a clear distinction between what she does and what happens to her, or to put it another way, between actions she is morally responsible for and events beyond her control. In trying to draw the boundary between person and world, these writers come to see both feelings and free will as problems. The difficulty of emotions is that although we today tend to imagine feelings as deeply personal—a legacy of Romanticism—the theories of sympathy posited by the "moral sense" philosophers in the eighteenth century portray a world in which emotions travel between persons all too freely. As a result, it is not always evident where passions come from or what they might mean. Thus sympathy produces interpersonal attachments that can make it difficult to distinguish the actions of one individual from those of another. Worse still, this unfettered communication of feeling means that the affective repercussions of actions (call it reward or punishment) always exceed the agent. Insofar as we cannot help feeling the feelings of others, we are bound to answer for their actions. What these eighteenth-century authors learn is that more often than not, our out-of-character actions are in some sense actions by other characters—but we are emotionally attached to them as if they were our own. Free will complicates the relationship

between person and action in a different way. In their efforts to demarcate a stable sphere of agency for the moral individual, the texts I study insist on the necessary connection between character and action. Yet as we shall see, these writers discover that actions are not immune to the vagaries of the outside world. Free will, accident, chance, fits of madness—for the authors of this dissertation, they are all cases of acting out of character—these all seem to weaken the necessary connection between character and action. External forces can sometimes change the complexion of an action so much that it becomes difficult to say to what degree the agent is liable for it (if at all).

Passions have another, larger part to play in these attempts to resolve the problem of acting out of character. Emotions have the power to blur the boundaries between persons, but they can also prop up those boundaries. If sympathy extends the reach of persons and actions by circulating the feelings they inspire, there are nonetheless "pure emotions in the soul" that resist such communication (THN, 367). This sounds very much like the distinction James Chandler draws between "sentiment" (or "distributed feeling") and the "vehement passions," a term he borrows from Philip Fisher.[15] Yet it is not quite right to call these "pure emotions in the soul" vehement passions. These feelings consolidate the self, but not in the way that anger or grief (Fisher's paradigmatic examples) does. That is, they do not outline the limits of the subject's rational will or her sphere of agency.[16] In the works I examine here, I argue, persons emerge as *objects* of passions, not *subjects* of them—although sometimes, as in the case of humility, the object and subject are one and the same. Persons are persons only insofar as they are able to

[15] See James Chandler, *An Archaeology of Sympathy: The Sentimental Mode in Literature and Cinema* (Chicago: University of Chicago Press, 2013), 11-12.
[16] See Philip Fisher, *The Vehement Passions* (Princeton: Princeton University Press, 2009).

arouse pride, humility, love, and hatred. Part of the work of this dissertation will be to explain exactly how moral feelings give rise to persons.

***

This dissertation begins with a consideration of Hume's treatment of character in the *Treatise*. In starting with Hume, I am not suggesting that his philosophical text is a historical source for the literary works of later chapters. The novels I talk about are not merely adaptations of or reactions to Hume's "science of man" (with the possible exception of Godwin's *St. Leon*). Rather they engage with a similar set of moral questions generated by the idea of acting out of character. For example, while Richardson draws on a contagion model of sympathy—a model that we today identify most closely with Hume—to explain why passions pose a problem for moral responsibility, he does not seek to restore the boundaries of the self the way Hume does. Instead, Richardson works through Hume's theory in order to investigate its implications for the classical understanding of justice. Similarly, Laclos's *Les Liaisons dangereuses* takes Hume's account of causation as a starting point for thinking about ethical issues rather than epistemological ones. Far from parasitic, the novels in this dissertation explore many of the questions and concepts that we now associate with Hume, but with different ends in mind.

Chapter One reads Hume's comments on character as an effort to reestablish the limits of the moral person. The question of what exactly Hume means by "character" has generated much critical debate. Readers have been especially troubled by his suggestion that persons might sometimes act out of character, a suggestion that would seem to undo the "necessary connection" between character and action he everywhere insists upon. This chapter seeks to make sense of this seeming inconsistency by drawing on Hume's seldom-discussed analysis of remorse in Book II of the Treatise. There he describes remorse as a "malice against the self," an "irregular

appetite" that neither desire nor character can explain (376). I argue that for Hume, this self-malice is an example of acting out of character—the only such example in the *Treatise*—and that it stems from an excessive and pathological sympathy with others. According to Hume, the remorseful person comes to identify more with someone else than with herself, to the point where she believes that she can reap a malicious pleasure from her own real pain. What this means is that under the influence of this too-perfect sympathy, we are not only strangers to ourselves, but enemies to ourselves as well. Thus although we tend to think of remorse as the correct moral response to wrongdoing, Hume insists that this passion is a dangerous and pathological vice. It is humility, not remorse, that we should be feeling when we have erred. Deaf to the opinions and sentiments of the outside world, humility shows us our vices and moral deserts as they really are. That is, it gives us a view of character from the inside. The passion of humility is then the cornerstone of Hume's ethics, for in telling the truth about character, humility establishes the boundaries of the moral self.

In Chapter Two, I show that for Richardson, as for Hume, passions destabilize the relationship between character and action. Taking its cue from classical tragedy and Hume's principles of association, *Clarissa* portrays familial ties as indissoluble bonds that serve as an unrestricted channel for emotion. The spontaneous circulation of sentiments among family members means that the guilt and suffering associated with a blamable action is felt not only by the agent, but also by everyone related to her. In Clarissa then, it makes no sense to talk about character and action in terms of the individual, and by extension, it makes no sense to talk about acting in or out of character. Agency is a collective enterprise and the moral consequences of actions are always shared.

In both Chapters One and Two I look at the way passions give rise to actions that are, with respect to the individual, neither necessary nor free. Chapter Three shifts the focus from the causal relationship between character and action to action itself. Critics have often remarked upon the theme of determinism that runs throughout Laclos's *Les Liaisons dangereuses*. In this section, I argue that while the novel eschews the idea of acting out of character, it takes seriously the notion that actions, once performed, have a life of their own. Action theorist Donald Davidson coined the term "deviant causal chain" to explain instances where an agent's goal is achieved in an unexpected way. I draw on Davidson's account to describe the novel's representation of causal chains. By showing that action can stray from its course without missing its mark, Laclos suggests that an individual might perform an action that is in character, but in a manner that is out of character. In doing so, he reminds us that how we do something is just as important as what we do.

Chapter Four brings together Laclos's concern with the fate of actions in the outside world with Hume and Richardson's preoccupation with the origins of actions, that is, the passions. As many critics have noted, Godwin's discussion of "Of Free Will and Necessity" in *Political Justice* (1793) is a recasting of Hume's theory of necessity. Yet if Hume looks to character to provide a grounding for moral judgments about persons, Godwin flirts with getting rid of persons altogether. He declares that "there is no such thing as action. Man is in no case strictly speaking the beginner of any event or series of events that takes place in the universe, but only the vehicle through which certain causes operate. … under the system of necessity the ideas of guilt, crime, desert, and accountableness have no place" (PJ, 166; 179). Character, for Godwin, is only "the vehicle" of human activity, for men are passive rather than active, and what they do is dictated by other outside forces. As we shall see, Godwin revises these views in his

later work. Chapter Four shows that *St. Leon* (1799), Godwin's second novel, is a pointed response to *Political Justice* and *Caleb Williams.* It devises a new theory of action that reconciles Hume's claim that action is motivated by passion with *Political Justice'*s insight that morality boils down to universal benevolence. In *St. Leon,* only other people are free—the self is and knows itself to be bound by necessity. What this means is that only others are agents and only they are candidates for moral approbation. In the world of the novel, the self counts for nothing, and in its absence, personal connections are dissolved. This world is an inversion of Hume's. Whereas Hume sees love as the result of proximity, for Godwin, love stems from moral admiration. We love those who perform actions we approve of. The novel's theory of action and the self-oblivion it posits thus enables Godwin to imagine a wholly new way to reconcile passion with impartiality. *St. Leon* brings together the partiality of the affections and the disinterestedness of universal benevolence by suggesting that we might genuinely love those who are morally deserving of our love with all the ardor usually reserved for our nearest and dearest.

I end with a brief Coda that focuses on a particular case of acting out of character in Jane Austen's *Persuasion* (1818). This case makes a final argument for the importance of treating acting out of character as first and foremost a moral problem rather than a formal one. What Austen's novel shows is that when Anne Elliot acts out of character, it is not a triumph of realist characterization, but rather a moment that threatens to rupture the fiction altogether. When a fleshed-out character like Anne harms a one-dimensional character like Mr. Elliot by acting out of character, she invites us to question the asymmetry of the realist novel that Alex Woloch describes as "the one versus the many." Thus while eighteenth-century authors use the idea of

acting out of character to discover what character is and how it should be depicted, Austen uses it to gesture towards the formal and ethical limits of the realist novel.

Chapter One

The View From Inside: Feeling Character in Hume's *Treatise*

"What, however, do we have in mind that a person must *be* to be the object of these moral attitudes?"[1]

For Thomas Nagel, "person" is, at bottom, a moral category. We intuitively distinguish persons from things not because they are rational or capable of language or made in God's image, but because they can be praised or blamed for the things they voluntarily do.[2] In other words, persons are free agents who are responsible for their actions precisely because they are free. Yet as Nagel is quick to point out, this definition does not get us very far since we are immediately confronted with the age-old problem of free will. As Nagel puts it: personhood presupposes agency, but are human beings really in command of their actions? The big difficulty of this question lies in the fact that, at the risk of sounding trite, the head and the heart cannot agree on the answer. For many philosophers, including Nagel, the greatest challenge to determinism in all its various guises is the fact that we *feel* free—but does feeling free necessarily mean that we *are* free? Nagel argues that if we put aside our inner sensations of liberty, we find that nothing we do is fully under our control. From our genetic makeup and the environmental conditions that shape

[1] Thomas Nagel, "Moral Luck" in *Mortal Questions* (Cambridge: Cambridge University Press, 2000), 36-37.
[2] We might map out these philosophical positions using their most famous proponents as reference points: Immanuel Kant (reason), Thomas Hobbes (language), and Moses Maimonides (God's image). See Kant, *Groundwork for the Metaphysics of Morals,* trans. Mary Gregor (Cambridge Cambridge University Press, 2002); Hobbes, *Leviathan* (Oxford: Oxford World's Classics, 2008); Moses Maimonides, *The Guide of the Perplexed,* trans. Shlomo Pines (Chicago: Chicago University Press, 1963).

our personalities to the accidents that generate unintended consequences from our actions, everything we do is tainted by the influence of the outside world. Under the scrutiny of the external gaze, the idea of agency shrinks to nothing, and notions of moral responsibility evaporate along with it. Thus to recognize that we are parts of the world—and parts of the world we undoubtedly are—rather than discrete, autonomous individuals is to forfeit our claim to agency and, by extension, to the moral category we refer to as personhood. In other words, if we think about it too much, what we discover is that there is really no such thing as a person.

Yet as Nagel points out, no amount of thinking can convince us that human beings are merely things. We simply cannot accept that this is a world without people, for everything about our experience says the opposite is true. This is why Nagel thinks there is no solution to the free will problem. Any attempt to answer the question of whether or not human beings possess free will is perpetually frustrated by the irreconcilability of our internal and external points of view.[3] From the outside, we see that human beings have very little control over who they are and what they do. When we turn our eye inward, however, we discover a power to choose our actions as well as certain limits to that power. Nagel explains that:

> [F]rom inside we have a rough idea of the boundary between what is us and what is not, what we do and what happens to us, what is our personality and what is an accidental handicap…. About ourselves we feel pride, shame, guilt, remorse—and agent-regret. We do not regard our actions and our characters merely as fortunate or unfortunate episodes—though they may also be that. (37)

[3] Note that for Nagel, the absence of determinism does not establish free will. Determinism and its opposite both require an external perspective, which means that actions are still events.

Emotions like pride and shame are moral evaluations of the self and they tell us exactly when we are performing actions and when we are caught up in events. After all, we do not feel remorse for other people's actions or for actions we could not have avoided. When we do feel remorse, it is a sure indication that we could (and should) have acted otherwise and that whatever blame there is lies squarely with us. Thus whereas an external point of view suggests that we human beings are passive things, the internal point of view insists that we are persons: free agents who are morally responsible for the things we do. This gap between the view from the inside and the view from the outside means, as far as Nagel is concerned, that it is impossible to say for certain what exactly a person is. If it is an error to grant human beings moral agency, then perhaps the idea of personhood is a fiction. On the other hand, if personhood is a fiction, then what *are* we who can feel pride, shame, guilt, remorse, and agent-regret about our actions?

Nagel's essay is a good starting point for thinking about what it means to be a moral individual because he helps us to see just how inseparable agency, responsibility, and psychological interiority have become in our current moment. These three terms are mutually dependent, and together, they make personhood possible. What Nagel shows, however, is that this model of personhood is fundamentally unstable. This is because in some real sense, human beings are not individuals. We are merely parts of the world, and as such, we are no more capable of independent action than a stone or a knife. Nagel believes we have reached an impasse on this issue because we cannot bring ourselves to believe what we know to be true. But before we admit defeat, we might ask ourselves whether Nagel's starting premise is sound. Are agency, accountability, and emotion as interdependent as they appear?

This chapter will read David Hume's *A Treatise of Human Nature* (1738) in order to think about what personhood looks like if the answer to this question is "no." I will argue that

Hume's famous rejection of free will leads him to a definition of personhood that has very little to do with action, and everything to do with the passions. For Hume, being a person boils down to feeling pride and humility about our virtues and vices, respectively. These two passions are moral responses that happen "from inside," and they reliably demarcate the outlines of who we are precisely insofar as they occur naturally and spontaneously, without any interference or input from either the faculty of understanding or the outside world. In emphasizing the private character of pride and humility, I am taking up a position that goes against the grain of current scholarship.[4] Since Annette Baier's *A Progress of Sentiments* (1991), it has become something of a commonplace in Hume studies to think of the passions as irreducibly social.[5] Part of my concern here will be to show that the critical preoccupation with sympathy and pride has led us to overstate the social aspect of the passions. By turning our attention to humility, we see that for Hume, certain passions draw a line between person and world. If sympathy and action each threaten to undermine personhood and moral responsibility—by affectively binding us to others, in the first instance, and by embedding us in the world of events in the second—pride and humility restore these concepts by consolidating the self from the inside.

I

The lynchpin of Hume's moral theory is his concept of character. Our moral evaluations of other persons are evaluations of their characters, that is, those durable principles of mind that tend to produce patterns of virtuous or vicious behavior. Without character to give them an enduring existence, persons would be agents, but not moral agents. Their actions would not "infix upon

[4] Eugenio Lecaldano is a notable exception to this trend. See Eugenio Lecaldano, "The Passions, Character, and the Self in Hume." *Hume Studies* 28, no. 2 (2002): 175-193.
[5] Baier, *A Progress of Sentiments: Reflections on Hume's Treatise* (Cambridge: Harvard University Press, 1991).

them" and they would be immune to the judgment of others as well as themselves. Despite its importance, however, Hume's theory of character has received surprisingly little attention in the secondary literature. One reason for this neglect is that Hume himself never takes the trouble to develop a precise definition of character. He uses the term inconsistently and sporadically, and any attempt to flesh out his meaning involves a cobbling together of different passages scattered throughout his works. Because Hume treats character as though it requires no explanation, critics have done the same. It is only recently that Hume scholars have begun to explore the idea of character in earnest. These studies have sparked a lively and productive debate about what, if anything, character can tell us about the fictional self of Book I. While some critics present character as an antidote to the skepticism about personal identity that so tortures Hume at the end of Book I, others insist that character and self are two distinct (and for some, incompatible) ideas. There is further disagreement within these camps about whether Hume is a realist about character or self or both (or neither).[6] All these disputes are raising new questions about Hume's

[6] John Bricke makes this point in "Hume's Conception of Character," but scholars have been slow to take up the challenge. See John Bricke, "Hume's Conception of Character." *The Southwestern Journal of Philosophy* 5, no. 1 (1974): 107-13. Currently there are three main critical strands that deal with this issue. The first attempts to reconcile character with the self described in Book I: see Jane L. McIntyre, "Character: A Humean Account." *History of Philosophy Quarterly* 7 (1990): 193-205 and McIntyre, "Hume and the Problem of Personal Identity" in *The Cambridge Companion to Hume,* eds. David Fate Norton and Jacqueline Anne Taylor (Cambridge: Cambridge University Press, 2009), 177-208; A.E. Pitson, *Hume's Philosophy of the Self* (London: Routledge, 2002); Donald Ainslie, "Character Traits and the Humean Approach to Ethics" in *Moral Psychology (Poznan Studies in the Philosophy of the Sciences and Humanities,* vol. 94), ed. Sergio Tenenbaum (Amsterdam: Rodopi, 2007), 79-111; and Erin Frykholm, "The Ontology of Character Traits in Hume," supplement, *Canadian Journal of Philosophy* 42, no. S1 (2012): 82-97. The second treats character as wholly unrelated to self: see for example Bricke, "Hume's Conception of Character"; Susan M. Purviance, "The Moral Self and the Indirect Passions." *Hume Studies* 23. No.2 (1997): 195-212; Lecaldano, "The Passions, Character, and the Self in Hume." The last reads character as a product of moral judgment: see Simon Blackburn, "Hume on the Mezzanine Level." *Hume Studies* 19, no. 2 (1993): 273-88; Timothy M. Costelloe, "Beauty, Morals, and Hume's Conception of Character." *History of Philosophy Quarterly* 21, no.4 (2004): 397-415. Baier also emphasizes the public

work, and scholars are now having to revisit some well-trodden ground—from the account of sympathy in Book II of the *Treatise* to the literary style of the *History of England* (c. 1754-1762) and *My Own Life* (1777)—in an effort to better flesh out Hume's account of character.

Given all this controversy, any consensus is striking. This chapter takes as its starting point a problem that is remarkable for how *little* disagreement it has generated. That is: does Hume think it is possible for a person to act out of character? This question comes with high stakes. An affirmative answer would seem to undermine Hume's doctrine of necessity, posing an even greater threat to moral responsibility than his claim that an action "perform[ed] ignorantly and casually" need not incriminate the person who caused it. Surprisingly then, Hume seems to argue that acting out of character exists as a real, exculpating condition. In an oft-cited passage, he observes that "Men are less blam'd for such evil actions, as they perform hastily and unpremeditatedly, than for such as proceed from thought and deliberation. For what reason? but because a hasty temper, tho' a constant cause in the mind, operates only by intervals, and infects not the whole character" (412). Scholars have tended to downplay the possibility of acting out of character implied by this example. They emphasize instead Hume's specifications that, firstly, we are not absolved but only "less blam'd" when we act hastily, and secondly, a hasty temper "infects not the *whole* character" (my emphasis), meaning presumably that it *does* infect *some* of it. Thus, John Bricke differentiates out-of-character actions from uncharacteristic actions, arguing that although Hume allows for the latter, he rejects the former.[7] According to Bricke's reading of Hume, all intentional actions come from character. Those actions we call "out of

aspect of character, but unlike Costelloe, for example, she does not deny that character is something real that inheres in the person. See Baier, "A Progress of Sentiments."

[7] Bricke explains that "Hume's position is not that a man is responsible only for characteristic actions, but that he is responsible for any action which flows from his character, i.e., any intentional action which he performs" ("Hume's Conception of Character," 112).

character" are the result either of accident—for example, when we are mistaken about how best to achieve a certain end—or of being in a new situation that activates certain traits that were always present, but dormant. Thus, Hume sees action as "a function of the interplay of… mental qualities and of the circumstances in which the individual finds himself," and more often than not, acting out of character occurs when we are placed in an unusual set of circumstances.[8] If those unusual circumstances were to become usual, then the out-of-character action would become usual as well. That is, it would become in character. Bricke maintains that in Hume's view, all intentional actions without exception come from character and, as a result, the only valid exculpating condition is the absence of intention. In the same vein, Annette Baier argues that "It is most unlikely that any of us ever act totally out of character…. When we act out of character in some respect, it will be still in character in another respect."[9] Like Bricke, Baier sees action as a product of different, sometimes conflicting mental principles operating within a specific context. Acting out of character happens when that context suddenly changes.[10] In such moments, a person's whole character may manifest itself in a new way.[11] Such actions are

[8] Bricke, "Hume's Conception of Character," 113.
[9] Baier, *Death and Character,* 20. Baier is primarily interested in the explanatory function of character, and her point is that character is part of "the chain of explanations [that] ends in different places in different cases" (21).
[10] Neither Baier nor Bricke go so far as to claim that circumstance is the *only* cause for acting out of character—equally important are variations in the interaction between passions, habits, and natural abilities. Nonetheless, they both focus on change of situation as the most obvious catalyst for out-of-character action. Thus Baier observes: "As Hume says, it is 'situation and temper' that regularly determine human action…. When he relates the life histories of the monarchs of England and other leaders, it is situation that introduces most of the novelty" (*Death and Character,* 14).
[11] I am being necessarily vague here. For Hume, Whole character includes character traits, but also potentially: natural abilities, passions, habits, custom, and general rules. What counts as a component of character is a major point of discussion among Hume scholars. Jane McIntyre convincingly argues that traits are passions (see "Character: A Humean Account"). Joel Kupperman points out that Hume's concept of character is not reducible to *moral* character, but rather includes natural abilities as well. See Joel Kupperman, "Character and Self-Knowledge."

certainly out of the ordinary, but they are not really out of character. This has become the standard reading of Hume, and Bricke and Baier speak for the majority when they pronounce acting out of character to be a misnomer rather than a suspension of moral necessity. The accepted line on Hume's view of acting out of character is that there is no such thing.

In this chapter, I will argue that for Hume, there is no such thing as acting out of character—with one crucial exception. This exception is the strange "kind of malice against ourselves" extended "even to [our] present fortune" that Hume describes in Book II of the *Treatise* (376). In this passage, seldom discussed in any detail, Hume observes that sometimes a person nurses a self-hatred so strong that it drives her to injure herself, overriding the "wonderful partiality for ourselves" that is natural to human beings.[12] On one level, in identifying self-harm as an instance of acting out of character, I am stating the obvious. It is simply impracticable for a person to exist in a permanent state of self-hatred and to be continuously hurting herself. Were we to meet with such a person, we would likely consider her condition a kind of pathology. Such seems to be Hume's meaning, at least, when he calls the various forms of this malice "irregular appetites for evil" (376). Of course, if Hume is saying nothing more here than "self-harm is contrary to human nature," then the lack of scholarly interest in this passage makes sense.

Proceedings of the Aristotelian Society 85 (1985): 219-238. Robert Heath Mahoney gives a great overview of this debate. See Mahoney, "Hume's Conception of Character." PhD diss., University of Southampton, School of Humanities, 2009: 42-82.

[12] In her recent review of the scholarship on passions in Hume, Elizabeth S. Radcliffe notes that the principle of comparison and the passion of malice have not been sufficiently studied. See Elizabeth S. Radcliffe, "Hume's Psychology of the Passions: The Literature and Future Directions." *Journal of the History of Philosophy* 53, no. 4 (2015): 574-81. James Baillie suggests that "self-malice" is a misnomer on Hume's part. He declares that "I am unclear why he calls these cases of *malice,* since malice involves a desire for harm to come to the person considered; a desire that is absent here. In fact, '*envy* against ourselves' is closer to the mark" (64). Baillie's reading makes sense for instances of malice against the past self, but not for the cases of malice against the present self that I discuss here. See James Baillie, *Routledge Philosophy GuideBook to Hume on Morality* (New York: Routledge, 2000), 64.

However, I think it unlikely that Hume is merely uttering a platitude about our innate instinct of self-preservation (not least because it would be out of character—nowhere in Book II does he express the least interest in self-preservation as a motivating force). I agree with Páll S. Árdal who notes, "Hume does not give us an adequate analysis of these concepts [the two cases of malice against the self], but the emotions to which he is drawing our attention are undoubtedly of interest to anyone who wants to understand 'the moral life.'"[13] The very least we can say is that Hume is flagging a psychological phenomenon that requires more explanation. Hume's account of malice against the self may not be as thorough as we would like, but surely this should pique our interest rather than dampen it.

In fact, Hume is making two substantial claims in this passage. I will briefly outline them before treating each one in detail. The first claim is about the threat that acting out of character poses to the self; the second is about what kind of moral attitude is appropriate when the object of scrutiny is one's own character. With regard to the first point, when Hume describes malice against the self as "irregular"—and importantly, this is the only moment in the *Treatise* where he calls a passion "irregular"—he is anticipating his later discussion "Of Liberty and Necessity" where he defines free will in exactly those terms. There he argues that the regularity and predictability of human behavior boils down to the necessary connection between character and action. Only madmen are (at least somewhat) exempt from moral necessity. As Hume observes, "'Tis commonly allow'd that mad-men have no liberty. But were we to judge by their actions, these have less regularity and constancy than the actions of wise-men, and consequently are farther remov'd from necessity. Our way of thinking in this particular is, therefore, absolutely

[13] Páll S. Árdal, *Passion and Value in Hume's Treatise* (Edinburgh: Edinburgh University Press, 1966), 60.

inconsistent" (404). We are accustomed to seeing the mentally ill as having no control over what they do because their actions are unpredictable and seemingly irrational. Simply put, their erratic behavior does not make sense to us. According to Hume, this is just another way of saying that the madman's actions seem to come from nowhere—we cannot discern a motive or reason behind his conduct—they appear to be completely detached from context or character. Thus, we get things backwards when we treat madness as a lack of liberty. The opposite is true: madmen are the only ones whose actions are not determined by preexisting conditions. Nonetheless, although he does not say it here, we are right to pity them. Hume makes this point some thirty pages earlier when he aligns the madman's freedom from character with malice directed towards the self. Somewhat paradoxically, it turns out that the ability to perform actions uncaused by character is itself a cause for a very specific action: self-harm. The implication here is that far from being a precondition for personhood (as Nagel believes), free will damages and perhaps even destroys the person.

Hume's second big claim in this passage is that in surveying one's own character, the correct moral response is to *do* nothing. He contends that a person who has committed a crime should feel either humility or nothing at all. What she should not feel under any circumstances is remorse. Hume's readers typically conflate the passions of humility and remorse, and as a result, neither has been adequately analyzed. This is partly Hume's fault as he sometimes uses the terms interchangeably. Nevertheless, at certain key moments he makes a point of keeping them separate. At stake here is the difference between the passion produced in response to vicious character (humility) and that which is a reaction to a vicious deed (remorse). As Hume sees it, the trouble with remorse is that we are judging our actions when we should be judging our character. This leads to the mistaken notion that we can compensate for an evil action by

cancelling it out with another action. More specifically, we imagine that we can atone for some pain we have inflicted on someone else by inflicting pain on ourselves. Hume describes this self-punishment as a case of malice against the self and he condemns it as both fruitless and mad. To expose the underlying perversity of remorse, he likens it to a similar impulse that is even more obviously wrongheaded. This second instance of malice against the self is when we cause ourselves pain because we cannot bear to be happy while a friend is in distress. In both these cases, the action stems from a false judgement, spreads pain rather than pleasure, and is of no use to anyone.

Worst of all, however, self-punishment involves a loss of self. We hurt ourselves because we sympathize more with someone else, be it a hypothetical self or a close friend, more than we sympathize with our present selves. Such sympathy is a form of violence against the self, for it inevitably "leaves us with no one to be."[14] As an alternative to this self-hating sympathy, Hume puts forward humility. Humility denies us the illusion of reparation and, in doing so, removes any possible motive for self-harm. Instead, it compels us to sit with the uneasiness that comes with having a constitutional and inexpungible vice. For Hume, to feel humility is to recognize that some of the most fundamental parts of our character are unchangeable and, even more disturbingly, that some of those parts are vices. Yet humility offers consolation too. To sit with our character is to sit apart from our surroundings. It is, in other words, to be a person distinct from both the social world and the natural world.[15]

---

[14] I take this formulation from Nagel, "Mortal Questions," 38.

[15] James Chandler gestures towards the dangers of excessive sympathy (and the antidote: humility) when he distinguishes "sentiment" from the "vehement passions": "[Sentiment] means something like distributed feeling. It is emotion that results from social circulation…. In such an understanding, therefore, sentiment is precisely not a 'vehement passion.' Vehement passion draws a line around a human subject, defining and intensifying personal will, personal limits, a person's strongest emotional attachments. Vehement passion signals our invisible depths.

The remainder of this chapter will examine Hume's account of malice against the self in order to flesh out his theory of character and moral responsibility. I argue that for all his talk of sympathy, Hume remains committed to a model of personhood that draws a firm line between "the destruction of the whole world" and "the scratching of my finger." While critics have tended to stress the social aspect of character and moral judgement in the *Treatise*, Hume's attack on malice against the self is at once an elaboration on and a reaffirmation of the crucial role that the "pure" passions of humility (and pride) play in the formation of the person. At bottom, it is an argument in favor of viewing ourselves from inside, where the passions naturally and reliably (and inexplicably) produce the idea of self. If Hume makes a number of counterintuitive claims here—that remorse is a vicious passion rather than a moral one; that free will undermines personhood instead of supporting it; that excessive sympathy with others leads to self-hatred, not humanity and benevolence—he does so in order to reach a very intuitive conclusion: however blurry it may appear to the outside world, there is a clear line between "what is us and what is not." Contrary to what he says in "Of Liberty and Necessity," Hume believes that it is the internal perspective, not the third-person perspective, that tells us where our responsibility begins and ends.

Sentiments, by contrast, might be said to spread us thin." In Chandler, *An Archaeology of Sympathy*, 11-12. Chandler takes the term "vehement passion" from Philip Fisher, who argues that Hume begins his discussion of the passions in Book II with pride because "a moment of asserted and felt self-worth [is] the primary ground of the passions" (*The Vehement Passions,* 177). Fisher does not address humility although he cites shame as one of the "structural opposites" of pride. He goes on to say that a person feels shame when she has performed an action that challenges her conception of herself. Shame means that she must either change her future actions or "mak[e] permanent the diminished sense of self-worth" that was, initially, "an unexpected and unwelcome event" (197-98). As I will argue, this notion of shame does not line up with Hume's account of humility as a passion that does not produce action.

II

Treatments of Hume's theory of character tend to focus on the relationship between doer and deed. This is hardly surprising, for as critics have correctly pointed out, the necessary connection between character and action is the mainstay of Hume's moral system and, in what amounts to the same thing, the source of his disagreement with the proponents of free will. As Bricke explains, "The crucial mistake of the libertarian, in Hume's view, is that he severs the link between intentional action and character, thus eliminating the kind of 'relation to the person or connexion with him'… which his actions must have if they are to be his responsibility" (112). According to Hume, to say that the will is subject to nothing—that in any given situation, a man might perform any action at all—is to deny that there is anything special about the relationship between an action and the person who performed it. It is to claim that this relationship is merely happenstance and that at any moment the agent is as likely to commit a heinous sin as a virtuous deed. Thus in what can only be described as a questionable interpretation of his opponent's argument, Hume declares that "according to the hypothesis of liberty, therefore, a man is as pure and untainted, after having committed the most horrid crimes, as at the first moment of his birth, nor is his character any way concern'd in his actions; since they are not deriv'd from it, and the wickedness of the one can never be us'd as a proof of the depravity of the other" (411). What the followers of the doctrine of free will fail to understand is that in rejecting the necessary connection between character and action, they are in fact saying that there is no connection at all. As a result, they not only misunderstand the nature of human agency, they also fail to grasp the ethical implications of their own beliefs. Without necessary connection, there is no link between person and action and, it follows, no grounds for moral responsibility.

This is already a big claim, yet Hume goes further. To deny a necessary connection between character and action is not just to deny a (morally significant) connection between them. It is to deny the existence of causes—which is to say, character—all together. Thus, Hume declares that "necessity makes an essential part of causation; and consequently liberty, by removing necessity, removes also causes, and is the very same thing with chance" (407). Character and motives must be necessarily motivating; otherwise, the very concept of motivation as an explanation for human action is meaningless. To put it another way, if persons are free to act or to not act on their motives—if they are free to act in or out of character—then character and motives cannot be causes. Causes, by definition, produce certain effects, which is precisely why they can help us to decipher human behavior. If character and motives only sometimes lead to specific actions, then we can neither infer their presence from action nor predict action based on their presence. Such a theory of motivation would utterly fail as an explanatory tool. In short, Hume claims that free will posits a world ruled by chance rather than necessity, where there is no causal connection between character and action. More to the point, however: in this world, there is no such thing as character, full stop. The problem free will poses for moral responsibility is not just that there is no way to link actions to persons, but rather that persons—as beings who are causes of virtuous and vicious actions and are for that reason the proper objects of our moral judgments—simply do not exist.

It is easy to see why Hume's critics find him so frustrating. Drawing on a series of definitions he develops in Book I, he builds a case for moral necessity that seems to render all counter-argument logically impossible. Thus in "Of the Inference from the Impression to the Idea," he claims that necessity and causation are interchangeable terms and that they both boil down to regularity. Causation is "constant conjunction": "we have no other notion of cause and

effect, but that of certain objects, which have been *always conjoin'd* together, and which in all past instances have been found inseparable" (93). In other words, causation is simply a coincidence of events so uniform, so constant, that we come to regard one to be the necessary consequence of the other. It follows that chance—that is, the absence of necessity—signifies irregularity. Hume explains that "chance is nothing real in itself, and properly speaking, is merely the negation of a cause" (125). The nonexistence of necessity, call it "liberty," is therefore the opposite of causation.[16] As Hume makes clear in Book II, liberty describes a situation in which a person is no more likely to perform one action than any other. Such a person is, as we have seen, no person at all. Thus, Hume builds on the definitions he lays out in Book I to argue that the very idea of character—and the ability to act in or out of it—presupposes necessity. Indeed, the only reason we can argue about the possibility of acting out of character at all is because whether we admit it or not, we are all necessitarians.

Michele Moody-Adams is one of the few readers to register, albeit obliquely, the boldness of Hume's claim that liberty destroys not just the connection between character and action, but character itself. Even though she disagrees with Hume's account of moral responsibility, she shares his sense of what is at stake when it is undermined. They both contend that far from being an act of kindness, failing to hold someone accountable for her guilty actions is to strip her of her personhood. After all, the only way we can argue that someone is not responsible for what she does is to say that she is not fit to be—returning to Nagel's

[16] Importantly, Hume accepts that people can have many different and possibly conflicting motives for doing what they do. His point, however, is that the motive upon which we eventually act is the cause of the action. Thus actions can be a matter of probability, but that does not make them free. Hume writes: "[I]f we affirm that one chance can, after any other manner, be superior to another, we must at the same time affirm, that there is something, which gives it the superiority, and determines the event to that side than the other: That is… we must allow of a cause, and destroy the supposition of chance" (125).

formulation—"the object of moral attitudes." We have to argue, in other words, that such a person is not a person in the moral sense of the word. Moody-Adams cites mental illness as a limit case that proves this point. She explains:

> [M]ental illness is often thought relevant to the question of responsibility, usually when it is believed to cause certain kinds of ignorance or (more controversially) to exert a kind of "irresistible" control over behavior. But mental illness disrupts one's normal patterns of functioning: it manifests itself in what one *does* (and often in what one thinks and feels) rather than what one *is.* To describe someone's *character* as an affliction is to make… a claim that one's affliction manifests itself in one's characteristic patterns of functioning—not just in the kinds of things one does but *in the kind of person one is*…. It is a short step from this to the claim that the only thing that those not similarly afflicted can do with such a person is to try to treat, control, restrain, or perhaps simply tolerate him. To take this sort of attitude toward someone is to see him as no longer fully human.[17]

Intuition tells us that it is not right that a kleptomaniac, for example, should receive the same punishment for stealing as someone who does not suffer from this particular condition. Yet as Moody-Adams shows, the attempt to treat the mentally ill with humanity can quickly turn into its opposite. It is not unreasonable to suppose that measures should be put in place so that a person who has a vice that she cannot change or control is prevented from acting on her vice—particularly where her actions would cause others harm. Thus the logic that would excuse the mentally ill person for actions beyond her control is the same logic that would justify divesting

[17] Michele Moody-Adams, "On the Old Saw that Character is Destiny" in *Identity, Character, and Morality: Essays in Moral Psychology*, eds. Owen J. Flanagan and Amelie Oksenberg Rorty (Boston: Massachusetts Institute of Technology Press, 1997), 124-25.

her of agency all together. For Moody-Adams—and for Hume, I argue—"to exempt someone from responsibility and blame for wrongdoing is simply to deny that person's humanity."[18]

Moody-Adams's discussion of mental illness sheds light on how Hume understands the problem of acting out of character. Indeed, my claim is that in Hume's view, acting out of character is a kind of madness and that this madness goes hand in hand with a loss of self. It is no coincidence, after all, that Moody-Adams's description of mental illness sounds a lot like acting out of character. Both cases raise the same question: what kind of action could undo the relationship between doer and deed without destroying the idea of personhood tout court? Or to put it in Hume's terms, what kind of action could violate moral necessity without subverting causation and character? A passing remark by Moody-Adams suggests an answer to this question. Listing the different presuppositions that prop up Hume's theory of necessity, Moody-Adams complains that Hume "simply assumes that only when someone acts without any motive at all will an action even seem capable of undermining the claim of a necessary connection between character and action" (119). Here Moody-Adams finally puts her finger on what acting out of character would really look like for Hume. The problem is not about whether people are able to perform out-of-character actions (they are). The question is whether people are able to act "without any motive at all."

Moody-Adams believes that Hume does not accept motiveless action as a real possibility. Yet Hume acknowledges that there is at least one kind of person who does act "without any motive at all." In "Of Liberty and Necessity," Hume observes that the actions of mad-men "have less regularity and constancy" than those of wise men. As I explained earlier, Hume defines the looseness of liberty as the negation of causation, or in other words, as the absence of motive. It

[18] Moody-Adams, "On that Old Saw that Character is Destiny," 125.

follows that the actions of mad-men are precisely those motiveless actions that suspend the connection between character and action.[19] What this means, according to Hume, is that if we want to think of such persons as "hav[ing] no liberty," then we must revise our concepts of liberty and necessity, for the one thing the mad-man most definitely has is "the want of determination… and a certain looseness" in his actions. Madness is then the exception that proves the rule. People of sound mind act in regular, legible, and predictable ways. Insofar as it exists at all, free will is the offspring of a diseased mind.

Hume packs this discussion of the mad-man's freedom into a single short paragraph, which perhaps explains why critics have tended to take it at face value, that is, as nothing more than an example of a logical inconsistency in the doctrine of liberty. In an earlier passage, however, Hume suggests that this exception to moral necessity occurs more frequently than we might expect. In "Of the Influence of Belief" in Book I, Hume describes madness in the following terms:

> When the imagination, from any extraordinary ferment of the blood and spirits, acquires such a vivacity as disorders all its powers and faculties, there is no means of distinguishing betwixt truth and falsehood; but every loose fiction or idea, having the same influence as the impressions of the memory, or the conclusions of the judgment, is receiv'd on the same footing, and operates with equal force on the passions. (123)

[19] At first glance, Hume's qualifiers—"less regularity" and "a certain constancy"—might seem to fall short of the complete absence of motive required for acting out of character. However, as Claudia M. Schmidt points out, "[I]n the *Enquiry* Hume compares the uncharacteristic actions of a person who is driven by a 'sudden and unknown frenzy' to perform murder and theft, to an unexpected natural event, such as an earthquake, which has an unknown physical cause" in Schmidt, *David Hume: Reason in History* (University Park: Penn State Press, 2010), 208.

Here Hume again aligns liberty with chance and irregularity. Mental illness is a disorder of the mind wherein imagination imbues fictional ideas with such force that they take on all the force of sense impressions (or ideas based on such impressions).[20] The sufferer can no longer differentiate ideas that have been fabricated wholesale by the mind from those grounded in experience. In this state, the monstrous hydra in my mind is as vivid and real as the rabid wolf that is standing just a few feet away, and being afraid of both, there is no telling which way I will run. If I am most terrified by the hydra and I flee from it in the direction of the wolf, a spectator who sees only the wolf will be hard pressed to make sense of my behavior. The important point here, however, is not simply that the mad-man's overheated imagination can make him act in inexplicable and unpredictable ways, but rather that his imagination is governed by chance. Indeed, it is just conceivable that madness, like virtue in rags, could exist without producing a single action by which we could recognize it. After all, the mad-man is no less mad for running away from the wolf in the direction of the hydra. In both scenarios, the mind's weakest perceptions (fictions) are made equal to the strongest ("the present impressions of the senses") and thereby acquire an influence over the passions far in excess of what they would normally possess. In short, madness describes a situation in which the passions are at the mercy of chance.

What all this means is that a mad-man's actions are motivated in one sense but unmotivated in another. When I flee the hydra, it is because I experience real fear—there is nothing fictitious about my passion. Yet insofar as my fear is based on a chimera, it seems to issue from nowhere. This helps us to see why fits of madness are not, for Hume, reserved for the

[20] Hume's theory of impressions and ideas is complex. The key difference between them, as many commentators have noted, is that impressions are "livelier" than ideas. See Baier for a lucid account of vivacity transfer in Hume. See Baier, "The Life and Mortality of the Mind" in *Death and Character*, 147-82.

mad. We are all of us susceptible to flights of fancy. Indeed, as Hume points out, the task of the poet is to encourage this tendency, that is, to provoke our passions by breathing life into the hydra. He notes that:

> We may observe the same effect [the elevation of fictions to the degree of reality associated with sense impressions] of poetry in a lesser degree; only with this difference, that the least reflection dissipates the illusions of poetry, and places the objects in their proper light. 'Tis however certain, that in the warmth of a poetical enthusiasm, a poet has a counterfeit belief, and even a kind of vision of his objects: And if there be any shadow of argument to support this belief, nothing contributes more to his full conviction than a blaze of poetical figures and images, which have their effect upon the poet himself, as well as upon his readers. (123)

Like madness, art can upset the hierarchical order of our system of perceptions. The only difference between the two kinds of disruptions (to use Moody-Adams's word) is that poetical fancies are easily corrected by means of sober reflection. The mad-man, by contrast, is hopelessly lost in a fog of equally potent perceptions.

By including poets and their listeners among those who suffer from a disordered imagination, Hume is making an important point about character. I have argued above that for Hume, free will and moral responsibility are incompatible. Insofar as the mad-man possesses the liberty of indifference—the capacity to act without motive—his actions do not belong to him and he is not responsible for them. Moody-Adams objects to this conclusion because in her view, as in Hume's, to say that a mad-man is unfit to be the object of praise or blame is to say that he is not a person. This is, for her, a morally reprehensible position to take. Hume seems less bothered. He makes no attempt to rescue the mentally ill from ethical oblivion. Yet if Hume

thinks the mad-man is beyond help, he is more sanguine about the plight of poets and their readers. Such persons experience madness (and may even act upon it), but only occasionally. What is more, there is a cure for their madness. According to Hume, a little contemplation is all that is needed to expose the poet's fictions as fictions and to reinstate the hierarchy of liveliness amongst the mind's perceptions. The case of the poet demonstrates that persons may be intermittently mad without losing their personhood.

From a moral point of view, Hume's account of poetical enthusiasm is perhaps of limited interest. This is because the increased liveliness of artistic fictions is short-lived and they are, as a result, less likely to end in action. Indeed, the example of the hydra and the wolf is mine—Hume does not explicitly argue that the disordered poetical imagination affects actions, although the link is implied in his claim that it influences the passions. Still, it seems that the poet does not *act* out of character in the same way the mad-man typically does. If this is right, then we are back where we started, for these cases appear to support the critical consensus surrounding the question of acting out of character in Hume. In a word, there is no acting out of character for Hume because the mad-man has no character and the poet does not act.

This is not, however, Hume's final word on the matter. I will now turn to Hume's account of "malice against the present self" to show that for Hume, the self-punishment occasioned by remorse and certain kinds of fellow-feeling has its roots in a disordered imagination. This self-malice is, in other words, a form of madness as well as an acting out of character. Persons who harm themselves do so because they have infused a fictitious person with so much vivacity in their imagination that the idea of this person eclipses what is normally the liveliest impression of all: the idea of the present self.

III

Hume first introduces the idea of malice in Part II of Book II of the *Treatise* in order to explain how it is that we can bear ill will towards a stranger whose qualities should, if anything, inspire love. Up until this point, he has been talking about passions directed towards persons who are in some way related to us. We feel pride and humility for ourselves and love and hatred for those who are members of our narrow circle of friends. While there is nothing in Hume's account to suggest that it is impossible to love or hate someone we do not know, practically speaking, these emotions are reserved for persons who are immediately present to us. Thus, Hume explains that I will love an acquaintance more than a stranger, even where the latter is more truly lovable. He writes: "When we have contracted a habitude and intimacy with any person; tho' in frequenting his company we have not been able to discover any very valuable quality, of which he is possess'd; yet we cannot forbear preferring him to strangers, of whose superior merit we are fully convinc'd" (352). Hume makes this point again when he notes that persons who are too far above us cannot be objects of our envy, for "the great disproportion cuts off the relation" between us and them (377). Simply put, it is our family, friends, and acquaintance who are the focus of our passions. Those who pass in and out of our lives never to return are largely a matter of indifference to us. Indeed, it is only thanks to the workings of sympathy that we develop any feelings for such persons at all. Thus, Hume observes that where a man's good qualities benefit those around him, "we approve of his character, and love his person, by a sympathy with the sentiments of those, who have a more particular connexion with him" (602). In other words, we love and hate strangers because sympathy transforms them into acquaintance, and it is only then that they become objects of our good will or ill will.

Yet Hume recognizes that occasionally, we hate strangers who are not hated by those around them. Hume names this seemingly inexplicable ill feeling "malice." He offers this definition: "*malice* [is] a joy in the misery of others, without any friendship or enmity to occasion this concern or joy. We pity even strangers and such as are perfectly indifferent to us: And if our ill-will to another proceed from any harm or injury, it is not, properly speaking, malice, but revenge" (369). Malice is then (to use Hume's word) a "counterfeit" hatred, and this for two reasons. First, because it is "unprovok'd," and second, because it is directed towards someone whose happiness or unhappiness should be irrelevant to us. Neither hatred nor sympathy can elucidate this phenomenon. According to the account of the passions Hume has given thus far, we will love someone if she possesses good qualities and hate her if she possesses bad ones. But malice is blind to the virtues, riches, power, and beauties of other people; its only concern is their happiness or unhappiness relative to our own. Sympathy fares no better as an explanation for malice. For Hume, sympathy is a communication of passions whereby my idea of another person's pain or pleasure is converted into a real impression of pain or pleasure. In other words, by thinking about another person's feelings I come to share those feelings.[21] Yet my experience in the case of malice is exactly the opposite: the other's misery gives me pleasure rather than pain and I can actually produce more pleasure for myself by augmenting her pain. It

[21] Hume's model of sympathy is often described as a contagion model because feelings move from person to person instinctively, as it were. We are bound to share in the feelings of whoever is near to us. For Adam Smith, by contrast, sympathy is not spontaneous. In *The Theory of Moral Sentiments* (1759), he famously states: "As we have no immediate experience of what other men feel, we can form no idea of the manner in which they are affected, but by conceiving what we ourselves should feel in the like situation. Though our brother is upon the rack, as long as we ourselves are at our ease, our senses will never inform us of what he suffers. They never did, and never can, carry us beyond our own person, and it is by the imagination only that we can form any conception of what are his sensations"; in *The Theory of Moral Sentiments,* eds. D.D. Raphael and A.L. Macfie (Indianapolis: Liberty Fund, 1984), 9. Smith's model of sympathy is often described as a projection model.

is in this sense that malice "imitates the effects of hatred," for like hatred, malice is associated with the passion of anger—"a desire of the misery and an aversion to the happiness of the person hated" (367). In short, malice means that I can enjoy and even actively desire the miseries of strangers without discerning anything in them that would justify my ill will.

Hume solves the problem of malice by positing a second psychological concept: the principle of comparison. Comparison is, at bottom, a distortion—both of ideas and of the passions that accompany them. Hume explains:

> [W]e seldom judge of objects from their intrinsic value, but form our notions of them from a comparison with other objects; it follows, that according as we observe a greater or less share of happiness or misery in others, we must make an estimate of our own, and feel a consequent pain or pleasure. The misery of another gives us a more lively idea of our happiness, and his happiness of our misery. The former, therefore, produces delight; and the latter uneasiness. (375)

According to Hume, our satisfaction vis-à-vis our situation depends in large part on how we perceive the situations of others. If the people around me appear to be richer, more powerful, more virtuous, and happier than myself, I will come to think of myself as poor, helpless, base, and miserable—even if I am (intrinsically) none of those things. Comparing my moderate happiness with someone else's great happiness will, in other words, make me see their happiness as even greater than it really is while my moderate happiness will seem painful to me. This new idea of my situation as being painful will in turn cause me real pain. Likewise, the idea of another's misfortune will make me conceive of my good fortune as better than it is in actual fact, and I will be more pleased as a result. This is then the logic behind the passion of malice: "malice is the unprovok'd desire of producing evil to another, in order to reap a pleasure from

the comparison" (377). The principle of comparison provides us with a motive for injuring strangers whose happiness would be otherwise insignificant to us. Thanks to comparison, our ideas about other people—their character, property, family, etc.—shape both our passions and the actions that follow in their wake.

On the face of it then, comparison is the opposite of sympathy, and likewise, malice is the opposite of pity. Indeed, Hume describes malice in exactly these terms when he says that it "is a kind of pity reverst…. In all kinds of comparison an object makes us always receive from another, to which it is compar'd, a sensation contrary to what arises from itself in its direct and immediate survey" (375). As many commentators have noted, it is a mistake to take this notion of malice as "pity reverst" too far. The principles of sympathy and comparison are opposite in their effects—that is, in the passions they produce—but the mechanism of comparison is not simply an inversion of the sympathetic mechanism. Indeed, as Gerald Postema has demonstrated in some detail, the two processes often work in tandem, and many of our "key human passions" exist "in a kind of mixed mode."[22] Nonetheless, it makes sense to ask under what conditions one principle dominates over the other. Why, in other words, do I ever feel malice rather than pity or vice versa? Critics generally concur that malice triumphs over pity when the idea of self is predominant in our thoughts and, by the same token, pity is produced when the idea of the other is center stage. As Postema explains, upon "observing the suffering of another person, one forms an idea of this suffering, but if the self… is already in one's view, the contrast principle simply boosts this sense, enhancing one's existing sentiments."[23] Here Postema is drawing on Hume's

[22] Gerald Postema, "Cemented with Diseased Qualities: Sympathy and Comparison in Hume's Moral Psychology." *Hume Studies* 31, no. 2 (2005): 279.
[23] Postema, "Cemented with Diseased Qualities": 275-76. Baier explains that "the workings of 'comparison'… depend on the blurriness of our vision of another's state, on a defect in our cognitive capacities. Once we see the others' state vividly, sympathy will operate." In Baier's

earlier claim that "the imagination passes easily from obscure to live ideas, but with difficulty from lively to obscure" (339). The idea of the self is the liveliest idea of all, and once we have it in our sights, our attention does not wander. As Hume observes, "when self is the object of a passion, 'tis not natural to quit the consideration of it, till the passion be exhausted" (341). Comparison wins out over sympathy when we are aware of another person's pleasure or pain, but only in relation to our own situation. It is the object of our focus that determines whether we respond to the hardships of others with malice or pity.

What then of Hume's claim that "a person may extend this malice against himself, even to his present fortune, and carry it so far as designedly to seek affliction, and encrease his pain and sorrows" (376)? I have argued above that Hume introduces comparison, the companion principle of sympathy, in order to explain why we sometimes feel a counterfeit hatred—malice—for strangers who have done nothing to deserve our enmity and anger. When comparison is at work, we experience the pain of others as pleasure (and vice versa). It is this inverse relationship that gives us an interest in harming our fellows. I then explained, following Postema's interpretation of Hume, that whether we respond to a stranger's emotions with malice or pity is a matter of where our focus lies. When we are already thinking of self, we see other people's situations only as a foil for our own. On the other hand, when the idea of self is not in the foreground of our minds, we sympathize with those around us, taking pleasure in their pleasure and pain in their pain. It should be clear that the concept of self-malice flies in the face of everything that has been said up until this point. Hume himself declares at the beginning of Part II that "[o]ur love and hatred are always directed to some sensible being external to us; and when

---

reading of Hume, sympathy is our default response to the emotions of others, and it is only when we cannot properly perceive those emotions that the principle of comparison becomes active. See Baier, *A Progress of Sentiments*, 150.

we talk of *self-love,* 'tis not in a proper sense" (329). The primary distinction between love and pride, on the one hand, and hatred and humility, on the other, is the object of the passion—self in the latter cases and, in the former, some other person. Nor is this merely a matter of taxonomy, for once malice is introduced, we see that self-hatred is a logical impossibility. If the joys of malice lie in the pains of others, the idea that we could somehow we take pleasure in our own pains—even to the point of harming ourselves—makes no sense at all. Whatever pleasure I might glean from perceiving and imagining my pain will be cancelled out by my firsthand experience of that pain. In short, an inverse relationship between pleasure and pain is impossible when the subject and object of the passion are one and the same. As a result, there is no obvious means by which self-malice can produce pleasure. If malice was supposed to account for the ill will that sometimes crops up among strangers, the question self-malice raises is this: how can we become so estranged from ourselves that we can become objects of hatred to ourselves?

Hume claims that there are exactly two situations in which malice against the present self occurs, and an examination of these cases will help us to understand the logic of this passion.[24] In the first of these, I hurt myself so as to reduce the contrast between my happiness and the sufferings of someone close to me. Hume explains:

> A person, who indulges himself in any pleasure, while his friend lies under affliction, feels the reflected uneasiness from his friend more sensibly by a comparison with the original pleasure, which he himself enjoys. This contrast, indeed, ought also to inliven the pleasure. But as grief is here suppos'd to be the predominant passion, every addition falls

[24] In this paper, self-malice refers specifically to malice against the *present* self. Hume also talks about a self-malice whereby we take joys in our past pains and pain in our past pleasures, but the distinction between past self and present self means that this kind of self-malice does not pose a problem for comparison. My concern here is with cases wherein the subject of the passion is identical to the object of the passion.

> to that side, and is swallow'd up in it, without operating in the least upon the contrary affection. (376)

Here it is the sorrow of my friend that is first in my thoughts, and my sympathy with him is so strong that my pleasure becomes a second source of pain not only for him, but for me as well. This phenomenon needs further explanation. As Postema notes, Hume's account "still leaves unexplained why one would seek afflictions for oneself in such cases, rather than attempt to relieve those of his friend" (280). Hume has already told us that love goes hand in hand with "a desire of the happiness of the person belov'd, and an aversion to his misery" (367). How is it then that my love for my friend manifests itself not as benevolence towards him, but as anger towards myself? Postema's answer to this question is that "in this case, sympathy has only a limited effect on the soul of the self-abuser: the self still looms large in the foreground of one's emotional experience and it spurs a further 'reversal,' now turning the 'hate' inward" (280). Postema explains self-malice by positing a middle focal point where the idea of self is sufficiently weak to prompt a sympathetic response, but still strong enough to ensure that the only solution to my friend's distress that I can imagine is framed in terms of my own pleasure and pain.

While this interpretation makes sense, I do not think this is what Hume has in mind in this passage. The trouble with Postema's account is that it cannot explain why Hume considers this self-harming sympathy a form of malice. The "reversal" he describes is not an inversion of emotion—the direct survey of my friend's pain causes me pain, not pleasure—but merely an inversion of the desire that we would expect from sympathy. In other words, Postema's "self-abuse" does not engage the mechanism of comparison that, for Hume, is the calling card of malice. Part of Postema's trouble is that he fails to understand the stakes involved in Hume's

claim that self-malice distorts our ideas and passions. As a result, he cannot imagine how someone else's passions could become more powerful than the passions I experience firsthand. Yet this is, I argue, exactly what happens in malice against the self. What Hume is describing in his example of the friend in distress is a three-step process marked not by a sympathy with "limited effect," but by a sympathy so complete that my idea of the other person's pain becomes more real to me than my idea of my own pleasure. The scenario Hume outlines is as follows: by sympathizing with my friend's unhappiness, that unhappiness becomes my own (I "feel[s] [her] reflected uneasiness); I compare this newly acquired unhappiness with my own happiness and my *unhappiness* increases ("the reflected uneasiness [is felt] more sensibly by a comparison with the original pleasure"); I harm myself in order to diminish my original pleasure and thereby halt the increase in pain caused by the contrast between my sympathetically received unhappiness and my own happiness. By breaking down Hume's description of self-malice in this way, we see that the truly odd thing about this passion is not the impulse to self-harm, but rather the dominance of the idea and feeling of the other's pain over the idea and feeling of the self's pleasure. The odd thing is, in other words, that we can feel someone else's emotions so keenly that we come to identify more with her situation than with our own.

This triumph of the other's passions over those of the self is counterintuitive, to say the least. If, as Postema claims, comparison operates when the focus of one's attention is self, then it should not be possible to compare my happiness with my friend's *from my friend's point of view*.[25] Comparison is meant to describe precisely those cases in which the self is unable to perceive the world in any other way but in relation to itself—that is, those cases where my

[25] However, Hume does believe that other people's pride can be, thanks to comparison, a source of humiliation for us. See *Treatise,* 597.

fixation on self prevents me from sympathizing with others. In malice against the self, however, my own situation and feelings are clearly in view, and yet I still identify more with my secondhand pain than my firsthand pleasure, even to the point of actively seeking to decrease the latter in order to decrease the former as well. Hume provides an explanation for this state of affairs, I think, but only obliquely. Much hinges on his description of the two cases of malice against the present self as "irregular appetites for evil" (376). While it is not immediately obvious what Hume means by "irregular" here, earlier in this paper I suggested that "irregular" is, more often than not, a technical term for Hume, and it specifically refers to that suspension of necessary connection he equates with liberty and madness.[26] Mad men are free, according to Hume, because their actions are irregular: unpredictable and seemingly motiveless. Indeed, they *are* motiveless insofar as it is purely a matter of chance whether any particular passion will result in action. As I argued above, for Hume, madness denotes a disordered imagination in which every idea is as real and as vivid as every other idea. This disorder among the mind's ideas is the foundation for malice against the self. I hate myself because I sympathize with my friend's distress so entirely that my idea of her situation becomes as lively as, nay, livelier than my own present experience. This is then the dark side of sympathy. At its limits, sympathy enables the passions of others to overwhelm me to such an extent that I become a stranger to myself—a stranger who will come to be, through the process of comparison, an object of my hatred and ill will.

[26] Incidentally, Hume is not alone in grouping these terms together. Consider Jean-Jacques Rousseau's claim in "The Profession of Faith of the Savoyard Vicar" that remorse is the voice of conscience and it is remorse that makes us free. See *Emile, or On Education,* trans. Allan Bloom (New York: Basic Books, 1979), 280.

Given that self-malice stands for irregularity, liberty, and madness, we might expect it to be a rarity in Hume's world, governed as it is by the laws of necessity. Hume's fear, however, is that this passion is all too common. I have discussed the first case of self-malice—when the sight of a friend's distress leads us to harm ourselves—at some length. The second case follows a similar logic to the first. Hume describes it thus: "When a criminal reflects on the punishment he deserves, the idea of it is magnify'd by a comparison with his present ease and satisfaction; which forces him, in a manner, to seek uneasiness, in order to avoid so disagreeable a contrast" (376-77).[27] In other words, this self-malice refers to those "remorses" that lead to "penances, which men inflict on themselves for their past sins and failings" (376). Because we are accustomed to thinking of remorse as a moral response, Hume's claim that it is an "irregular appetite for evil" might strike us as strange. However, the analogy between the two instances of self-malice can help us to make sense of this statement. In Hume's eyes, it is a mistake to read remorse as merely a recognition of one's past faults and a desire to atone for them. Instead the passion of remorse is, as in the first case of self-malice, the product of a complex interaction between imagination, the principle of sympathy, and the principle of comparison. For Hume, "when a criminal reflects on the punishment he deserves," he is not just imagining his offense and its corresponding punishment; he is imagining the punishment as experienced by a hypothetical self. As Hume's first example of self-malice makes clear, comparison always requires a second person, a second point of view. The only way we can become objects of hatred to ourselves is by sympathizing with someone else. The difference between the two cases of self-malice is just that the friend in the first has, in the second, been replaced by a fictional self who is suffering the just retributions that the criminal has somehow avoided. Thus by way of analogy,

[27] Postema's article does not discuss this second case of malice against the present self.

we can see that remorse unfolds in three steps: thinking of my past wrongs prompts me first to imagine and then to sympathize with a fictional self who is being properly punished for those wrongs; I compare this sympathetically acquired pain with my own present pleasure, and my *pain* increases; I harm myself in order to diminish my present pleasure and, by this means, decrease the pain caused by the contrast between my hypothetical self's pain and my actual pleasure. Like the person who loses herself in the feelings of her friend, the remorseful criminal sympathizes with another self so fully that he comes to treat his real and present self as a stranger whose pain can give him pleasure. For Hume, remorse makes us not just strangers to ourselves, but the enemies of ourselves.

Having sketched out the underlying mechanism of self-malice, we are now in a position to see what makes this passion "irregular" and "evil" in Hume's eyes. In the normal order of things, it is impossible to hate oneself on someone else's behalf (as I do when I harm myself to soothe the distresses of my friend, for example). This is because self is the most vivid idea we are capable of conceiving and no other perception can equal it—certainly not those vicarious ideas animated by sympathy. It is worth pausing on this point, for recent scholars have tended to emphasize the infectiousness of the emotions at the expense of the primacy of the idea of self.[28] As Baier notes, however, sympathy "is fueled by one's maximally vivacious impression of self.… Of course, we influence each other's thoughts and passions, and often replicate them, but that does not affect the self-contained nature of the vivacity transfers in each person's mental history."[29] Here Baier reminds us that however much we feel the passions of others, according to Hume, their passions will never equal the liveliness of our own. Our sympathetically received

[28] See Adela Pinch, *Strange Fits of Passion,* for example.
[29] Baier, *Death and Character,* 163-64.

passions are parasitic on the real passions that are part and parcel of self, and it is only because the self—the assemblage of our impressions, ideas, and passions, or as Hume famously puts it, "a bundle or collection of different perceptions" (252)—transmits its vivacity to our ideas of others' emotions that they can become emotions for us at all. Thus although Hume argues for a transpersonal model of feeling, he also insists on the integrity and boundaries of the self (as Baier says, vivacity transfers are "self-contained"). Indeed, it is only because we can discern between our own emotions and those we acquire from others that we are capable of benevolence. After all, if we could not recognize our experience of another person's suffering as *her* suffering, then sympathy could never become a desire to relieve her pain. All our pity would be for ourselves.[30]

If then the self and its passions are always the foremost ideas in our imagination, malice against the present self should be impossible. I have argued, however, that self-malice is the product of a disorder of the imagination wherein all our ideas—from immediate perceptions of hot and cold to thoughts of hydras and centaurs—are put on a plane. When the imagination is in this disordered state, every passion is on a par with every other passion, which is also to say that no motive is any stronger or any weaker than any other motive. Our actions are, in consequence, irregular. That is, they are unpredictable and arbitrary, for whatever emerges from this whirlwind of passions does so only by pure chance. It is precisely this state of mind that gives rise to self-malice. In order to bear malice towards myself, I must first sympathize with the passions of another so thoroughly that I come to see myself as an object whose pain can give me pleasure

[30] There is much critical debate about how sympathy becomes concern for the other in the *Treatise*. See Árdal, *Passion and Value;* Philip Mercer, *Sympathy and Ethics* (Oxford: Clarendon Press, 1972); Pitson, *Hume's Philosophy of the Self;* Postema, "Cemented with Diseased Qualities."

and whose pleasure gives me pain. While it is not clear from Hume's account whether this complete sympathy with the other *necessarily* leads to self-malice, I am inclined to interpret his use of the term "irregular" as a sign that there is nothing necessary about this process and that comparison may or may not occur in my mad and disordered state. The surprising effect of perfect sympathy is then that it is as likely to lead to self-hatred and self-harm as it is to love and concern for the other.

That self-malice is an evil may seem obvious. Hume's example of remorse should, however, give us pause. After all, for many of us, a certain amount of self-recrimination in the aftermath of some misdeed is not only appropriate, but required.[31] Why is it then that Hume believes that "'tis by none esteem'd a virtue to feel any fruitless remorses upon the thoughts of past villainy and baseness" (298)? In order to make sense of this statement, we first need to understand Hume's particular conception of evil. Hume defines evil as any quality that inspires humility (when the evil quality belongs to the self) or hatred (when the evil quality belongs to someone else). As he explains, "Moral good and evil are certainly distinguish'd by our *sentiments,* not by *reason:* But these sentiments may arise either from the mere species or appearance of characters and passions, or from reflexions on their tendency to the happiness of mankind, and of particular persons" (589). Evil is, in other words, anything that causes others or me immediate pain or that is not useful either to others or to me. Thus when Hume calls remorse "fruitless," he is calling attention to the fact that unlike malice, self-malice is incapable of producing pleasure either for myself or for anyone else.[32] As I argued above, the inverse

[31]For some clear examples, see Nick Smith's recent work on the role of apology in criminal sentencing. Nick Smith, *Justice through Apologies: Remorse, Reform, and Punishment* (Cambridge: Cambridge University Press, 2014).

[32] This is a useful reminder, too, that for Hume, malice is not "evil" in any straightforward sense. It is worth noting that Hume believes that we can come to be ashamed of qualities in ourselves

relationship between the malicious self's pleasure and the other's pain disappears when self and other are one and the same. Whatever happiness I manufacture for myself through comparison is offset by the unhappiness I suffer in order to activate this comparison. In fact, the pleasures of self-malice are *more* than offset when we consider its effects on my friend in distress in Hume's first case. For although I might imagine that I am easing my friend's unhappiness by hurting myself, because she is my friend and not a stranger, she is more likely to sympathize with me than to "reap pleasure from the comparison." If this is correct, then inflicting harm on myself will only compound her pain, and this increased "reflected uneasiness" will then affect me still more. The situation of remorse is no different. In remorse, I hope to ease the pain of a hypothetical self who is experiencing the suffering I *should* be experiencing. As this hypothetical self does not exist in reality, however, this enterprise is absolutely "fruitless." Here I am engaging in what is, for Hume, the worst kind of abstract thinking. I imagine that I can increase pleasure without its being some specific person's pleasure—as if feelings could exist without persons to feel them. Thus, the evil of self-malice is that although it aims to produce pleasure, it only multiplies unhappiness.

IV

In this section, I will argue that although Hume considers remorse to be an evil—as disagreeable as it is fruitless—he stops well short of William Godwin's hope that in the future, "We shall therefore no more be disposed to repent of our own faults than of the faults of others" (PJ, 170). Against Godwin, Hume believes that our actions are in some important sense *our* actions, and

---

that are immediately agreeable to us, but disagreeable to others. He writes: "And this sympathy we sometimes carry so far, as even to be displeas'd with a quality commodious to us, merely because it displeases others, and renders us disagreeable in their eyes; tho' peraps we never can have any interest in rendering ourselves agreeable to them" (589).

that as a result, we must answer for them. This answer, for Hume, is manifested as humility, for it is the passion of humility that forces us to confront our vices head-on. That is, humility forces us to face our moral flaws without taking refuge in remorse's fantasy that these flaws belong to some other person—a person whose suffering can produce pleasure for others, including our own sympathetically generated and ultimately fictional selves. Whereas remorse blurs the boundaries between persons thanks to a disordered imagination that makes perfect sympathy possible, humility does the opposite. The passion of humility is, first and foremost, an original and natural response to moral character. As such, it is a reliable indicator of where the self begins and ends, of which actions are ours and which are not. In other words, humility tells us who we are from the inside, thereby circumventing the confusions of sympathy and misjudgments of comparison. For Hume, the correct moral response to having done wrong is humility because humility is the only passion deaf to the ideas of the outside world. To feel humility is to recognize our vices as well as our virtues and to accept that there is no simple remedy for the evils that come from character. I will suggest that for Hume, the best we can do is to look inward, not with an eye to changing our character—acting out of character—but with a determination to acknowledge ourselves for what we really are.

Hume's commitment to humility as the ideal moral response is, admittedly, not immediately obvious. Indeed, if critics have, as I will argue, failed to recognize the importance of humility to the moral philosophy of the *Treatise,* it is in part because elsewhere, Hume attacks this passion in no uncertain terms. In *An Enquiry concerning the Principles of Morals* (1751), for example, Hume condemns humility as one of a "whole train of monkish virtues" that "serve to no manner of purpose; neither advance a man's fortune in the world, nor render him a more valuable member of society; neither qualify him for the entertainment of company, nor encrease

his power of self-enjoyment."[33] For Hume, again, a virtue is a quality that is "*useful* or *agreeable* to the *person himself* or to *others."* Humility—as debilitating as it is disagreeable—fails to meet any of these four conditions. Humility is a kind of "meanness," a defect in character that turns man into a stupid and sour-tempered creature so unhappy with himself that he is incapable of serving either his own interests or those of his fellows. Indeed, Hume goes so far as to compare "the absence of [self-value]" to the "want of a nose, eye, or any of the most material features of the face or members of the body" (E 61). Like a mutilated sense organ, "an abjectness of character" incapacitates the individual while at the same time making him an immediate object of disgust to both himself and others.

Using the same logic, Hume celebrates (due) pride as the first of the virtues. Always agreeable to its possessor, pride is also a prerequisite for useful—that is, moral—action (E, 69). After all, a person must be conscious of her accomplishments before she can employ them.[34] And although Hume recognizes that self-praise is, generally speaking, an ugly sound to every ear but the speaker's, he insists that, "Where a man has no sense of value in himself, we are not likely to have any higher esteem of him" (E, 61). If a person cannot find anything in his character to be satisfied with, it is unlikely that we who are, inevitably, less invested in this enterprise than he is, will find anything either. We oblige men to temper their pride in company, but this is not so much because we disapprove of proper pride, but rather to guard against unfounded pride. According to Hume, human beings have a natural tendency to overestimate their own virtues. He

[33] David Hume, *An Enquiry concerning the Principles of Morals,* ed. Tom L. Beauchamp (Oxford: Clarendon Press, 1998), 73. Hereafter cited parenthetically in the text and abbreviated E where disambiguation is necessary.
[34] See also Book III of the *Treatise:* "… nothing is more useful to us in the conduct of life, than a due degree of pride…. Whatever capacity any one may be endow'd with, 'tis entirely useless to him, if he be not acquainted with it, and form not designs suitable to it" (THN, 596-97).

explains that "Men have, in general, a much greater propensity to over-value than under-value themselves…. This makes us more jealous of the excess on the former side, and causes us to regard, with a peculiar indulgence, all tendency to modesty" (E, 69). In other words, a person's laudable pride can quickly slide into arrogance, and it is for this reason that we have developed social customs to keep pride in check. This proclivity for exaggeration aside, however, Hume insists that "a certain degree of generous pride" is useful and agreeable to its possessor as well as to those around him. Just as humility is vicious to every eye, pride is a virtue in every sense of the word.

Not surprisingly, intellectual historians have often pointed to Hume's celebration of pride and attack on humility in the *Enquiry* as an important precursor to Nietzsche's "transvaluation of values." Like Nietzsche, Hume turns Christian morality on its head, relegating the enervating "virtue" of humility to the catalogue of vices and elevating pride, the first of the cardinal sins, to the status of virtue.[35] However, this reading of Hume as part of a particular strand of moral iconoclasm that begins with Spinoza and reaches its climax with Nietzsche only makes sense if we give priority to the "cool reflexions" of Hume's later philosophy over the more wide-ranging and less polished musings of the 1738 *Treatise*. In his earlier work, Hume is much more circumspect when it comes to compiling a catalogue of virtues (and vices).[36] Whereas the *Enquiry* argues that pride meets all four criteria of virtue—it is useful and agreeable both to its possessor and to those around him—and humility fails to fulfill any of them, in the *Treatise,* things are not so clear-cut. There, in place of the inversion of values promoted in the *Enquiry,*

[35] See Craig Beam, "Hume and Nietzsche: Naturalists, Ethicists, Anti-Christians," *Hume Studies* 22, no.2 (1996): 299-324; See Peter Kail, "Hume and Nietzsche" in *The Oxford Handbook of Hume,* ed. Paul Russell (Oxford: Oxford University Press, 2016), 755-77.
[36] Hume, *Treatise,* 297. However, see Baier for an argument about how Hume "moralizes" pride in the third book of the *Treatise* ("A Catalogue of Virtues," *A Progress of Sentiments,* 198-220).

Hume suggests that pride and humility occupy an ambiguous moral position: "the former impression [pride] is not always vicious, nor the latter [humility] virtuous."[37] Thus in Book II of the *Treatise,* Hume announces that while we ourselves take pleasure in our pride and displeasure in our humility, when we meet with these qualities in another person, "Humility exalts; but pride mortifies us" (295). This statement is somewhat softened in Book III when Hume distinguishes "an over-weaning conceit" from "a due degree of pride" and, likewise, humility from modesty (596-97). Yet Hume never claims as he does in the *Enquiry* that others find our pride immediately agreeable and our humility immediately disagreeable.[38] At most, he allows that pride is not always offensive to the spectator and humility is not always painful to its possessor: "tho' pride, or self-applause, be *sometimes* disagreeable to others, 'tis always agreeable to ourselves; as on the other hand, modesty, tho' it give pleasure to every one, who observes it, produces *often* uneasiness in the person endow'd with it" (597, my emphasis). Hume's moral evaluation of pride and humility in the *Treatise* is, then, inconclusive. Although he draws attention to the positive aspects of the former and the negative aspects of the latter, he stops short of the complete inversion of values he will propose thirteen years later.

[37] Baier describes the *Enquiry concerning the Principles of Morals* as an "anti-Christian manifesto," noting that "in *Enquiry concerning the Principles of Morals*, the main opponent is identified as the theological moralist." See Baier, "*Enquiry concerning the Principles of Morals:* Incomparably the Best?" in *A Companion to Hume,* ed. Elizabeth S. Radcliffe (Singapore: Blackwell, 2008), 315.

[38] Perhaps the most glaring difference between the discussion of pride and humility in the *Treatise* and that in the *Enquiry* is the portrayal of humility in the later text as an "abjectness of character" akin to a debilitating physical deformity. This analogy grounds Hume's claim that humility is immediately disagreeable to others as well as to ourselves. Without this image, it is difficult to understand why someone else's low opinion of himself would cause us pain. The *Treatise'*s account of humility—it is a source of pleasure for the observer—is much more convincing.

The *Treatise* calls then for a more nuanced reading of humility. To begin, we must recognize that just as humility is not merely a synonym for remorse, it is not merely the opposite number of pride. This may seem counterintuitive, for at times, Hume seems to plainly claim that the only difference between the two passions lies in the fact that one ends in pleasure and the other in pain. Thus having outlined the underlying mechanism of the emotion of pride at the beginning of Book II, he declares:

> What I have said of pride is equally true of humility. The sensation of humility is uneasy, as that of pride is agreeable; for which reason the separate sensation, arising from the causes, must be revers'd, while the relation to self continues the same. Tho' pride and humility are directly contrary in their effects, and in their sensations, they have notwithstanding the same object; so that 'tis requisite only to change the relation of impressions, without making any change upon that of ideas. (288-89)

Hume's point here is that both pride and humility are produced courtesy of "a double relation of impressions and ideas." In each case, the self becomes the object of a passion through its connection to something whose qualities inspire either pleasure or pain. This "something"—Hume calls it the "cause" or "subject"—might be any number of things, including character, action, body, property, and family. As Hume explains, "Any thing, that gives a pleasant sensation, and is related to self, excites the passion of pride, which is also agreeable, and has self for its object" (288). So when I am proud of my house, for example, it is because this house (the subject that causes the passion) has the quality of beauty (which gives pleasure) and it is mine (it is connected to me by virtue of being my property). Thus while the sight of any beautiful house will give me pleasure, when the beautiful house belongs to me, it makes me think of myself, and in doing so, produces an additional pleasurable feeling: pride. Likewise, any ugly house will pain

me, but if the ugly house is mine, then it will be a source of humility for me as well. Thus both of these passions are generated by a "double relation of impressions and ideas." The subject (the idea of the house) is related to the object (the idea of self) and the impression of sensation (in the case of a beautiful house: pleasure) is related to the secondary impression (the passion of pride). This double relation is a precondition for all of the "indirect passions," and without it, there would be no such thing as pride or humility.

Hume's sketch of the mechanism that generates pride and humility explains why Book II begins with a discussion of the indirect passions rather than what would appear to be, as many critics have noted, a more natural starting point: the simpler direct passions. According to Hume, pride and humility generate the idea of self, and in this sense, they provide a solution to the skepticism about personal identity that marked the end of Book I. As Adela Pinch succinctly puts it:

> … though pride is the passion that seems most to originate *from* a sense of personal identity, that seems unimaginable without one, in fact for Hume the reverse is true: the passion of pride produces a sense of self.… The self as 'object'… is not the 'cause' of the passion, but rather its *effect*…. the cause is rather the feelings aroused by other objects and the qualities of objects that are close to self—the self's house, horse, or clothes, the temperature of the climate in which one was born, the fertility of one's native soil…. [Passion] ties the "bundle of perceptions" into a recognizable human form we can claim as our own. (22; 24)

In other words, the idea of self is not the source of pride. It is, rather, the product of it.[39] What this means is that self is crucially dependent on the outside world. Indeed, as critics have been quick to point out, it is no coincidence that Hume first illustrates the operations of pride by way of a property relation. This example usefully emphasizes the fact that it is an outside object—my beautiful house (and as Pinch continues: one's horse, clothes, native climate…)—that inspires the emotion that fixes the self. In other words, pride outlines self from the outside, so to speak, and it is only thanks to my relations to the object world that I am able to think of myself as a "self" at all.

Expanding on this picture of self as a product of external relations, critics have highlighted the social nature of Humean pride. Hume's self is, at bottom, a social self, for the passions that give it shape need propping up by the passions of others. Thus Hume writes:

> But beside these original causes of pride and humility, there is a secondary one in the opinions of others, which has an equal influence on the affections. Our reputation, our character, our name are considerations 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. (316)

According to Hume, I may hold the deeds to the most beautiful house in the neighborhood, but unless there are other people to admire the beauty of my house and, by extension, to respect and sympathetically enjoy my power and wealth, I will feel little in the way of pride. I need

[39] Or rather, as Baier notes, pride *reactivates* the idea of self: "Although the particular passions of pride and humility are said to 'produce the idea of self'… all Hume here means is that they reliably reactivate that idea, not that it is derived from them, as ideas are from impressions which they copy. Generalization or what Hume terms 'abstraction' of the objects of pride and humility no more generates the idea of self than generalization of the objects of love and hate generates the idea of other people" in *Death and Character*, 132.

reassurance from others that my house really is as beautiful as I think it is and that I really am as happy as I feel. Indeed, the force of sympathy is so strong, says Hume, that another man's pride in his rather average house can make me see my house as less beautiful than it is—thereby diminishing my pride and, by extension, my sense of self—even when I know quite well that his house is inferior to mine and that his pride is unwarranted (595). Thus for Hume, as Pinch observes, "Not only is our sense of self constituted through having feelings of pride inspired by contingent objects, but the feelings we have about ourselves may be largely versions of other people's feelings about us" (25). In a sense then, my idea of self is as much a product of others' love and respect as it is of pride. Such are the effects of sympathy.

In noting the external aspect of pride, critics have tended to downplay Hume's repeated reminder that although we can take pride in or be humiliated by anything related to us—our houses, horses, clothes—the primary source of these emotions are our virtues and vices. Thus Hume observes that "the most obvious causes of these passions" are "VICE and VIRTUE," and again: "pride and humility have the qualities of our mind and body, that is *self,* for their natural and more immediate causes" (295; 302). In other words, houses, horses, and clothes are secondary and weak causes of pride or humility. These passions are first and foremost concerned with our moral character, and precisely for that reason, self is *not* dependent on "contingent objects." It is character—our durable qualities of mind—that most immediately defines the self. What is more, in the case of humility, sympathetic reinforcement is shunned. This marks an important difference between pride and humility. Whereas pride seeks out the approval of others, humility would bury itself in obscurity—as in Hume's example of those men who choose to live among strangers rather than suffer the contempt of those who know the reasons for their feelings of humility and, to some degree, share in that pain. The sympathy of others amplifies our

feelings, and it makes sense that when we are feeling unhappy, we do our best to avoid this magnification of our pain. Thus humility hides from the world's opinion for exactly the same reason that pride seeks it out: sympathy reinforces what the self already knows.

To say "the self already knows" might seem like a wrong turn of phrase, yet I believe that in some real sense this is the best representation of what Hume is trying to say about humility. Insofar as pride and humility stem from moral character, they are neither of them particularly concerned with the beauty of houses or the eminence of family members. In their purest form, they are exclusively preoccupied with what is truly part and parcel of the moral self. When we feel these passions, we are both acknowledging that we have such and such character trait and admitting that this character trait is a virtue or a vice. Nor is there any possibility of error here. For Hume, pride and humility are infallible in their evaluations of the self's character. He explains:

> … the peculiar object of pride and humility is determin'd by an original and natural instinct, and… 'tis absolutely impossible, from the primary constitution of the mind, that these passions shou'd ever look beyond self, or that individual person, of whose actions and sentiments each of us is intimately conscious…. For this I pretend not to give any reason; but consider such a peculiar direction of the thought as an original quality. (286)

This is, in other words, an internal view of the self that is accurate precisely *because* it is internal, which is to say original and natural.[40] There is no interference here from the principles of comparison or sympathy. This is self-perception without feedback from the external world.

[40] Eugenio Lecaldano makes a similar argument, but he argues that persons adopt a general point of view vis-à-vis themselves. My claim is that humility gives us a glimpse of our real selves precisely because it is indifferent to the feelings and opinions of the outside world. See Lecaldano, "The Passions, Character, and the Self in Hume": 175-193

This is all the more true for humility, for this passion is not reliant on social-referencing in the way that pride is. Humility is, for Hume, both the most independent of the passions and the most truthful.

Humility tells the truth about moral character. It does not, however, call for action. One of the qualities that distinguishes pride and humility from the rest of the passions is that they alone are "pure emotions in the soul." What Hume means by this is that pride and humility are fundamentally non-motivating. They do not, in other words, point to any action or end. He states:

> The passions of love and hatred are always followed by, or rather conjoined with benevolence and anger. It is this conjunction, which chiefly distinguishes these affections from pride and humility. For pride and humility are pure emotions in the soul, unattended with any desire, and not immediately exciting us to action. But love and hatred are not compleated within themselves, nor rest in that emotion, which they produce, but carry the mind to something farther. Love is always followed by a desire of the happiness of the person beloved, and an aversion to his misery: As hatred produces a desire of the misery and an aversion to the happiness of the person hated. (367)

Nature decrees that love and hatred will move us to benevolent and angry actions, respectively. By contrast, pride and humility are complete in themselves and therefore inert. Baier has argued that this claim does not make sense for humility. She reasons that while pride might be a pure emotion, humility seems to naturally imply a wish to escape whatever is causing it. As she observes, "one would expect the person who is ashamed of his ragged coat to desire a better one. Hume himself describes humility as unlike pride in being an 'uneasy' passion. He may, then,

have erred in claiming purity for humility as well as pride."[41] Hume does not, so far as Baier is concerned, give a good reason for how humility can be both uneasy and non-motivating. Yet what kind of desire could accompany the shame attached not to a ragged coat, but to a constitutional vice? Hume allows for moral reformation, but he also believes that some of our most essential character traits are too firmly rooted to change. To feel humility in relation to such qualities, when they are vicious, is inevitable. To try to change them by punishing them is evil as well as futile. It turns out that for Hume, humility is moral exactly because it is does not motivate action. As such, it presents a real alternative to the passion of self-malice, which tells us—wrongly—that we can undo past injuries by injuring ourselves. Instead, humility asks us to rest a moment with our vices and to feel the uneasiness that goes hand in hand with being a moral individual. For Hume, this is punishment enough.

[41] See Annette Baier, "Master Passions," in *Explaining Emotions,* ed. Amelie O. Rorty (Berkeley: University of California Press, 1988), 408-09.

Chapter Two

So Close a Connection: Painful Associations in Samuel Richardson's *Clarissa*

Near the beginning of Samuel Richardson's *Clarissa* (1748), Anna Howe warns her friend that her love for her family is clouding her judgment. She writes:

> But you are so tender of some people who have no tenderness for anybody but themselves, that I must conjure you to speak out. Remember that a friendship like ours admits of no reserves. You may trust my impartiality. It would be an affront to your own judgement if you did not; for do you not *ask* my advice? —And have you not taught me that friendship should never give a bias against justice?[1]

Usually it is Clarissa who chides Anna for making her the object of a "partial love," but here the roles are reversed. Anna suspects that, in order to protect the Harlowes from the censure they deserve, Clarissa has told only part of her tale. Following in Anna's footsteps, Richardson's critics have frequently observed that the novel fails to fulfill its promise of being a whole, unified story. Each narrator gives a different account of events so as to curry sympathy from the reader.[2] But Anna's claim is not quite this. She believes that Clarissa's narrative is biased, but not in her own favor. Clarissa's natural inclination is to minimize her family's faults as much as possible,

---

[1] Samuel Richardson, *Clarissa, or The History of a Young Lady*, ed. Angus Ross (London: Penguin, 2004), 67-68. All references to this text are hereafter cited parenthetically by page number.

[2] There is hardly a critic who has not paused over the problem of the novel's competing narratives. Thomas Keymer puts it most succinctly, perhaps: "Clarissa makes available no such thing as a neutral text, no writing that does not in the first place serve its writer." *Richardson's Clarissa and the Eighteenth-Century Reader* (Cambridge: Cambridge University Press, 1992), 55.

even at her own expense. According to Anna, Richardson's heroine is "so tender of some people" that rather than do herself justice, she would do them *more* than justice.

In some respects, Anna's comment seems wide of the mark. If indeed Clarissa was initially determined to exonerate the Harlowes in the eye of the World—regardless of the cost to herself—by the end of the novel, her primary concern is "to leave behind [her] such an account as may clear up [her] conduct to several of [her] friends" (1173). In other words, as the story wears on, Clarissa's desire to defend the Harlowes from public censure is overshadowed by her need to defend herself to her family—and to herself. The change is gradual. As Clarissa becomes more and more isolated, she becomes more and more introspective. When finally she finds herself well and truly on her own, one of the few consolations she has left is that, having turned her eye inward, she can "acquit [herself] to [*herself*]" (1139). As Clarissa shrinks into herself, her attachment to others becomes weaker, and although she never ceases to love the Harlowes, ultimately she comes to value her own reputation over theirs. Partial love has its limits.

Given this progression, it is small wonder then that Anna's insight about Clarissa's partiality has failed to make an impression on readers. For most critics, it is *Clarissa's* representation of first-person experience—of Clarissa's shrunken world—that makes the novel so extraordinary. Samuel Johnson pronounced Richardson's novel "the first Book in the world for the knowledge it displays of the human Heart" and, in doing so, gave birth to a critical tradition in which Anna's concept of an outward-looking "partial love" has no obvious place.[3] Readers have sought to mine Clarissa's heart, so to speak, delving ever deeper in an effort to find

[3] Samuel Johnson quoted in "Recollections of Dr. Johnson by Miss Reynolds," in *Johnsonian Miscellanies*, ed. George Birbeck Norman Hill, 2 vols. (New York: Harper, 1897), 2:251.

out what exactly the novel knows (and displays) about individual human psychology.[4] Following Ian Watt's *The Rise of the Novel,* critics have been preoccupied with the way the "novel's subjective mode of vision" allows us privileged access into the inner lives of its characters, particularly Clarissa's.[5] Many have, like Watt, detected a psychological tension within Clarissa that licenses doubts concerning her true motives; these doubts then operate almost tautologically, serving to further "deepen" her character by figuring her as a split subject driven by unconscious desires.[6] For these critics, as for Clarissa herself, private passions hold the key to who she is and what she does. Such readers submit to the challenge of judging Clarissa's heart rather than her actions, and questions about Clarissa's moral character—and, by extension, her culpability—are recast as an epistemological problem: how can we know what Clarissa's true intentions are?

There is another line of criticism, however, that pulls against the "rise of the individual" account made popular by Watt. William Warner's generalized deconstructionist approach paints

---

[4] William Beatty Warner, Leopold Damrosch, and Nancy K. Miller, all argue that Clarissa does not know her own heart and harbors repressed feelings for Lovelace. See Warner, *Reading Clarissa: the Struggles of Interpretation* (New Haven: Yale University Press, 1979); Damrosch, *God's Plot and Man's Stories: Studies in the Fictional Imagination from Milton to Fielding* (Chicago: University of Chicago Press, 1985); Miller, *The Heroine's Text: Readings in the French and English Novel, 1722–1782* (New York: Columbia University Press, 1980). For more recent interpretations, see Katherine Binhammer, "Knowing Love: The Epistemology of *Clarissa," ELH* 74 (2007): 859-79; Hina Nazar, "Judging Clarissa's Heart," in *Enlightened Sentiments: Judgment and Autonomy in the Age of Sensibility* (New York: Fordham University Press, 2012), 59-81. *Clarissa* has also caught the attention of critics working from a cognitive science angle. See for example Lisa Zunshine, "Richardson's *Clarissa* and a Theory of Mind," in *The Work of Fiction: Cognition, Culture, and Complexity,* eds. Alan Richardson and Ellen Spolsky (Burlington: Ashgate, 2004), 127-46.

[5] Ian Watt, *The Rise of the Novel: Studies in Defoe, Richardson, and Fielding* (Berkeley: University of California Press, 1962), 238.

[6] Many have seen Clarissa as the first modern secular subject. Watt argues that Clarissa wonders whether her original hope that she could reform Lovelace was a mask for sexual desire. Terry Eagleton's claim that Anna is part of Clarissa's unconscious is particularly suggestive. See Eagleton, *The Rape of Clarissa: Writing, Sexuality, and Class Struggle in Samuel Richardson* (Minneapolis: University of Minnesota Press, 1982).

a world in which action, and thus culpability, can transcend the individual. For Warner, the novel enacts a "struggle for interpretation," and plot takes a backseat to the battle for narrative authority between Clarissa and Lovelace. Intention, causation and, ultimately, blame are all determined by the competition playing out on the "textual plane." According to Warner, it is because it is Clarissa who ultimately wins this competition that we read the novel not as a comedy (as Lovelace would like us to), but as a tragedy wherein everyone has a hand in the heroine's mysterious death. As he explains, "by dying without any explicit 'cause' for doing so, Clarissa allows blame for her death to spread and infect everyone who has touched her life. Everybody becomes a subordinate 'cause,' in the orchestration of one great 'effect'—the death of Clarissa Harlowe."[7] In other words, because no one is directly responsible for Clarissa's death, everyone is indirectly responsible. Thus unlike the Watt line of criticism that seeks to individuate actions and their authors, Warner's deconstructionism sees in Clarissa's death a transpersonal action in which the novel's characters are all complicit.

Warner's work anticipates that of recent critics who have begun to look more closely at the ways in which the novel frustrates our impulse to demarcate persons and their actions. As Frances Ferguson, Sandra Macpherson, Jonathan Kramnick, and others have noted, *Clarissa* invites us to think about a world in which there is no straightforward correspondence between intention and action and, as a result, no firm psychological grounds for attributing a particular action to a particular person. For Ferguson and Kramnick, this break between intention and

[7] Warner, *Reading Clarissa*, 113. Warner's deconstructionist approach has been much criticized and his claim that Clarissa is complicit in her rape is still a sore point for many. See especially Terry Castle, *Clarissa's Ciphers: Meaning and Disruption in Richardson's Clarissa* (Ithaca: Cornell University Press, 1982).

action undermines familiar notions of moral accountability.[8] Macpherson takes a slightly different tack. Building on Warner's insight that our inability to identify a specific cause or person behind an action does not necessarily frustrate our ability to lay blame, she argues that "the novel complicates the question of… agency but not the question of… responsibility" (82). According to Macpherson, *Clarissa* follows the logic of strict liability, holding persons accountable for the effects of their actions regardless of how unexpected or unintentional or removed they are. What this means, ultimately, is that the very notion of "their actions" becomes blurry as each deed is incorporated into a single causal chain, i.e. the plot. Macpherson explains: "[Strict liability] means… that there is nothing we can unequivocally claim is *not* our doing. …the novel *does* divide blame, distributing it between Lovelace and Clarissa and among all those whose actions are, however accidentally, the sine qua non of the novel's tragic plot" (76-77). In other words, persons are inseparable from each other and the plot that links them *because* there is a plot that links them.[9] For Macpherson then, as for Warner, because every particular

[8]See Frances Ferguson, "Rape and the Rise of the Novel." *Representations* 20 (Autumn 1987): 88-112, and Jonathan Kramnick, *Actions and Objects: from Hobbes to Richardson* (Stanford: Stanford University Press, 2010). Discussing Clarissa's rape, Ferguson explains that despite Lovelace's attempts to wrest retroactive consent from Clarissa (which would negate the rape and thereby absolve him of guilt), the fact that Clarissa was unconscious during the event itself means that, under the law, it was and will always be impossible for her to give her consent, even if she would like to. In other words, although we often define rape in terms of the victim's mental state, i.e. non-consent, the law holds Lovelace responsible regardless of how Clarissa herself sees the experience. Kramnick offers an interpretation of Clarissa's death that would exonerate everyone, including Lovelace: "as it considers the problem of suicide, *Clarissa* presents something like a world with events yet no actions. There are still mental states in this world; they just observe events as they transpire. The end of the novel may be an invitation to consider this possible world" (*Actions and Objects*, 230). Things that happen—injuries—have either mental or physical causes. If the former, then there is a guilty agent, if the latter, then there is an unfortunate event. If minds have no efficacy at all—and this is to take Clarissa's commitment to her "guiltless intentions" to its logical limit—then we have a world with events, but no actions (and no responsibility).

[9] The difference, however, is that whereas Warner argues that events become linked through storytelling, Macpherson holds that they are connected through material causation. Macpherson's

action in the novel tends to the "one great effect," everyone is responsible for causing Clarissa's rape and death.

Clarissa is, as Warner and Macpherson show, a novel about the association of persons and the spread of actions. I share their feeling that in the text, blame somehow seems to exceed the bounds of the individual. However, while Warner and Macpherson argue that more or less all the novel's characters are caught up in the tangled plot (conceived either as a narrative structure or as a causal chain) that leads to Clarissa's death, I want to focus on the way guilt appears to "infect" (to use Warner's word) some persons but not others. After all, the mystery surrounding Clarissa's slow demise did nothing to prevent either Richardson or his audience from censuring some characters but not others. Most readers are only too willing to point fingers, thinking little about the actual physical cause of Clarissa's death. And while Lovelace's and Clarissa's complicity might be points of contention, other cases appear clear-cut: Mrs. Sinclair deserves blame, for example, and Mrs. Norton, praise. The novel complicates these lines somewhat—writing to Belford, Anna seems to exonerate Sinclair, exclaiming that Lovelace's "guilt… will be adjudged all his own"—but not that much: Anna quickly backtracks, observing of "those common wretches" that "they too may meet with their punishment: as it is already begun" (1454).[10] In any case, there is no avoiding the verdict decisively pronounced by the book's Conclusion. Each character gets what he deserves and rare is the reader for whom Belford's final summation is a surprise. In its last pages then, *Clarissa* is committed to firming up rather than

account is an inversion of Warner's. He holds that all the characters are indirectly involved in the absence of a material cause. Macpherson agrees with Warner's conclusion, but because she sees material causes as all intertwined.

[10] For a discussion of Anna's hedging, see E. Derek Taylor, *Reason and Religion in 'Clarissa': Samuel Richardson and 'the famous Mr. Norris, of Bemerton'* (Farnham: Ashgate Press, 2009), 118.

breaking down the distinction between "good" (innocent) persons and "bad" (blameworthy) persons. It seems therefore intuitively wrong to group all the characters together.

The intelligibility of the novel's account of moral responsibility—and ultimately, of justice—turns on Richardson's ability to lay bare the logic behind what looks like the limited circulation of guilt. Accordingly, the following reading revisits the notion of transindividuality outlined by Warner and Macpherson, but with an eye to defining its contours. The danger of thinking about guilt as selectively transpersonal is that it starts to feel disingenuous, a means for us to cherry-pick our villains according to our inclinations rather than their actions. Richardson's novel confronts this problem head-on: it is precisely this kind of bias that Anna is attributing to Clarissa when she accuses her of censoring her story to shield her family from blame. Without exactly conceding Anna's point, Clarissa's defense throughout the novel will be that however badly they treat her, her kin are entitled to her protection. What she realizes—and this insight belongs to Clarissa alone—is that there is no undoing a familial bond. The biological, social, and affective ties that link the Harlowes are indissoluble, and if Clarissa would spare her family pain, it is partly because she is compelled to share their suffering just as if it were her own. The argument of this essay is that in *Clarissa*, keeping persons separate is impossible not because a single action can, regardless of its specific cause, implicate everyone, but because persons are always affectively connected to (only) some other persons: namely, their kin. It is suffering, not blame, that exceeds the person. What the novel shows then is that although we tend to think of justice as a balancing of the books (we restore order by repaying harm with harm), in fact, punishment and culpability never quite line up. One cannot inflict pain upon a guilty party without hurting the innocent people around her. It is this insight—that owing to our sympathetic

associations, we are punished for the faults of our friends and vice versa—that makes *Clarissa* tragic.

I

Clarissa's account of suffering suggests that the retribution we associate with justice is neither universal nor individual. This is a familiar idea. While we may find it difficult to accept a model of justice that is selectively transpersonal, so to speak, the Greek tragedians had no such qualms. The notion that one might be fairly punished for the actions of some persons but not others made sense to them. Indeed, the story of Oedipus—the paradigmatic example of Aristotelian tragedy—can be read as a quest to find the middle ground between Watt's individualism and the blanket complicity suggested by Warner and Macpherson. The play begins with the kind of generalized punishment seen in *Clarissa*'s ending. An unexpiated act of regicide, Laius's murder, has brought a plague upon Thebes. This sprawling blight is presented as a problem that must be solved by finding and banishing the particular individual responsible for the king's death. Order will be restored by moving away from the collective suffering of the people of Thebes to a specific punishment that targets Laius's killer. On the face of it then, the story of Oedipus appears to favor an account of moral responsibility that (and this sounds more like Watt) revolves around individual agency. But to read this way is to understand Oedipus's plight only by halves. It is to forget that it is precisely his lack of agency, his powerlessness to thwart Fate, that makes him tragic. In defying the oracle's admonishment to die childless, Laius condemns himself to a punishment that Oedipus cannot but share. Oedipus commits parricide, incest, and regicide for no other reason than that he is his father's son. In some real sense then, Oedipus's crimes (actions) belong more to his parents than to him, and it is his bond with them that makes him a suitable candidate for punishment.

This loosely oedipal model of justice—what I will here call punishment by association—has hardly figured in the secondary literature on *Clarissa*. Macpherson's recent study on the novel as a tragic form, which draws on Greek tragedy and eighteenth-century legal doctrine to show that "the realist novel is a project of blame not exculpation," comes closest to the reading I am proposing here.[11] However, whereas Macpherson is interested in the detachment of agency from responsibility—how persons can be *blamed* for the actions of others—I am interested in the detachment of responsibility from punishment—how persons can be *punished* for the actions of others (without their being blamed for them).[12] I want to turn here to a critical discussion that will help to bring *Clarissa*'s particular account of punishment by association into sharper focus. George Jeffreys and William Duncombe stand alone among Richardson's contemporaries for their extended conversation about Aristotelian tragedy and, more importantly, for their intuition that the novel's asymmetrical justice—too one-sided to be either impartial or haphazard—has a forerunner in the classical notion of punishment by association. Jeffreys was the first to suggest this parallel, urging Duncombe to consider that in Greek tragedy (*Oedipus* being a case in point) punishment is not reserved for the guilty. He writes:

> … here, to qualify it still more, there is so close a connection between them [Oedipus and his parents], that they cannot be separated; a consideration, which, in other cases, as well as those of parents and children, will, by way of exception to a general rule, reconcile us

[11] Macpherson, 13.

[12] Punishment by association looks very similar to a sub-branch of felony murder: vicarious liability. Macpherson explains that under the doctrine of vicarious liability, "principals [are held] accountable for the acts of accessories and accessories accountable for the acts of principals" (*Harm's Way,* 14). Thus in law (and in *Clarissa*), a master can be held responsible for the deeds of his servant regardless of whether she was acting on his express commands. Again, the key difference between vicarious liability and punishment by association is that in the former instance, people are *blamed* for the actions of others, in the latter, they are *punished* for them.

> to the distress of the innocent, occasioned by their attachment to the guilty: Monimia, for instance, is entirely innocent, and yet, though her death moves pity, and strikes terror, we behold it without repining at Providence; her destiny is united to that of Castalio; so that the natural consequence of his fault are the ruin of both…[13]

Oedipus's lot is hard, but although we sympathize with him, explains Jeffreys, we can also see a logic to it. Oedipus may in some sense be faultless, but Laius and Jocasta are guilty, and the fate of the child is inextricably intertwined with that of the parents. Likewise, Castalio's woes are also his wife's, and his misery must be hers as well.[14] In short, sometimes an innocent person may be so closely connected with a guilty one that she is bound to share the latter's fate. Being punished does not imply culpability then, nor does virtue guarantee reward. The innocent may suffer together with the guilty for no other reason than because there is no way of parting them.

Not everyone will be "reconcile[d]… to the distress of the innocent" as simply as Jeffreys. If it is easy for us to see the tragedy in punishment by association, it is correspondingly difficult to understand the justice in it. Why should persons be stripped of their individuality at the very moment when it most concerns them? The story of Oedipus provides no real explanation—but the Monimia example does. The character interactions in Thomas Otway's play allow Jeffreys to make the argument that punishment by association is in some sense the most effective form of punishment. He explains that "by involving [Monimia] in that ruin, [Castalio] is punished still more severely, than if he had suffered alone."[15] Monimia's misery means that Castalio is punished twice over: the pain he gives to Monimia redounds back upon

---

[13] George Jeffreys to William Duncombe, 28 February 1749, in *Clarissa: The Eighteenth-Century Response 1747-1804*, ed. Lois E. Bueler, 2 vols. (New York: AMS Press, 2010), 1:36. Jeffreys seems to have read the Postscript before finishing the novel.
[14] Monimia and Castalio are characters from Thomas Otway's *The Orphan* (1680).
[15] Jeffreys, Letter to Duncombe, 36.

him and thereby becomes a second source of agony. Punishing the innocent is, in other words, a means of punishing the guilty still further.

Although the parallel between Monimia and Clarissa is not made explicit in Jeffreys's letter, we can look to Lovelace to fill the gap. After her death, Lovelace claims Clarissa (and her pain) as his own. He insists that "I was her husband. What though I treated her like a villain? Do I not pay for it now?" (1384). Lovelace aligns Clarissa's story with Monimia's. He is her Castalio, so to speak, and they are both punished for his faults: he doubly so, through her death as well as his own. In other words, Lovelace's suffering is compounded by Clarissa's, and it is exactly this transmission of pain that justifies punishment by association. Characters like Monimia and Clarissa matter less as individual moral agents and more as vehicles for suffering. They are, first and foremost, instruments for punishing the guilty persons to whom they are attached. According to this logic, Clarissa has it backwards when, relatively early on, she complains that she is "but a *cipher,* to give *him* [Lovelace] significance, and *myself* pain" (567). The reverse is true: Clarissa gives herself significance by giving Lovelace pain.

In calling herself a cipher, Clarissa seems to imagine that she exists solely in terms of her relationship with Lovelace. This moment is, however, an exceptional one. Throughout the novel, Clarissa categorically denies any attachment to Lovelace. Lovelace's account of their union leaves room for this challenge. It is immediately after her death that Lovelace is most adamant about his claim upon Clarissa. He performs his grief "in the air of a tragedian," reciting his mad letter—the counterpart to the delirious fragments Clarissa composes after her rape—to an audience of family and friends. In doing so, he transforms epistle into theater and mere rhetoric into tragic drama (1382). Lovelace makes a Monimia of Clarissa just as, in his younger days, he would transform a lady into "a Cynthia, a Stella, a Sacharissa" in order to compose a "sonnet,

elogy, and madrigal" in imitation of his favorite poets (143).[16] There is, in other words, something artful in Lovelace's distraction. Clarissa's death is an opportunity for misinterpretation. It allows Lovelace to imagine her forgiveness as love, and further still, as marriage. He challenges Belford: "For am I not her husband? And being forgiven by her, am I not the chosen of her heart? What else signifies her forgiveness?" (1384). In fact, before her death, Clarissa specifically states that her forgiveness precludes attachment. She explains to Colonel Morden that "I can indeed forgive him [Lovelace]. But that is because I think his crimes have set me above him. Can I be above the man, sir, to whom I shall give my hand and my vows?" (1301).[17] Clarissa is far from accepting Lovelace's proposals. For her, there is an unbridgeable gap between them.

If, as Jeffreys claims, "so close a connection" is a precondition for transpersonal punishment, then much hinges on whether Clarissa really is, as Lovelace would have it, his "wife; … to all eternity" (1385). Clarissa insists that Lovelace is claiming a false intimacy with her, and her repudiation of him would seem to undermine Jeffreys's reading. In fact, however, the novel's most explicit articulation of the concept of punishment by association comes not from Lovelace, but from Clarissa, who describes her relationship with her family in exactly these terms. Thus in one of her most lucid attempts to make sense of the "strange situation" that has descended upon Harlowe Place, Clarissa imagines that she has been chosen as a medium for divine retribution. She tells Anna that "surrounded by our heaps and stores, hoarded up as fast as

[16] In the same letter, Lovelace adopts lines from Otway's tragic play, *The History and Fall of Caius Marius* (1680) and inserts Clarissa's name for Lavinia's: "CLARISSA!—Oh, there's a music in the name…" (144).

[17] And to Anna: "*Beneath* my pity as he is, I nevertheless pity him!—But this I could not do if I still loved him: for, my dear, one must be greatly sensible of the baseness and ingratitude of those we love. I love him not, therefore! My soul disdains communion with him" (1116).

acquired, we seemed to think ourselves out of the reach of the bolts of adverse fate…. Who knows what the justice of Heaven may inflict in order to convince us that we are not out of the reach of misfortune" (333). Here Clarissa speaks in terms of "we," putting herself on a level with her family. She, like them, has been "too happy" to be humble and they are all being punished for their impious presumption. Later on, however, Clarissa seeks solace in the hope that she will be merely a conduit for unhappiness and not the reason for it: "If *I* am to be singled out to be the *punisher* of myself, and family, who so lately was the *pride* of it, pray for me, my dear, that I may… be enabled to support my character, so as to be *justly* acquitted of willful and premeditated faults" (333). Clarissa's formulation here is telling. In asking that she be able to "support [her] character," Clarissa is effectively clearing herself of blame. After all, if she can only continue as she is, she will be "justly acquitted." [18] Thus Clarissa comes to see herself as the instrument of "a perverse fate" that works against her family and, because she too is a Harlowe, against herself. Clarissa is blameless, but as the vehicle of Providence, she is destined to suffer alongside her guilty relations.

With Clarissa's intuition in mind, I want to bring Jeffreys's theory of punishment by association back to the fore by shifting attention away from Clarissa's relationship with Lovelace. Although this precarious romance is, for many readers, the central axis of the story, this is less true for Richardson's heroine. For most of the narrative, it is the Harlowes who dominate Clarissa's thoughts, not Lovelace. Almost as soon as she steps out from Harlowe Place, she begins to petition for a reconciliation, offering to give up Lovelace's proposals if they will give up those of Mr. Solmes. She seeks intimacy with her family and distance from Lovelace in

[18] It is worth noting that "support" means "to maintain in being and action" as well as "to provide evidence or authority for, or corroboration of… to bear out" (*Oxford English Dictionary,* s.v. "support," accessed August 14, 2015, http://www.oed.com.)

equal measure; indeed, these two impulses add up to the same thing. Thus Clarissa consistently refers to Lovelace as "Mr. Lovelace"—and later, as "that man." In contrast, when she talks about her relations, it is always in terms of their familial and familiar association with her: "father," "mother," "uncle," "aunt," "brother," "sister." These titles are much more than social roles for Clarissa, and the people who bear them are not merely placeholders. Anna is wrong therefore to interpret Clarissa's reluctance to settle marriage terms with Lovelace as an impractical modesty ("You must be father, mother, uncle to yourself…. What room for delicacy now?" urges Anna [588]). Clarissa is not standing on punctilio. She believes that a father's—her father's—authority is unassailable, and any exchange of vows with Lovelace would be meaningless without his consent.[19] For Clarissa, Mr. Harlowe's power over her is dictated by indissoluble bonds.[20] As she explains to an incredulous Lovelace: "*Duty* and *nature,* sir, call upon me to make the submissions you speak of: there is a father, there is a mother, there are uncles… to justify and demand those submissions" (592). While Anna, Lovelace, and "the world" more generally all argue that her family's ill-treatment of her mitigates the obedience she owes them, Clarissa argues that familial ties are such that breaking them—independence—is impossible. Even in the immediate aftermath of what she believes is the fulfillment of her father's curse, her rape, Clarissa insists upon her right to call herself his daughter: "But your heavy curse, my papa—Yes,

[19] Clarissa believes that her father retains the power to confirm or absolve her (his child's) promises. Richardson's footnote explains her logic. Under "the Old Law," "the vows of a single woman, and of a wife, if the father of the one, or the husband of the other, disallow of them, as soon as they know them, are to be of no force" (361).

[20] Critics have long acknowledged Clarissa's ever-present sense of duty to her father. See Nazar, "Judging Clarissa's Heart"; Wendy S. Jones, *Consensual Fictions: Women, Liberalism, and the English Novel* (Toronto: University of Toronto Press, 2005); Gordon Fulton, *Styles of Meaning and Meanings of Style in Richardson's 'Clarissa'* (Montreal: Mcgill-Queen's University Press, 1999); Lois E. Bueler, *Clarissa's Plots* (Newark: University of Delaware Press, 1994); Florian Stuber, "On Fathers and Authority in *Clarissa*." *Studies in English Literature, 1500-1900* 25 (1985): 557-74.

I *will* call you papa, and help yourself as you can—for you are my own dear papa, whether you will or not—and though I am an unworthy child—yet I *am* your child—" (890). Mr. Harlowe might withdraw his protection, but he cannot disown her.

Like Anna and Lovelace, readers have often found Clarissa's unwavering devotion to the Harlowes perplexing.[21] After all, even her friendship with Anna is not without reservation. Clarissa frequently praises her confidante by telling her that "I should not have loved you so well as I really do [if…]," and just as frequently, she worries that she is no longer worthy of Anna's favor.[22] The affection she feels for Anna seems to have more in common with her "conditional liking" of Lovelace than with the boundless love she has for her family. Clarissa's strange preference for her cruel relations over her "truly sympathizing and unalterable friend" feels almost perverse. Yet the idea that familial ties necessarily trump other social relations was a familiar one in the period. As Ruth Perry points out, the notion of *cri du sang*—"a primal and palpable" sympathetic link between blood relatives—was a popular plot device in eighteenth-century fiction, where the reunion of estranged family members was routinely characterized by instant mutual attraction and affection.[23] In this context, Clarissa's partial love is neither

---

[21] See Keymer for an account of Richardson's extensive correspondence with his female readers, many of whom were puzzled by Clarissa's sense of filial duty (Keymer, "Casuistry in *Clarissa*" in *Richardson's Clarissa*, 85-142).

[22] Clarissa's and Anna's love for one another is asymmetrical. After Clarissa elopes, Anna cries, "Love you still!—How can I help it, if I would?" (371). Clarissa's affection, on the other hand, is conditional: "for if you, even you, were really so loath to own a mistake… I believe I should not have loved you so well as I really do love you" (287).

[23] Ruth Perry's examples include Daniel Defoe's *Moll Flanders* (1722),Tobias Smollett's *Roderick Random* (1748), Frances Burney's *Evelina* (1778), and Matthew Lewis's *The Monk* (1796), to name a few. Perry argues that the trope is mainly associated with the sentimental fiction of the second half of the period. See Perry, *Novel Relations: The Transformation of Kinship in English Literature and Culture, 1748-1818* (Cambridge: Cambridge University Press, 2004), 98. Perry's larger argument is that over the course of the eighteenth century, conjugal relations matter more and more and consanguineal ones less and less. My reading of Richardson suggests that in Clarissa's mind (although perhaps not in her brother's), that shift has yet to

irrational nor arbitrary, but rather proof that, as Perry argues, "in a society in which consanguineal ties were increasingly attenuated… blood relations still counted, at least in the imaginative realm of fiction."[24]

Yet Richardson's commitment to the primacy of familial relations is more than a sentimental trope. His careful representation of the Harlowe clan exemplifies David Hume's theories of sympathy and association. Hume sought to dissipate the mystery surrounding the affinity of blood by giving it a philosophical underpinning.[25] In *A Treatise of Human Nature* (1738), he argues that "whoever is united to us by any connexion is always sure of a share of our love, proportion'd to the connexion, without enquiring into his other qualities. Thus the relation of blood produces the strongest tie the mind is capable of in the love of parents to their children, and a lesser degree of the same affection, as the relation lessens. Nor has consanguinity alone this effect, but any other relation without exception."[26] Relations between persons are, according to Hume, a matter of degree and our concern for others is directly dependent on how close a connection we share with them. Because "blood produces the strongest tie the mind is capable of," there is no bond tighter than that of kinship. Thus for Hume, there is no mystery in the fact that Clarissa loves her family despite their avarice and cruelty. The persons with whom we are most intimately connected are our blood relations and our love for them exists independent of any "other qualities" they may have.

---

happen. At the other end of the spectrum, in his recent book, Marshall Sahlins demystifies the blood bond, arguing that kinship is a symbolic (and affective) relation, not a biological one. See Sahlins, *What Kinship Is—And Is Not* (Chicago: Chicago University Press, 2013).

[24] Perry, 98.

[25] Annette Baier notes that Hume had a predecessor in Locke, but points out that Locke did not give consanguinity the special status that Hume did. See Baier, *A Progress of Sentiments*, 48.

[26] Hume, *Treatise*, 352.

Hume explains the priority of family by way of his three principles of association. All persons are more or less related to us depending on their proximity (spatial and temporal contiguity) and the extent to which they resemble us. However, as Annette Baier astutely emphasizes, Hume believes that kinship differs from other kinds of relations in that it carries the additional force of causal connection. Baier observes that in the *Treatise,* the blood relation is the "foundational relation," for "all other relations are 'cousins-german' of 'the relation of blood,' aspects of it, variants of such aspects, or abstract descendants of them. The three Humean 'natural relations' come together in it."[27] The bond between parent and child is stronger than other ties because it combines all three modes of association: contiguity, resemblance, and causation. "Blood produces the strongest tie the mind is capable of" because not only do family members share similar traits and inhabit the same sphere, they also quite literally—that is, biologically—cause each other.[28] Richardson's novel taps into this logic. The Harlowes are, according to Clarissa, a particularly tight-knit clan precisely because the begetting relationship between parent and child is mirrored in other familial ties. As she explains, "never was there a family more united in its different branches than ours. Our uncles consider us as their own children" (56). In the Harlowe family, the uncle-niece (-nephew) bond is reinforced by the special causal relation Hume had reserved for parents and children.[29]

[27] Baier, 48.
[28] Furthermore, Hume reminds us that causation is a stronger association than either resemblance or contiguity: "… there is no relation, which produces a stronger connexion in the fancy, and makes one idea more readily recall another, than the relation of cause and effect betwixt their objects" (Hume, *Treatise,* 11).
[29] Keymer notes that Clarissa's elder siblings take on the parental role: "Clarissa's father['s]… authority over her is often mediated, or perhaps usurped, by his son and elder daughter" (Keymer, *Richardson's Clarissa*, 118).

Framed by Hume's account of the association of persons, Clarissa's "family fondness" makes sense. According to Hume, our love for our family is such that their happiness and welfare is what matters most to us. As he explains, "tho' it be rare to meet with one [man], who loves any single person better than himself; yet 'tis as rare to meet with one, in whom all the kind affections, taken together, do not over-balance all the selfish" (THN, 487). In other words, while human beings are naturally animated by self-interested impulses, the sum total of our loves checks those impulses. That those loves are specifically familial loves is made clear by the example Hume uses to illustrate his point: a father who spends his money on his wife and children rather than on himself.[30] Moreover, Hume immediately goes on to warn that although our generous affections "must be acknowledg'd to the honour of human nature, we may at the same time remark, that so noble an affection, instead of fitting men for large societies, is almost as contrary to them, as the most narrow selfishness."[31] Our kind passions do not extend to the whole of mankind. They are directed towards those near to us, and as such, they are themselves a kind of self-interest.[32] According to Hume then, there is nothing strange in Clarissa's partial love for her family. Our special bonds with our blood relations ensure that we care more for them than the rest of the world—including ourselves.

[30] Hume writes: "Consult common experience: Do you not see, that tho' the whole expence of the family is generally under the direction of the master of it, yet there are few that do not bestow the largest part of their fortunes on the pleasures of their wives, and the education of their children, reserving the smallest portion for their own proper use and entertainment" (*Treatise,* 487).
[31] Hume, *Treatise,* 487.
[32] For an excellent account of the eighteenth-century debate about whether the domestic affections are compatible with universal benevolence, see Evan Radcliffe, "Revolutionary Writing, Moral Philosophy, and Universal Benevolence in the Eighteenth Century." *Journal of the History of Ideas* 54 (1993): 221-240.

Hume's theory of sympathy supports this radical family bias. Explaining the mechanism underlying fellow-feeling, Hume writes that "when any affection is infus'd by sympathy, it is at first known only by its effects, and by those external signs in the countenance and conversation, which convey an idea of it. This idea is presently converted into an impression, and acquires such a degree of force and vivacity, as to become the very passion itself, and produce an equal emotion, as any original affection."[33] In other words, although we initially encounter other people's feelings as mere ideas, sympathy transforms those ideas into impressions, making them come alive, so to speak, as feelings in their own right. Through sympathy, we come to experience the passions of others as if they were our own.[34] If Clarissa would exculpate the Harlowes at any cost, it is in part because she feels their suffering with all its original force. To ask her—as Anna and Lovelace both do—to look upon her family with an objective eye is to ask the impossible. Faced with the unadulterated pain of those she most cares about, Clarissa would do anything to relieve it.

If then Clarissa's attachment to her family seems excessive to us, it is because we fail to understand its Humean logic. Like Anna, we imagine that Clarissa's story might have ended differently if only she had shown a little spirit. Anna describes what could have been: "Your insolent brother, what has *he* to control you? –Were it me (I wish it were for one month, and no more), I'd show him the difference. I'd be in my own mansion…. I'd set up my own chariot. I'd

[33] Hume, *Treatise,* 317.

[34] Adela Pinch points out that Hume sometimes stretches this claim even further. In suggesting that certain passions are more the products of sympathy than of direct experience, Hume opens the door to the possibility that, in Pinch's words, "our feelings are always really someone else's" (*Strange Fits,* 19). I find Pinch's reading of Hume compelling, but I would argue not so much that all passions are vicarious, but rather that vicarious passions can be more powerful than those we experience firsthand. If this is true, one could posit that it is because Clarissa feels her family's suffering more acutely than her own that she would help them even at her own expense. As attractive as it is, I do not see this logic at work in Richardson's novel.

visit them when they deserved it. But when my brother and sister gave themselves airs, I'd let them know that I was their sister, and not their servant" (129). For Anna, sympathy is not an involuntary communication of passion. It is, rather, a form of mental projection whereby one person willfully imagines herself in the place of another ("Were it me," she says). Here Anna is anticipating the model of sympathy Adam Smith would elaborate eleven years later in *The Theory of Moral Sentiments* (1759). Smith's treatise argues that although "the passions upon some occasions, may seem to be transfused from one man to another, instantaneously, and antecedent to any knowledge of what excited them," in fact, the senses "never did, and never can, carry us beyond our own person."[35] Against Hume, Smith explicitly denies that passions are contagious. Other people's passions come to us by way of the imagination, that is, indirectly. Likewise, his theory leaves no room for the possibility that we might care about someone else's emotions more than our own. We can never experience the passions of others as intensely as those we experience firsthand, and as a result, we will always be more concerned about the latter than the former. What this means is that so long as Anna thinks of sympathy in Smith's terms, Clarissa's partial love for the Harlowes will remain unintelligible to her. It is only by recognizing that passions are, as Hume suggests, transpersonal—infectious—that we can begin to understand why Clarissa cares more about her family's suffering than her own.

What all this means is that Lovelace is not quite correct when he complains that Clarissa is "above flattery, and despises all praise but that which flows from the approbation of her own heart" (423). Clarissa is above *his* flattery, but her heart is by no means as self-sufficient as he imagines.[36] By her own account, much of her sense of self-worth stems from her sympathetic

[35] Smith, *The Theory of Moral Sentiments,* 11; 9.
[36] Nazar argues that the novel comes out in favor of individual self-judgment as opposed to peer judgment. Clarissa's decision not to marry Lovelace after the rape, despite the urgings of Anna,

connection with her kin. She confesses to Anna that until lately, "I was the pride of all my friends, proud *myself* of *their* pride, and glorying in my standing" (333). Clarissa possesses the family fault—which is, of course, what the Harlowes have been arguing all along—indeed, she cannot avoid it.[37] Their pride produces hers, mirroring the causal association inherent to the blood relation.[38] Thus although Clarissa is (most of the time) indifferent to the praise of the world, she is not—and cannot be—insensible to the regard of her relations. She takes special delight in their esteem, and likewise, it is their contempt that hurts her most. Sentiments are a family affair, and being a Harlowe means, at some very basic level, putting the passions of other Harlowes first.

I have been arguing that Hume's particular account of the association of persons provides an explanation for how punishment comes to be selectively transpersonal. There is, however, a discrepancy between Clarissa's sense that she has been infected with the Harlowe pride and her later claim that she has been "singled out to be the *punisher*" of her family's faults. The problem is that under Clarissa's account, the Harlowe fault is a passion. What this means is that in sharing her family's feelings, Clarissa takes on their fault: this is guilt by association, not punishment by

---

her family, Lovelace's relations, and Lovelace himself, shows that "ultimately, it is one's own eyes, rather than the world's eye or even the friend's eyes that matter in moral discrimination" ("Judging Clarissa's Heart," 96). It seems important, however, that Clarissa has little trouble convincing Anna and the Montagues that refusing Lovelace is the right thing for her to do. Moreover, Clarissa does not believe that she needs to choose between satisfying her friends and satisfying her conscience. She is sure that so long as she supports her character, she will regain her family's good opinion (albeit after her death).

[37] The Harlowes trace Clarissa's disobedience back to her inflated sense of self-worth. They are not completely off the mark. Her rejection of Solmes smacks of vanity: "I to have that Mr. Solmes!—No indeed!" (104). Hence her mother's constant retort, "No more of your merits, Clary!" (94). See Warner for a reading of Clarissa's merciless caricature of Solmes in *Reading Clarissa.*

[38] See Pinch for a compelling discussion of how in Hume, the self is produced by pride rather than the other way around ("The Philosopher as Man of Feeling: Hume's Book of the Passions" in *Strange Fits*, 17-51).

association. Thus a strict Humean reading of Richardson's novel might go something like this: Clarissa is connected to the Harlowes through ties of blood and, as a result of the strong sympathy between them, she takes on their pride and makes it her own. If this is right, then there is no need to look very far to explain the novel's ending. There is nothing so troubling about Clarissa's death, for she *deserves* punishment. Clarissa's pride belongs to her—its provenance is beside the point—and she is justly repaid for her "haughty airs" and her foolish self-reliance. According to this view Clarissa is punished wholly on her own account; she is not, as she elsewhere imagines, the unlucky instrument of Heaven's justice.

Yet the novel takes great pains to avoid this reading. To label Clarissa guilty by association would be to dismiss what is, for many readers, the central problem of the novel: just how faulty is Clarissa? Clarissa's pride is a case in point. Richardson repeatedly reminds us that her pride is different from that of the Harlowes. As Anna explains, "some superficial persons… impute[d] pride to [Clarissa]. But she knew not what pride, in the bad sense of the word, was" (1466). Clarissa's pride is a "laudable pride," a "native dignity" "—just another virtue that sets her apart from the other characters. In short, her pride is not a blamable pride. Furthermore, although Clarissa condemns her "*secret* pride" more loudly than anyone, she minimizes this fault by focusing on what looks like a greater one. For her, the source of her woes lies more in her disobedience than her pride. Shortly after her elopement, she writes: "I ought not to have corresponded with him…. —This last evil, although the remote, yet sure consequence of my first—my prohibited correspondence! By a father, at least, prohibited!... I suppose it concerned me more than any other to be the arbitress of the quarrels of unruly spirits—and now I find my presumption punished!" (381). This is Clarissa's casuistry at its best. While admitting to an ill-conceived (though well-intended) "presumption," this confession is overshadowed by her

insistence that it was the exchange of letters with Lovelace that has brought her to this point. It is this claim that quietly lets Clarissa off the hook. "By a father, at least, prohibited!" might look like a rhetorical flourish, but in fact, Clarissa is reminding us that Mrs. Harlowe expressly encouraged the correspondence proscribed by Mr. Harlowe Sr. This contradiction means that in disobeying her father, Clarissa was simultaneously obeying her mother. In other words, what looks like an admission of error is in fact self-vindication. All this to say what readers have been saying in one way or another from the start: it is virtually impossible to ascertain (the extent of) Clarissa's guilt.

The question of who is responsible for Clarissa's bad actions—refusing Solmes, corresponding with Anna and Lovelace, eloping with a libertine—is, as critics have frequently noted, far from clear. More often than not, the question of Clarissa's guilt hangs in the balance as she wavers between reproaching her family and/or Lovelace and denouncing herself. Part of the difficulty of judging Clarissa lies in the fact that she seems to be all too willing to take responsibility for the actions of others, a tendency that has frustrated her defenders both within the novel and without.[39] As commentators have observed, however, it is a mistake to interpret this impulse as mere masochism.[40] Macpherson lays bare the subtext of such a reading:

> At stake in the vocabulary of masochism and the argument aligning Clarissa's abjection with Richardson's sadistic subjection of her are two related assumptions: that Clarissa blames herself in order to harm herself and that to ally innocence and culpability is at best

[39]Most famously, Clarissa takes responsibility for her rape, excusing Lovelace on the basis that he was only acting according to his nature. See the parable of the lady and the bear in the third Mad Paper (891).

[40] For discussions of masochism in the novel see for example Laura Hinton, "The Heroine's Subjection: Clarissa, Sadomasochism, and the Natural Law." *Eighteenth-Century Studies* 32 (1999): 293-308, and Judith Wilt, "He Could Go No Farther: A Modest Proposal about Lovelace and Clarissa." *PMLA* 92 (1977): 19-32.

> mad, at worst perversely inhumane. To think so, however, is to think like Lovelace—to cultivate the obverse idea that blamable persons are not innocent, and if they are not innocent they have not been harmed.[41]

Macpherson's point is that it is only by recognizing that Clarissa can be both sinning and sinned against, so to speak, that we can insist upon Lovelace's guilt. To suppose that she is either a blameless victim or a culpable accomplice is to absolve her rapist, for the novel consistently denies us the proof we need to clear Clarissa of any wrongdoing. As Macpherson shows, Clarissa undermines this dichotomy by insisting on her faultless intentions while at the same time taking responsibility for the unintended outcomes of her actions. Thus although we might find the heroine's urge to find fault with herself unsettling, that urge grows out of a commitment to strict liability, that is, to the only logic of accountability according to which Lovelace is unequivocally guilty.

While I share Macpherson's sense that it is a mistake to read Clarissa's tendency to find fault with herself as a form of self-harm, I see this tendency not as evidence of her adherence to the logic of strict liability, but rather as the necessary consequence of partial love. Clarissa takes responsibility for the injuries she suffers not in order to punish herself but, rather, to absolve the persons who inflict those harms. After running away with Lovelace, for example, Clarissa tells Anna that: "… now I have this *one* consolation left me (a very sad one, you'll say), that I have cleared *them [my family]* of blame, and taken it all upon *myself!*" (382). Clarissa understands that the more innocent she is, the more her family is guilty. And as her situation becomes more and more desperate, she realizes that there is no question of reconciling the two such that they may all be innocent. If she appropriates their actions, therefore, it is because incriminating herself

[41] Macpherson, *Harm's Way,* 91.

means excusing them. Clarissa is not interested, at least where the Harlowes are concerned, in the kind of justice that grounds itself in the causal relationship between agent and action. She is not interested, in other words, in the kind of justice that would cause her family pain, even where it is warranted. This is an unapologetic commitment to partial love that flies in the face of the ideal of impartiality that Clarissa espouses elsewhere. As such, it asks us to think about what justice might look like when familial sympathy matters more than either individual agency or material causation.

At times, Clarissa does seem to invoke the logic of material causation and strict liability. In her attempt to clear her family of blame, Clarissa often represents herself as the author of the "first" evil in a chain of moral errors. She draws a causal connection between her actions and those of her family in order to identify herself as the source of their faults. However, this tactic is only half successful. The trouble is that as the Harlowes' actions become more and more blamable, Clarissa's first fault seems less and less relevant. Clarissa recognizes this. Upon learning that Anna has heard something so terrible that it cannot bear repeating ("bad as what I have mentioned is, I have not told you all; nor now, perhaps ever will" [587]), Clarissa exclaims: "Surely my Father has not renewed his dreadful malediction! –Surely, if so, my Mother has not joined in it! Have my Uncles given it their sanction, and made it a family act? *And themselves thereby more really faulty, than even* THEY *suppose me to be, tho' I the cause of the greater fault in them?*"[42] Clarissa is horrified by Anna's intimation, worried that her father's awful curse

[42] Samuel Richardson, *Clarissa, Or, the History of a Young Lady,* 3rd ed., 8 vols. (London: Printed for S. Richardson, 1751) 4: 85. The last sentence of this quotation ("And themselves thereby more really faulty…") does not appear in the first edition of the novel. When Richardson published the third edition in 1751, he promoted it as a more complete and authentic version of the text: "It is proper to observe, with regard to the *present Edition,* that it has been thought fit to restore many Passages, and several Letters, which were omitted in the former merely for shortening-sake" (preface to *Clarissa*, 3rd ed., 1: ix). Regardless of whether this line is a

has been repeated, this time as a chorus of condemnation. In her worst imaginings, her relations have come together to reiterate her father's curse, united in wishing her harm in this life as well as the next. Clarissa seems to think that a group action outweighs an individual one and that a curse uttered against her by her whole family is an even greater punishment than her father's alone. This is more than she deserves, and importantly, Clarissa's fear is not all for herself. She believes that were they to take their wrath too far, the Harlowes would be committing a fault even greater than hers.

Unwilling to assume blame for a crime so great, but still intent on defending her family, Clarissa searches for a solution. On the one hand, Clarissa's claim that she is "the cause of the greater fault in them" argues for a continuity between her actions and those of the Harlowes. It locates the first fault in her, making her responsible for the "greater fault" that follows. Clarissa thereby lays claim to the last evil as well as the first—and everything in between. On the other hand, she draws a clear distinction between herself and her relations in this passage. She portrays herself as the (albeit not wholly innocent) victim of a hypothetical family act and insists that her guilt is quantitatively different from theirs. Clarissa may be the cause of their acting badly, but they remain "more really faulty" than her. Pronouncing herself the cause of her family's "greater fault" shifts the brunt of the blame onto Clarissa, but it does not entirely exonerate the Harlowes. To achieve this, Clarissa must go further. She seeks to appropriate their actions in order to

---

restoration of unpublished manuscript material or a new addition, it is a helpful articulation of Clarissa's logic in this scene. For discussions of *Clarissa*'s textual history, see Fulton, *Styles of Meaning;* Shirley van Marter, "Richardson's Revisions of *Clarissa* in the Third and Fourth Editions." *Studies in Bibliography* 26 (1973): 107-32; Mark Kinkead-Weekes, "*Clarissa* Restored?" *Review of English Studies* 10 (1959): 156-71. See also Toni Bowers and John Richetti's introduction to Broadview's abridged text, which is based on the third edition (Bowers and Richetti, introduction to *Clarissa Or, The History of a Young Lady*, by Samuel Richardson [Peterborough: Broadview Press, 2011]).

completely eliminate the Harlowes from the equation. Prior to their appointment at the garden gate, Lovelace looks to overcome Clarissa's qualms by assuring her that people expect her to elope with him. Public opinion, he tells her, is all on her side, and if she were to flee her father's house, it would be seen as nothing more than justice. None of this is new (Anna Howe has already said something to the same effect) but Clarissa's reaction is rather curious. Convinced that Lovelace is too polite to "say all he might on the subject," Clarissa reflects:

> I have no doubt that I am the talk, and perhaps the byword of half the county. If so, I am afraid that I can now do nothing that will give me more disgrace than I have already so causelessly received by their indiscreet persecutions: and let me be whose I will, and do what I will, I shall never wipe off the stain my confinement and the rigorous usage I have received have fixed upon me; at least in my own opinion. (349)

Lovelace justifies Clarissa's escape by condemning her family for their harsh usage, but Clarissa justifies it by condemning herself; she is sunk so low that whatever she does now will have little effect one way or another. She reads into Lovelace's words ("as he hints," she explains), twisting them such that it is she who is disgraced by the persecutions she has "causelessly received." The confinement and "rigorous usage" Clarissa suffers fastens to and reflects upon her instead of the Harlowes. The Harlowes confine Clarissa but, in (only) her eyes, the action of confining fixes upon her instead of them and she is the one who bears its guilty stain.

Clarissa is universally lauded for a wisdom beyond her years, but here she seems to get it wrong. We can see how she falls into error by tracing the trajectory of her thought. Imagining the county gossip, she complains that her relations have ruined her reputation through their "indiscreet persecutions." As long as Clarissa is looking at her situation from a third-person point

of view, she sees herself as the guiltless victim of her family's violence. Ventriloquizing Lovelace, she announces that:

> 'But that, nevertheless, as to the world's opinion, it is impossible to imagine that the behaviour of my relations to me has not already brought upon my family those free censures which they deserve, and caused the step which I am so scrupulous about taking to be no other than the natural and expected consequence of their treatment of me.'
>
> Indeed, I am afraid all this is true… (349)

Importantly, Clarissa is both more and less than a mouthpiece here. This moment stands out insofar as it is the *only* sentence in Clarissa's recitation of Lovelace's long letter that is not punctuated by "he" or "his." Lovelace's constant presence is briefly suppressed as Clarissa inhabits a third-person perspective, "the world's" point of view. Seeing her family's actions as others do, she can conclude of Lovelace's claims that "all this is true": the Harlowes are at fault and she is innocent. This passage is the last extract from Lovelace's letter and quotation marks clearly signal the switch. Over the course of the next paragraph, Clarissa moves from a proclamation of innocence (she begins, "all this is true") to complete self-abasement; the paragraph ends: "I shall never wipe off the stain my confinement and the rigorous usage I have received have fixed upon me; at least in my own opinion." It seems that thoughts about the ill-treatment she has suffered inevitably lead Clarissa from the world's opinion to her own, a shift in perspective that results in a reallocation of guilt. Lovelace, Anna, and the world all blame the Harlowes for their mistreatment of Clarissa, but Clarissa blames herself. So when Clarissa adopts the world's point of view, she is the object and recipient of the Harlowes' cruelty, but when she looks at things through her own eyes ("at least in my own opinion"), the bad actions of the Harlowes belong to her.

Clarissa thinks the Harlowes' actions fix upon her, as if she were their cause. If this seems back to front, it is only because we, like Anna and Lovelace, are looking from the outside in. Clarissa's bid to take responsibility for the cruel persecutions of her relations is, at bottom, an attempt to mitigate her own suffering. Nothing hurts Clarissa more than the thought that she has made her family the object of public censure. When Anna speaks against the Harlowes, Clarissa not only defends them, she blames herself for having given her friend fodder for criticism. She writes:

> I am very angry with you for your reflections on my relations, particularly on my father…. One cannot one's *self* forbear to write or speak freely of those we love and honour; that is to say, when grief wrings the heart. But it goes against one to hear anybody else take the same liberties…. when passion has subsided and I come by reflection to see by *your* severity what I have given occasion for I cannot help condemning myself. (134)

Clarissa understands herself to be the "occasion" for making the Harlowes blameworthy in the eye of the world. If Clarissa's guilt automatically implies her family's innocence, it is likewise true that her innocence implies their guilt. As a result, Clarissa cannot plead her case without hurting the Harlowes. That Clarissa ultimately condemns herself for defending herself is a testament to her partial love.

Clarissa's attempts to spare the Harlowes pain cannot but fail. Readers have tended to denounce the Harlowes' avarice, yet it is not true that Clarissa is merely a bargaining chip in their quest for social status (though she is certainly partially that).[43] The Harlowes are, as

[43] A notable exception to this trend in the criticism is John Allen Stevenson's claim that "the Harlowes are only using what appears to be a property marriage to disguise their true aim, which is to escape society's powerful imperative that parents give up their daughters in marital

Lovelace frequently sneers, first and foremost a proud family, and the source of that pride is Clarissa. To lose Clarissa is to lose that very thing that makes the Harlowes who they are. As Mrs. Harlowe explains, "The dear creature, you know, my Norton, gave an eminence to us all. And now that she has left us, so disgracefully left us! we are stripped of our ornament, and are but a common family!" (584). Clarissa's fall deprives her family of that "ornament" which makes them different to and better than the rest of the world. Without Clarissa, the Harlowes are "but a common family" and they are forced to forfeit the notorious pride that once distinguished them. What all this means is that although Clarissa would readily save her family's reputation at the cost of her own, this will do nothing to relieve their suffering. Clarissa is indissolubly attached to the Harlowes and her disgrace and suffering means theirs as well. The Harlowes are trapped. Either Clarissa or her kin is guilty, but what the novel finally shows is that in the end, it does not really matter which one of them is to blame. To punish one Harlowe is to punish them all. *Clarissa* thus reveals the seamy-side of fellow-feeling: our sympathetic associations—inescapable and unchosen—mean that we must suffer for the faults of our loved ones and they in turn must suffer for ours. The association of persons is a painful business.

II

*Clarissa'*s "knowledge… of the human heart" culminates in this insight: just as the heart is the seat of our private conscience, it is also that which binds us to the people we love. If the novel's ending feels unsatisfactory to us, it is because we are imagining that we want a justice that is both impartial and individual. But Richardson reminds us that we would willingly forego

exchange" ("The Courtship of the Family: Clarissa and the Harlowes Once More," *ELH* 48 [1981]: 760).

this ideal in order to spare our loved ones pain. Thus in the wake of her father's curse, Clarissa reflects:

> … whatever shall be my destiny, that dreadful part of my father's malediction, that I may be punished by the man in whom he supposes I put my confidence may not take place! That this for *Mr. Lovelace's* own sake, and for the sake of *human nature*, may not be!... Otherwise, my fault will appear to be doubled in the eye of the event-judging world. And yet, methinks, I would be glad that the unkindness of my father and uncles, whose hearts have already been too much wounded by my error, may be justified in every article, excepting in this heavy curse… at least that most dreadful part of it which regards futurity! (566)

Here Clarissa is not far off from Hume's famous claim that "it is not contrary to reason to prefer the destruction of the whole world to the scratching of my finger."[44] Crucially, however, it is not her own welfare that most concerns her, but her family's. Clarissa wants the Harlowes to be right in "the eye of the event-judging world" in order to lessen their suffering—at her expense and even at the expense of "human nature." This prejudice makes a case both for Hume's theory of the association of persons and for the sympathetic mechanism that ensures that, finally, we are never really on our own. Persons are connected to other persons and, in consequence, they share each other's pain. The logical consequence of such sympathy is then that suffering—and, by extension, punishment—is transpersonal. One cannot punish a guilty person without harming her friends.

Clarissa shows Richardson working through the classical notion of punishment by association in order to understand what we, given our "partial loves," would consider justice.

[44] Hume, *Treatise,* 416.

That justice turns out to be Clarissa's death is indicative of just how strong our personal associations are. The novel depicts familial bonds as indissoluble and although Clarissa may leave home, she cannot leave her family. Clarissa's dilemma is that she cannot accept Solmes but, so long as she refuses him, she cannot protect the Harlowes from censure. There seems to be no way out of this problem. As Clarissa laments, "What a world is this! What is there in it desirable? The good we hope for, so strangely mixed, that one knows not what to wish for: and one half of mankind tormenting the other, and being tormented themselves in tormenting!—For here in this my particular case, my relations cannot be happy, though they make me unhappy!" (224). What Clarissa comes to hope for, of course, is that her "closing scene" may be happy. Only death will sever the ties between her family and herself, making her pain—and happiness—all her own. This vision of justice is anything but, to use Richardson's word, "*poetical*." What *Clarissa* shows us is that because suffering is always shared, there is no such thing as fit punishment. There is no such thing, in other words, as "an eye for an eye, a tooth for a tooth." Retribution does not negate harm; it spreads it.

Chapter Three

Deviant Connections in Choderlos de Laclos's *Les Liaisons dangereuses*

J'ai vu les mœurs de ce siècle et j'ai publié ces lettres.

—Jean-Jacques Rousseau, Préface de *Julie, ou La Nouvelle Héloise*

In Hume's posthumously published *Dialogues Concerning Natural Religion* (1779), Cleanthes claims that it makes no sense to look for the first cause in a causal sequence:

> In such a chain, too, or succession of objects, each part is caused by that which preceded it, and causes that which succeeds it. Where then is the difficulty? But the WHOLE, you say, wants a cause. I answer, that the uniting of these parts into a whole, like the uniting of several distinct counties into one kingdom, is performed merely by an arbitrary act of the mind…. Did I show you the particular causes of each individual in a collection of twenty particles of matter, I should think it very unreasonable, should you afterwards ask me, what was the cause of the whole twenty.[1]

A whole is not more than the sum of its parts, says Cleanthes, and there is no need to investigate its cause for it is we, the observers, who put that whole together through an "arbitrary act of the mind." Although Cleanthes is not thinking about character in this passage, his description of a "whole" matches Hume's definition of "whole character" from the *Treatise*: a whole made up of characters and causes—simply put, a collection of causes. The two accounts differ, however, in that while Cleanthes seeks to *explain* the relationship between the whole and its parts—the whole

[1] David Hume, *Dialogues Concerning Natural Religion: And Other Writings* (Cambridge: Cambridge University Press, 2007), 65-6.

is a haphazard chimera produced by the mind—in the *Treatise*, Hume merely describes it. Hume avoids, in other words, the question of how whole character relates to its parts (causes).

In some ways, Cleathes's assertion that a causal chain is merely a collection of contiguous causes is even more counterintuitive, perhaps, than Hume's famous claim that self is a fiction. After all, it seems obvious that a causal chain is not the same thing as a group of particles, and the first cause of the former is not the same as the overarching cause of the latter. Yet the distinction between a temporal collection of causes and a spatial one does not really affect Cleanthes's larger argument, which is simply that both these entities stem from an arbitrary act of the mind. Indeed, if Cleanthes's analogy feels strained, it is more because it paradoxically insists that causation and arbitrariness are both inherent to causal chains. Cleanthes, more concerned with the question of a First Cause than with the mechanism of causal chains, does not explore this tension, and Hume does not take it up elsewhere. However the juxtaposition of these contradictory terms—"cause" as against "arbitrary"— is suggestive, to say the least.

I begin this chapter by considering the possibility that, borrowing Cleanthes's logic, the boundary that demarcates a person's "WHOLE" character—the sum of those traits that motivate her behaviour—is at best ambiguous and at worst, completely arbitrary. In the previous chapter, I argued that Richardson's *Clarissa* challenges our intuitive understanding of moral responsibility by showing how sympathy can connect persons to actions that are not properly their own. Clarissa is punished for the faults of her family not because she is somehow complicit in them, but because she is affectively attached to her kin, and as a result, their punishment becomes hers as well. Similarly, Pierre-Ambroise Choderlos de Laclos's *Les Liaisons dangereuses* (1782) destabilizes the notion of moral accountability by showing how causal chains can complicate the

relationship between character and action. In the novel, actions and their outcomes are rarely a direct consequence of someone's character. Rather, characters realize their intentions by way of causal chains, wherein external obstacles and chance events influence how their actions unfold without throwing them off course. The argument of this chapter is that in *Les Liaisons dangereuses*, causal chains can loosen necessary connection without breaking it, thereby blurring the line between action and event, character and world, without losing the distinction all together. That is, *Les Liaisons dangereuses* thinks about the question that Hume's theory of necessary connection implies without explicitly raising: what if *how* one does something is just as important as *what* one does? And what if one cannot fully control how one does something? *Les Liaisons dangereuses* provides a case study for thinking about how causal chains might enable chance to operate within necessary connection and what that might mean in terms of moral accountability.

I

It is a striking feature of Laclos's *Les Liaisons dangereuses* that no one ever seems to perform an action wholly contrary to his desires or intentions. Ingenious plotting only accomplishes what was, on some level, willingly anticipated. The notorious rake Prévan succeeds in separating the inséparables only because the three friends are already bored to death with one another's company. Valmont gains access to Cécile because she wants to be able to engage in amorous exchange.[2] And as Jean-Luc Seylaz points out, "La tricherie de Laclos consiste donc ici à nous faire oublier que la victoire de Valmont était impossible sans la complicité de Mme de Tourvel,

[2] Cécile gives Valmont the key to her room. As Catriona Seth notes in the introduction to the 2011 Pléiade edition, in the jargon of the day, "key" and "lock" were code for the male and female genitals.

c'est-à-dire sa passion."[3] Thus when Mme de Volanges laments, in the final lines of the novel, that: "j'éprouve en ce moment que notre raison, déjà si insuffisante pour prévenir nos malheurs, l'est encore davantage pour nous en consoler," she is not claiming that "our misfortunes" are unforeseeable.[4] The distinction between prévoir [to foresee] and prévenir [to warn] matters, for in the world of *Les Liaisons dangereuses,* everything is predictable; the fictional Publisher presses this message home by giving away the novel's ending (Tourvel's death and Cécile's retreat to the convent) in his preface. But the Publisher's spoiler (and this holds true for the logical and hence unsurprising effects in the diegesis, too) spoils very little, and this is precisely Laclos's point. The novel's moral universe is, like that of Hume's *Treatise,* Newtonian, which means that human motive and action operate like cause and effect in the natural world. In presenting a world where every action is both necessary and motivated, *Les Liaisons dangereuses* positions itself within a Humean framework of necessary connection—every action presupposes its motive, which is to say that there is no such thing as unintentional action.

Laclos's novel asks us to consider where accident might fit within that framework. The text engages with the Humean theory of causation in order to look at how the meaning of an action might not necessarily be exhausted by the intention that motivates it. Hume, as I have discussed earlier, explained that what we call causal connection is merely a connection that we

[3] "Laclos's art consists in making us forget that the victory of Valmont would have been impossible without the complicity of Mme de Tourvel, which is to say her passion." Translation mine. See Jean-Luc Seylaz's seminal work, *Les Liaisons dangereuses et la création romanesque chez Laclos* (Paris: Minard, 1958), 143.

[4] Pierre-Ambroise Choderlos de Laclos, *Les Liaisons dangereuses,* ed. Catriona Seth (Paris: Gallimard, 2011), 459. Hereafter cited parenthetically in the text as L. "I feel at this moment that our reason, so incapable of [warning us of] our misfortunes, is still less capable of consoling us for them." Laclos, *Les Liaisons dangereuses,* trans. Richard Aldington (London: Routledge, 2010), 407. Hereafter cited parenthetically in the text as A. All English translations are taken from this text unless otherwise noted. My changes are indicated in square brackets.

have seen with such regularity that we take it to be a rule. A causal process, by extension, is simply a chain of constant conjunctions. What *Les Liaisons dangereuses* makes clear is that causal chains leave room for intermediary accidents because the same cause and effect might be linked in an infinite number of different ways, each of which will produce a different meaning. The point here is that there is no such thing as a "safe" connection. The danger of liaisons is that what looks like a simple cause-effect relation might in fact be a causal chain, which opens up the possibility for multiple incidental causes. This is a moral problem. In complicating the Humean concept of causation, Laclos's novel raises questions about the relationship between accident and moral accountability. Am I responsible when an action I perform takes an unexpected turn, accruing meaning along the way before it finally achieves my intended effect?

That the question of causation is at the heart of the novel is clear from the first. After all, the very authenticity of Laclos's text hangs upon a causal connection. In his preface, Laclos's fictional Publisher declares the letters decisively false on the basis that:

> Sans doute les mêmes causes ne manqueraient pas de produire les mêmes effets, et que cependant nous ne voyons point aujourd'hui de Demoiselle avec soixante mille livres de rente, se faire Religieuse, ni de Présidente jeune et jolie, mourir de chagrin. (L, 14)[5]

The logic of the fiction is good, he admits, but that it *is* fiction—at least in terms of setting—is borne out by the fact that in this day and age, we never see such sad cases (a rich young heiress taking the veil or a pretty young wife dying of grief). The Publisher's "proof" implies that from the fact that there are no Céciles or Mme de Tourvels, one can deduce that there are no Valmonts or Merteuils. But Christine Roulston puts forward an alternative interpretation, suggesting that

[5] "It is that the same causes could not fail to produce the same effects, and that today we never see a young Lady, with an income of sixty thousand livres, become a nun, nor the wife of a Président die of grief while she is still young and pretty" (A, 4).

"the improbability of the novel resides in its idealized representation of virtue, not its immoral articulation of vice."[6] Accordingly, Roulston argues that the Publisher's statement functions as social critique. It implies that it is enlightenment morality that falls short of the models portrayed in the letters and not the other way around. But the bigger problem, of course, is that each effect the Publisher singles out has more than one possible cause. The terrible fates of Cécile and Mme de Tourvel might equally be a consequence either of extreme depravity or of particular virtue, so that likewise, it is possible that today's ladies are living in an enlightened society free of vice, but just as likely that they are simply too wicked to want to consign themselves to the nunnery or the sickbed. In other words, there is no way of knowing in what sense the social mores depicted in the narrative are out of proportion to those of the framing text—whether the ladies are excessively good or their tormentors especially bad. Instead of establishing the relative difference between the moral values contained in the letters and those of Enlightenment society, the Publisher ends by raising questions about the means he uses to determine what those moral values are. Thus the Publisher's conclusion is no conclusion at all—and this is exactly the point. His logic is good—following Hume, he infers cause from effect. But a "too voluminous" narrative—whether true or false—cannot be explained in terms of a single causal connection.

The Publisher's Note, insofar as it seeks to establish a relationship between the letters and the extradiegetic world, sounds like an echo of the novel's epigraph, a quotation taken from Jean-Jacques Rousseau's first preface to *Julie, ou La Nouvelle Héloise* (1761): "J'ai vu les mœurs de mon temps, et j'ai publié ces lettres."[7] However whereas the Publisher's preface challenges the veracity of the letters based on anachronism, the epigraph seems to leave the truth

[6] Christine Roulston, *Virtue, Gender, and the Authentic Self in Eighteenth-Century Fiction: Richardson, Rousseau, and Laclos* (Miami: University Press of Florida, 1998), 146.
[7] "I have seen the morals of my time, and I have published these letters."

status of the letters ambiguous while framing them as a response to current social conditions. An overlap worth exploring, though, is the repetition of the word "siècle" in the preface, which draws attention to a small change Laclos made to Rousseau's text. In the manuscript, Laclos inserts "ce siècle" in the place of "mon temps" so that the epigraph in fact reads: "J'ai vu les mœurs de ce siècle et j'ai publié ces lettres."[8] Critics have not made much of the revision. Laurent Versini's explanation in the 1979 Pléiade is that Laclos was quoting Rousseau from memory and later corrected his mistake. Since then, critics have been content to read the epigraph almost exclusively in dialogue with Rousseau.[9] But the correction is not in the manuscript and there is no way of knowing when—or who—made the correction.[10] Given that

[8] "I have seen the morals of this century, and I have published these letters."

[9] See Laclos, *Œuvres complètes*, ed. Laurent Versini (Paris: Gallimard, 1979), 1163. Thomas Kavanaugh and Susan Winnett both note the change but do not comment upon it. See Kavanaugh, *Enlightened Pleasures: Eighteenth-Century France and the New Epicureanism* (New Haven: Yale University Press, 2010), 233; Winnett, *Terrible Sociability: the Text of Manners in Laclos, Goethe, and James* (Stanford: Stanford University Press, 1993), 45. Versini, Peter Brooks, and Ronald Rosbottom all give interesting readings of the epigraph. For Versini, the epigraph identifies Laclos's philosophical position in relation to Rousseau's. Versini argues that Laclos agrees with Rousseau that society is the source of evil, but that he believes, contra Rousseau, that man is naturally sociable. Brooks argues that "By this epigraph, Rousseau meant to suggest that his novel constituted a lesson to his contemporaries; Laclos uses it to imply that his novel is an exemplum of contemporary behavior," (Brooks, *The Novel of Worldliness,* 212). Rosbottom sees the epigraph as Laclos's way of establishing an ironic juxtaposition between the two texts. See Laurent Versini, *Le Roman le plus intelligent: Les Liaisons dangereuses de Laclos* (Paris: H. Champion, 1998); Peter Brooks, *The Novel of Worldliness: Crébillon, Marivaux, Laclos, Stendhal* (Princeton: Princeton University Press, 1969); Robert Rosbottom, "Dangerous Connections," in *Laclos: Critical Approaches to Les Liaisons dangereuses,* ed. Lloyd R. Free, (Madrid: José Porrúa Turanzas, 1978), 183-221.

[10] Laclos, Pierre Choderlos de. *Les Liaisons dangereuses.* Manuscript c. 1779-1781. Bibliothèque Nationale de France, Département des Manuscrits, Doc. 12845, fol. 35r., accessed October 24, 2016, ark:/12148/btv1b60002397. The title page shows the alteration to the title (*Les Dangers des liaisons*, originally), but the epigraph remains untouched. Dorothy Thelander argues that the BNF's MS was followed by a later draft (now lost) that served as the printer's copy. As she admits, however, there is no way of proving that the BNF MS is not the printer's copy. It has been argued that the last note by the Publisher, which leaves room for a sequel, is in fact a note from Durand Neveu, Laclos's publisher. The lines are not in Laclos's hand. See Thelander, *Laclos and the Epistolary Novel* (Genève: Droz, 1963).

the Publisher's note consistently gives preference to "siècle" over "époque" or "temps" in order to emphasize the fact that "notre siècle" is the "siècle de philosophie, ou de lumières," the change is worth pausing over. [11]

The word "siècle" invites the reader to break from the Rousseauist perspective and give scope to a larger context. One can imagine a worse place to start than with the epigraph. How might one read the epigraph without Rousseau? Perhaps Laclos's statement does not mimic the logic of Rousseau's—perhaps it is not causal. Laclos does not publish the letters *because* he has seen the manners of his time; he sees them *and* he publishes them. Insofar as the epigraph sets the tone for what is to follow then, *Les Liaisons dangereuses* is neither didactic nor exemplary. It is nothing more than a collection of observations, observations from an outsider's perspective: it is "ce siècle" (this century), not "mon siècle" (my century). That is, the epigraph is not a declaration of the text's orientation relative to Rousseauist ideology, as critics have suggested, but rather it signals an affirmation of Enlightenment values, specifically Newtonian science and Humean empiricism.[12] As the fictional Editor says, "Plusieurs personnes pourront compter encore pour quelque chose un assez grand nombre d'observations, ou nouvelles, ou peu connues, et qui se trouvent éparses dans ces lettres" (L, 9).[13] What the letters offer, despite all their faults, are a great number of observations, and this is, in the end, all they offer. Although the novel's introductory materials (the epigraph, Editor's Preface, and Publisher's Note) have often been

[11] "Époque" and "temps" each appear just once in the Publisher's preface, both in reference to the "fiction" rather than to the Publisher's enlightened, philosophical age.
[12] Versini points out that Laclos was a man of his times—apart from his frequent references to Rousseau, other citations show that he read Helvétius and that he owned almost all of Voltaire. Versini convincingly shows the influence of Locke, Condillac, La Mettrie, and d'Holbach. See Versini, *Le Roman le plus intelligent,* pp. 137-45; *Laclos et la tradition, essai sur les sources et la technique des Liaisons dangereuses* (Paris : Klincksieck, 1968).
[13] "Some persons may count as a merit the considerable number of observations, either new or little known, which are scattered through these letters" (A, 7).

viewed as being in tension with one another, there is a sense in which they are very much in agreement, at least with regard to how the novel should be read.[14] The Publisher's causal realism and the emphasis the epigraph and the Editor both place on observation all meet in Hume's empiricism. All three framing texts read action and character from the outside. The letters are descriptions that "feign no hypothesis": they reveal *how* immoral people corrupt others—but not why.

II

Critics have not, however, been content to stop at the "how." Laclos's novel is widely acknowledged to be the stylistic zenith of the epistolary genre, and comparisons with Samuel Richardson and Rousseau inevitably generate interpretations founded in psychological realism. Lacanian theorist Slavoj Žižek has recently pointed to Valmont as a classic case of the paradox of the superego (the more he compromises his desire—by falling in love, by renouncing Mme de Tourvel—the greater is his guilt).[15] Even those readings that dismiss "minor" characters like Cécile and Danceny as flat cannot resist treating the libertine protagonists as "round characters," an example of the asymmetry that Alex Woloch has described as the one vs. the many.[16] But as

[14] Henri Duranton, for example, argues that the two prefaces cancel each other out. David McCallam argues that the publisher undermines the editor, but that the final publisher's note (which hints at a sequel depending on the reception of the letters) hands over his authority to the public. See Henri Duranton, **"***Les Liaisons dangereuses* ou le miroir ennemi," *Revue des sciences humaines* 153 (1974) : 125-43; and David McCallam, *L'art de l'équivoque chez Laclos* (Genève: Droz, 2008).

[15] Slavoj Zizek, *Metastases of Enjoyment: Six Essays on Women and Causality* (London: Verso, 1994), 68.

[16] Thus Charles Baudelaire dismisses the correspondents one by one—Cécile is a "type parfait de la detestable jeune fille," Mme de Tourvel is an "Ève touchante," Danceny an "Homme d'honneur." Merteuil, however, is harder to pin down, and Baudelaire ends up advising the reader to see for himself, directing him to Merteuil's self-portrait, Letter XXXVIII. See Baudelaire, notes on *Les Liaisons dangereuses* (1856) in Laclos, *Les Liaisons dangereuses*, ed. Seth, 632.

Dorothy Thelander astutely points out, however much *Les Liaisons dangereuses* has been called a "psychological novel," it is not "a psychological novel in the usual sense of the term. It does not take us into the minds of the characters…. It does not take us, we have to get there ourselves. Few other novels offer us such an opportunity to make up our own story, to stop and think, to look for hidden motives."[17] Laclos's novel does not present interior consciousness; so strictly speaking, we are never "in" the minds of the characters. The characters are presented as figures who perform actions and talk about them afterwards, not figures whose thoughts are visible to us. We see them only from the outside. For Thelander, this means that it is up to the reader to make up the difference.

If Laclos's text pushes us to make sense of character and action from an external perspective, Thelander's insistence that we make up the difference shows just how stubbornly readers have resisted that move. One of the more receptive strains of criticism, however, has been that which looks upon the driving force behind the novel as some species of determinism. Critics have talked about this fatalism in various terms: the machinations of the libertine protagonists; a rigid social order; and the semiotic imperatives of the epistolary form, to name a few.[18] Anne C. Vila has recently added "constitutional determinism" to the list. She argues that *Les Liaisons dangereuses* was informed by the work of contemporary moral anthropologists who posited a dimorphic doctrine of human types based on sensibility. The human constitution existed in one of two modes: soft, womb-like, and feminine or firm, cerebral, and masculine.

---

[17] Thelander, *Laclos and the Epistolary Novel*, 153.

[18] For psychological determinism, see Georges Daniel, *Fatalité du secret et fatalité du bavardage: La Marquise de Merteuil, Jean-François Rameau* (Paris: A.-G Nizet, 1966). For social determinism, see Peter Brooks, *The Novel of Worldliness*; Susan Winnett, *Terrible Sociability: The Text of Manners in Laclos, Goethe, and James.* For semiotic determinism, see Tzvetan Todorov, *Littérature et signification* (Paris: Larousse, 1967).

Vila claims that "the *type* of sensible constitution with which [Laclos's] novel's characters are endowed is what determines how they function in the world: they either resist the moral and physical stimuli that surround them, or they give in to them…. Resistance… is a constitutional quality that defines every character's fate."[19] What you are (born as) determines what you will do. Thus against critics who focus on Valmont's inner conflict, for example, Vila maintains that Valmont's "grave character flaws"—his lack of originality and, more importantly, his inability to resist sentiment—are inherent to his physiological constitution. Valmont's fate is sealed by his inborn predisposition to "involuntarily softness." Under this reading, character and action are more or less equivalent ("soft" is both a character trait and a way of responding to the world) and both boil down to an immutable constitution.

Vila's interpretation is attractive here because it seems to genuinely do away with the psychological lens. Characterization in *Les Liaisons dangereuses* is based on physiological makeup: what someone will do entirely depends on whether he is built to resist or yield to the world around him. This sounds like Hume in that action can be predicted from character, and likewise, character (or, in Vila's terms, constitution) is inferred from action (resisting or yielding). However this relationship between "doing" and either resisting or yielding is less clear than it at first appears. Vila seems to link resistance with action. Those born with a masculine constitution are able to stand firm against the world and carry out their plans. In other words, a capacity for performing actions is either part of your corporeal makeup or it is not; if you have a firm constitution, that is, a firm character, then you will be able to use that firmness to move others. This picture is complicated, however, by Vila's acknowledgment that "Valmont

[19] Anne C. Vila, *Enlightenment and Pathology: Sensibility in the Literature and Medicine of Eighteenth-Century France* (Baltimore: Johns Hopkins University Press, 1998), 270.

continues to exert a good deal of control over events, even in his death scene," despite the fact that over the course of the novel, his innately soft constitution leads him to suffer a complete loss of agency ("in part 4 of the novel, therefore, he is not redeemed but reduced, at least in terms of the libertine philosophical system to which he had previously subscribed").[20] By the end then, Valmont is both active and yielding, undoing the straightforward correspondence between agency and resistance.[21] This break poses a problem for Vila since it is, in part, this connection that permits her to sidestep the problem of psychology. Without physiological constitution to ground action, the necessary connection between motive and action seems to break down and, without it, Vila cannot explain why the characters do what they do.

Vila's reading breaks down because Valmont undergoes a change. It is Valmont's metamorphosis—the cold, cerebral libertine reluctantly turned lover—that seems to require a psychological explanation. The issue is, at bottom, that as the novel progresses, Valmont acts less and less like Valmont. That is to say, he acts out of character more and more frequently until finally, he no longer seems to be the same person. More accurately, however: he is not who he seemed to be or said he was. One of the assumptions underlying this hermeneutical problem is that we know what is in character and what is out of character for Valmont. That Valmont falls in love is by now a standard reading of the novel, but if one takes seriously the Editor's claim that virtually all ("presque tous") the sentiments expressed in the letters are either "feigned or dissimulated," then it becomes much more difficult to say what happens to Valmont or, indeed, if

[20] Ibid., 278.
[21] Thus Vila's argument follows in the same vein as that of Susan Dunn. Dunn holds that Valmont is a passive actor. His "sensibility makes his system—or lack of system—of acting more complicated…. The intrusion of emotion into Valmont's performances confuses, in his own mind, theater and self-identity." Valmont's performance becomes a reaction rather than an action (e.g. real tears in response to Mme de Tourvel's). See Dunn, "Valmont, Actor and Spectator." *French Review* 58 (Oct. 1984): 42.

anything happens to him at all. After all, he professes and denies his love for Mme de Tourvel countless times throughout the novel. Moreover, as Peter Brooks notes, the very term "love" is rendered suspect from the moment Valmont announces (in his first letter) that he must have Mme de Tourvel to save himself from the embarrassment of falling in love with her. That statement, argues Brooks, "is ambiguous… since 'love' has here only a conventional public meaning…. The nature of his sentiments toward Mme de Tourvel is, at this point in the novel, obscure, and must remain so, for the only accounts we receive are those designed for the Marquise."[22] That Valmont's sincerity is suspect at any given moment bears testimony to the opacity of his character. In fact, the strongest argument for Valmont's transformation from roué to infatuated lover is the novel's ending, and it is by reading backward that one pieces together how he gradually succumbs to Mme de Tourvel's sensibility.[23] However, reading character in this way takes development more or less as a given. It also requires setting aside those actions that do not fit this particular narrative: actions that do not contribute to Valmont's love story and cannot, within the framework of that story, be properly classified as either in or out of character. Thus the question of what it means to act out of character is merely deferred. Critics

[22] See Brooks, *The Novel of Worldliness*, 193.

[23] Patrick Byrne maintains that it is the fact that Valmont continues his affair with Mme de Tourvel even after he has had her that betrays his true sentiment. Thus Versini argues that Mme de Merteuil refuses to reward Valmont because he does not produce the written proof—and shows every outward sign to the contrary—of having conquered. Critics often point to the "missing" letter from Valmont to Mme de Volanges, which is in the MS but was replaced by Letter CLIV in the published version. The letter asks Mme de Volanges to assure Mme de Tourvel of his "repentance, regrets, and above all, my love." Critics point to this letter as a "true" expression of Valmont's feelings—hence omitted. Brooks explains: "we surely have sufficient evidence that Valmont's despair at loss of the Présidente is far from feigned—the artistic advantage gained by suppressing the letter is evident. For who could write such a letter? What would an unfeigning Valmont sound like? Where could he find the terms, the vocabulary, the code to express himself?"(Brooks, *The Novel of Worldliness,* 208). See Byrne, *Les Liaisons dangereuses: A Study of Motive and Moral* (Glasgow: University of Glasgow, 1989); Versini, *Le Roman le plus intelligent.*

psychologize Valmont in order to explain his contradictory actions, but this approach merely leaves other behaviors unaccounted for.

One way to properly deal with these contradictory behaviors might be, simply, to not account for them at all; that is, to forego questions of interiority. While psychological interpretations have dominated the Laclos criticism, reading Laclos's characters from the outside reframes conflicting causes so that instead of existing synchronically within character they stretch out diachronically in the form of a causal chain. What this means is that a character's action might seem to have multiple causes—it might look like an accident, too. Valmont sends Mme de Tourvel a letter breaking things off, a letter composed by Mme de Merteuil, and it has "an effect," just as he expected: "ce dernier m'a paru original et proper à faire de l'effet" (L, 391).[24] But although it is his intention to break with Tourvel (and to break with her at Mme de Merteuil's bidding, moreover), he fails to grasp the full import of Mme de Merteuil's missive:

> Je ne sais si j'ai mal lu ou mal entendu, et votre Lettre, et l'histoire que vous m'faites, et le petit modèle épistolaire qui y était compris. Ce que je puis vous dire, c'est que…. Aussi je l'ai copié tout simplement, et tout simplement encore je l'ai envoyé à la céleste Présidente. (L, 391)[25]

As a result, while his intention and action correspond—he intends to leave Mme de Tourvel by sending her Mme de Merteuil's letter—his action ends up carrying more weight than he had intended, dealing Mme de Tourvel a "fatal blow." Instead of reading this episode as Mme de Merteuil instructs us to, i.e. the triumph of Valmont's libertine vanity over his love for Mme de

[24] "[I]t seemed to me original and likely to make an effect" (A, 347).
[25]"I am not sure whether I have misread or misunderstood your letter, the story you relate and the little epistolary model which accompanied it. What I can tell you is that…. I simply copied it out, and still more simply I sent it out to the heavenly Madame de Tourvel" (A, 346-7).

Tourvel, we can then see it as a causal chain. Valmont's action is intentional, but it also seems accidental because it is overdetermined—it has picked up meaning along the way.

Valmont's action carries excess meaning because it happens in *a wrong way*. Action theorist Donald Davidson was one of the first to describe the problem of deviant causal chains, recognizing the fact that there are a number of ways to get from cause A to effect B (a number of chains, that is). In his classic essay, "Freedom to Act," Davidson cites Daniel Bennett's scenario of a man who attempts to shoot his enemy, aims and misses, but with the noise of the gunshot causes a stampede among a sounder of wild boar which trample his enemy to death.[26] The pig-stampeder's action (firing a gun) realizes his intention (to kill his enemy), only not in the way he had envisioned. The paradoxical consequence of a deviant causal chain is, then, that one might perform an action that one intended without intending the particular performance that makes the action a success. Thus in Valmont's case, although he leaves Mme de Tourvel just as he intended, we can take his subsequent mystification—"Étonné de ce silence, auquel je ne m'attendais pas…" (L, 393)[27]—as a sign that somehow, things were not supposed to go quite like that. In the novel's Newtonian moral universe, the reactions of others are (must be) entirely predictable. Valmont's surprise indicates that Mme de Tourvel is responding to something he had not anticipated. Because it does not occur in "the right way," it is a deviant causal chain. The larger point in terms of the novel, however, is that the injection of the unknown into the causal process raises questions about moral accountability: to what extent is the pig-stampeder responsible for a murder that he could not possibly have foreseen (death by boar)? Likewise, to what extent is Valmont accountable for Mme de Tourvel's response, which he did not predict?

[26] See Donald Davidson, "Freedom to Act," reprinted in *Essays on Actions and Events* (Oxford: Oxford University Press, 1980), 78.
[27] "Surprised by this silence, which I did not expect…" (A, 349).

The situations of Valmont and the pig-stampeder are not exactly parallel, of course. Whereas the action and intention of the pig-stampeder exactly match, Valmont's action gathers surplus meaning by virtue of the causal path it follows and, as a result, his action realizes his intention while at the same time exceeding it. Contemporary philosophers who, like Davidson, are primarily interested in deviant causal chains for the challenge they pose to a definition of intention, draw a strict line between the problems of causal deviance and overdetermination.[28] *Les Liaisons dangereuses,* on the other hand, conflates the two precisely in order to explore the moral implications of accident. That is, the novel provides a way for thinking about how excess meaning is generated through a deviant causal chain and what this means for moral accountability. Thus, how should we understand, in moral terms, the fact that Mme de Merteuil guides Valmont's hand to strike the mortal blow? This action can only be figured as a causal chain since Mme de Merteuil's letter is not by itself a sufficient cause—the letter must come from Valmont—but the case is not perfectly analogous to that of the pig-stampeder since Mme de Merteuil's letter has significantly altered the meaning of Valmont's action. Because the action is yet faithful to Valmont's aim, however, the deviation is not an accident in the usual sense. At the heart of *Les Liaisons dangereuses* is, then, a question: is a person responsible for his action if it turns out to be just what he wanted—and then some?

It is a question that puts pressure on the way Hume's theory of causation determines moral accountability. Deviant causal chains never crossed Hume's mind—and for good reason. For Hume, because causation is essentially a fiction, there is no reason why any particular chain should be the considered the "right" one. Moreover, since Hume believes that "cause and effect

[28] Philosophers generally define overdetermination as a case in which two independent and individually sufficient causal chains produce the same event.

must be contiguous in space and time," each effect can be linked to its antecedent, but not to causes further up the chain.[29] Another way to think about this is to take an empiricist point of view: I can observe the connection between the phenomena A and B, and between B and C, but if I have not actually seen the connection between A and C, I cannot say they are linked. For Hume, moral judgment is a relatively easy business simply because it does not take causal chains into account. The relationship between motive and action is always one-to-one, which is to say that the relationship between character and action is also one-to-one. *Les Liaisons dangereuses,* on the other hand, begins from the premise that the first cause and final effect of a causal process are linked.[30] In imagining causal chains as prying open the connection between intention and action such that unexpected causes might occur in between, the novel suggests an unsettling possibility, which is that a person might perform an action that is in character but perform it in a manner that is out of character, thanks to a causal chain gone wayward.

III

Valmont himself admits to one wholly uncharacteristic action and, as passionate about self-analysis and self-portraiture as he is, he wastes little time wondering why. His disclaimer is simply that he never blames himself for a good action when it gives him exertion or amusement: "je ne me reproche pas une bonne action, pourvu qu'elle m'exerce ou m'amuse" (L, 170). His good action, in this case, involves saving a married woman from being caught *in flagrante delicto*. Having spent the night with the Vicomtesse de M-- with her lover, Vressac,

[29] Hume, *Treatise,* 173.
[30] This logic is inbuilt to the sentimental novel, whose plot revolves around that all-important "first fatal step." See for example Richardson's *Clarissa* (1748); Rousseau, *Julie* (1761); Henry MacKenzie, *Julia de Roubigne* (1777); Maria Edgeworth, *Belinda* (1801).

and her husband sleeping just across the hall, Valmont escorts her back to her room only to find that she has locked herself out. He describes the scene to Merteuil:

> Vous n'avez pas l'idée de l'expression de désespoir avec laquelle la Vicomtesse me dit aussitôt : "Ah! Je suis perdue." Il faut convenir qu'il eût été plaisant de la laisser dans cette situation : mais pouvais-je souffrir qu'une femme fût perdue pour moi, sans l'être par moi? Et devais-je, comme le commun des hommes, me laisser maîtriser par les circonstances? Il fallait donc trouver un moyen. (L, 172-73) [31]

This scene is rarely discussed in any detail, but it recalls an earlier act of kindness on Valmont's part that is a favorite in the criticism.[32] In Letter 21, as in the case above, Valmont is faced with a figure of "despair." Valmont arranges for Mme de Tourvel's spy to witness him saving a poor peasant family from eviction, the idea being that she will believe such benevolence is de rigueur with him. This action is not out of character—Valmont has an ulterior motive—but his *reaction* is. The gratitude of the poor people is a spectacle more moving than Valmont had anticipated:

> Après cette action si simple, vous n'imaginez pas quel chœur de bénédictions retentit autour de moi…. Quelles larmes de reconnaissance coulaient des yeux du vieux chef de

[31] "You can have no idea of the expression of despair with which the Vicomtesse said to me at once: 'Ah! I am ruined.' It must be admitted that it would have been amusing to leave her in this situation, but could I allow a woman to be ruined *for* me without being ruined *by* me? And ought I, like the majority of men, to let myself be dominated by events? A way out had to be found" (A, 149).

[32] Thelander, for example, groups the Vressac episode together with the Prévan tale and dismisses both as superfluous "display pieces": "From the point of view of the novel, Laclos needed something to fill up time while Valmont could get ready to seduce Cécile and while the infatuation of La Présidente might develop.…. Having one's two best narrators with nothing to narrate is an awkward situation" (*Laclos and the Epistolary Novel*, 59).

cette famille, et embellissaient cette figure de Patriarche, qu'un moment auparavant l'empreinte farouche du désespoir rendait vraiment hideuse! (L, 57) [33]

Critics tend to point to Valmont's response as a sign of his innate sensibility, his inability to separate performance from passive spectatorship and, ultimately, the chink in his libertine armor.[34] If part of Valmont's enjoyment comes from the way the peasants apotheosize him—"Tombons tous aux pieds de cette image de Dieu" (L, 57) [35]—an equal measure comes from the delicious emotion the Greuzian tableau inspires.

But if the despair of the aged Patriarch draws tears from Valmont, the Vicomtesse's despair has the reverse effect. This response fits ill with Vila's claim that Valmont suffers from "unplanned feeling" when he comes face to face with sentimental stimuli. If Valmont's experience with the peasants suggests a yielding constitution, his reaction to the Vicomtesse's distress implies otherwise, for there is no sign of sympathy here. For Valmont (and perhaps for the reader as well), it would have been amusing indeed ("il eût été plaisant") to leave the Vicomtesse in her predicament. Be that as it may, he chooses to exert himself on her behalf. He provides an explanation. *Pour* [for] and *par* [by] might seem like a fine distinction, but Valmont is not splitting words. Accustomed to performing actions, Valmont refuses to submit to

[33] "After so simple an action you cannot imagine what a chorus of benedictions echoed round me from the spectators! What tears of gratitude flowed from the eyes of the aged head of the family and embellished this patriarchal face which a moment before was rendered truly hideous by the wild imprint of despair!" (A, 47-48).

[34] Brooks states "the admission to an involuntary emotion is an important indication of the extent to which parody can become emotional reality for Valmont," (*The Novel of Worldliness,* 191). Thelander: "The unpleasant feeling produced by this scene is not based on fake sentimentality hiding behind the real thing. There are three levels of emotion in it: the feelings expected by society, those Valmont is willing to allow himself to feel, and those he actually experiences. There is a closer connection between the first and third reactions than Valmont can admit and herein lies his dilemma" (*Laclos and the Epistolary Novel,* 88). See also Vila, *Enlightenment and Pathology* and Dunn, "Valmont, Actor and Spectator."

[35] "Let us all fall at the feet of this Image of God" (A, 48).

circumstance, and instead of playing an incidental role in the Vicomtesse's humiliation, he chooses to orchestrate her escape. Suellen Diaconoff argues that Valmont is a narcissist and that his seduction of the Vicomtesse is "necessitated in his mind so that the event [will] revolve about him and his desires rather than the woman." Valmont's decision to intervene, she says, "makes clear to the reader [his] need to maintain absolute control over external events, both because of an obsessive need for power and for his own self-protection."[36] Diaconoff's assertion rings true and critics have more or less taken Valmont at his word when he insists that he will not be "dominated by events."[37] Valmont would, perhaps, rather be the cause of a good action than instrumental to a bad one.

Is it that simple? Valmont's preface—"si j'ai le talent de perdre les femmes, je n'ai pas moins, si je veux, celui de les sauver" (L, 170)[38]—dismisses the incident as a (potentially) routine event, but he has played this trick before. Readers who interpret Valmont's good deed toward the Vicomtesse in terms of his narcissistic "character" run the risk of misinterpreting the scene in much the same way as Mme de Tourvel did previously. She, upon learning that Valmont's man had gone ahead of him to find out who in the village was in need of help, concludes that "ce n'est meme plus seulement une compassion passagère et que l'occasion

---

[36] Suellen Diaconoff, *Eros and Power in Les Liaisons dangereuses: A Study in Evil* (Paris: Librairie Droz, 1979), 68.

[37] Lester Crocker argues that "the point is to work his will. He will harvest whatever satisfactions come with it, but working his will is the *summum bonum.*" See Lester Crocker, "The Status of Evil in *Les Liaisons dangereuses,"* in *Laclos: Critical Approaches to Les Liaisons* dangereuses, Lloyd R. Free, ed., (Madrid: José Porrúa Turanzas, D.L. 1978), 84. Brooks argues that in the diegesis, every conquest one makes forces other people to conceive their roles in relation to your self-representation (*The Novel of Worldliness,* 184).

[38] "If I have a talent for misleading women I have none the less, when I wish, that of excusing them" (A, 147).

determine: c'est le projet formé de faire du bien" (L, 59).[39] She is right, of course, that Valmont's act is not a matter of being in the right place at the right time, but the inference she draws is, as we know, wide of the mark. Valmont's action here, the rescue of his bed partner, is no more routine (that is, in character) than the one-off act of charity that Tourvel, in all her naïveté, takes to be habit.

Upon closer inspection there is something rather odd about Valmont's logic. It is not quite accurate to say that without his interference the Vicomtesse would not have been ruined *by* him. Contrary to his fears, he would have "ruined" her very properly indeed in the very literal sense of "dishonoring a woman by seducing and abandoning her." She would have been "perdue," again, in the very real sense of losing her honor and social station, and in another sense, too: "perdre" signifies the act of profiting from any occasion that presents itself.[40] Valmont has achieved his goal by any standard. At the same time (and despite the fact that he rescues her), it is equally inaccurate to say that Valmont does not want to ruin the Vicomtesse. After all, he owes her nothing. Valmont rewards the grizzled peasant ten louis for the pleasure he received from his performance, but after their night together, Valmont is more pleased with himself than with the Vicomtesse. As he tells Mme de Merteuil: "vous me connaissez, et j'ai été content de moi" (L, 172).[41] Moreover, he has, in the fashion of a typical rake, no qualms about either circulating the story or putting names to the parties involved, which means that in fact he has only deferred the Vicomtesse's ruin—postponing her humiliation to another day (which will

[39] "It is not merely a passing compassion provoked by accident; it is a preconceived plan of doing good" (A, 49).
[40] Definitions of "perdu(e)" include: "lui ôtant l'honneur"; "dont la situation est désespérée"; and "employer ou profiter de (qqch.) sans en tirer ce qu'on attendait." See *OED* «ruin v. trans. 2 e; *Le Grand Robert* "Perdre."
[41] "You know me, and I was pleased with myself" (A, 149).

not be far off, Mme de Merteuil assures him). Nothing that occurs is out of keeping with Valmont's intentions or desires. It is only the turning point—precipitated by Valmont's impression that he has been the victim of circumstances—that strikes us as strange.

Why then does Valmont ally himself so closely with the Vicomtesse's (un)lucky star? Diaconoff claims that as a narcissist, Valmont must "be at the center of any event," but over the course of the novel, Valmont frequently takes up the role of dispassionate observer. His fondness for spectatorship should only increase his appreciation for the symmetry afforded by the locked door: that Valmont is rooming across the hall is, he observes, bad luck ("mauvais destin") for the husband and lover because it allows the Vicomtesse to slip out of her chamber; her bad luck is that she is unable to slip back in. And although Valmont claims that he will undertake a good action if it gives him exertion *or* amuses him, here he exerts himself at the *expense* of his amusement. Thelander and Joyce O. Lowrie offer one explanation as they both discuss the Vressac episode as a counterweight to or mirror of the Prévan affair, noting how Valmont uses the story to position himself relative to Mme de Merteuil and Prévan respectively.[42] In their readings, Valmont's escapade is a means for him to establish bragging rights, to remind both Mme de Merteuil and the reader of his prowess.

Within the context, however, Valmont is remarkably humble. According to Valmont, his desire for the Vicomtesse is sparked not by the love quadrangle, as Diaconoff argues, but by the Vicomtesse's insistence that it is "impossible" to meet him: "Cela m'est impossible, me répondit-elle…. ce mot d'impossible me révolta comme de coutume. Je me sentis humilié d'être

[42] See Thelander, *Laclos and the Epistolary Novel*; Joyce O. Lowrie, "The Prévan Cycle as Pre-Text in Laclos's *Les Liaisons dangereuses," Sightings: Mirrors in Texts—Texts in Mirrors* (Amsterdam: Rodopi, 2008), 77-101.

sacrificié à Vressec, et je résolus de ne le pas souffrir: j'insistai donc" (L, 170).[43] While Valmont treats the mere word as a challenge, his next letter to Mme de Merteuil reveals that there is really only one feat he considers "impossible." Again warning her to beware Prévan, he confesses that:

> Je ne me crois pas plus bête qu'un autre; des moyens de déshonorer une femme, j'en ai trouvé cent, j'en ai trouvé mille : mais quand je me suis occupé de chercher comment elle pourrait s'en sauver, je n'en ai jamais vu la possibilité. (L, 183)[44]

Valmont knows an infinite number of means by which to ruin a woman, but not one by which she might save herself. Thus he corrects himself when he asks Mme de Merteuil to drop Prévan so that they can talk of other things: "D'autre chose! Je me trompe, c'est toujours de la même; toujours des femmes à avoir ou à perdre, et souvent tous les deux" (L, 183).[45] There are no other things—women had or ruined or both: this is the only topic of conversation for a pair of libertines.

It is for that very reason that the Vicomtesse's shame must be published as quickly as possible. As Mme de Merteuil points out:

> Comme vous dites, il faut qu'on en parle; car si l'occasion a pu vous engager, comme je le conçois, à préférer pour le moment le mystère à l'éclat, il faut convenir que cette femme ne méritait pas un procédé si honnête. (L, 179)[46]

---

[43] "'Impossible,' she replied… That word 'impossible' roused me, as it always does. I felt humiliated at being sacrificed to Vressac and I resolved not to endure it; so I insisted" (A, 147).
[44] "I do not think I am more stupid than others. I have found a hundred, I have found a thousand ways of dishonouring a woman; but as hard as I look, I have never been able to imagine a way for her to avoid it" (A, 159).
[45] "Of something else! I am wrong, it is always the same thing; always women to have or to ruin, and often both together" (A, 159).
[46] "As you say, it must be talked about; for if the occasion led you, as I can well understand, to prefer mystery to scandal for the moment, it must yet be admitted that the woman did not deserve such good treatment" (A, 155).

The Vicomtesse cannot be allowed to escape ruin altogether, says Mme de Merteuil, for "that woman does not deserve such good treatment." In saving the Vicomtesse Valmont is not guilty, says Mme de Merteuil, but things must be put right. "That woman" has escaped ruin without saving herself, which is something that even Mme de Merteuil cannot do since it is not something that can be "done." Escaping ruin without *escaping*, that is, without acting, is a feat that relies upon the actions of others over whom one seemingly has no influence (recall that Valmont is untouched by the Vicomtesse's despair) and, what is more, it counts on those actions being seemingly out of character (a rake turned hero). It is out of a woman's control: everything depends upon a door accidentally locked. Mme de Merteuil's complaint is that things have fallen into place for the Vicomtesse through no will of her own: "procédé" also means "design." Thus the Vressac episode poses a problem not because Valmont "acts out of character"—he has, he must have, reasons for doing so—but because the event itself is an anomaly. Chance has intervened in a Newtonian moral universe.

Returning then to Valmont's original anxiety ("mais pouvais-je souffrir qu'une femme fût perdue pour moi, sans l'être par moi?"), it becomes clear that his formulation is not quite correct. The problem is not that the Vicomtesse will be ruined *for* him but not *by* him, but rather that she will be ruined by him *and* for him. There is no question of a failed intention here, only an intention-plus, so to speak. A straightforward causal relationship (Valmont ruins the Vicomtesse) has turned into a deviant causal chain (Valmont seduces the Vicomtesse and she accidentally locks herself out and she is ruined: once again, she is both had and ruined because she is seduced and abandoned), which Valmont transforms into a non-deviant causal chain (Valmont seduces the Vicomtesse, rescues her, and ruins her). There is something amusing about the whole situation after all—for the reader, that is. In taking seriously the idea that the locked

door might, not just detract from, but take credit for his exploit, Valmont has supplanted the Vicomtesse as the butt of the joke. Valmont has got things upside down in that it is not he who is instrumental to the Vicomtesse's ruin, but the lock (the accident). He fails to recognize that he is by himself a necessary and sufficient condition for the Vicomtesse's dishonor; the lock is simply icing on the cake. The bigger point, however, is that despite Valmont's jealous rivalry with a locked door, his triumph (or guilt, depending on your point of view) is indisputable. In all three cases—the simple cause-effect relation, the deviant causal chain, and Valmont's corrective, the non-deviant causal chain—Valmont is, regardless of how many events occur in between, the cause of the Vicomtesse's ruin. It seems then that the first cause in a causal chain is responsible for everything that follows.

I have been arguing that if we read *Les Liaisons dangereuses'*s treatment of character not as an attempt to represent mind, but as an empirical investigation into the nature of causation, we can see that the novel both fleshes out Hume's theory of character and troubles it. By thinking through cases where actions are completed not immediately, but by way of causal chains, *Les Liaisons dangereuses* shows us that *how* we do something is just as important as *what* we do. For Laclos, it is this "how"—always prey to the influences of the external world—that accounts for acting out of character. When I act not like myself, it is because some causal chain has altered my relationship to my action. In this way, acting out of character is a reminder that actions, once performed, have a life of their own.

Chapter Four

"I Was the Hero's Father!": Liberty, Partiality, and Justice in William Godwin's *St. Leon*

The story which I now take up the pen to relate, derives no interest from myself…. But, though insignificant in myself, and uncharacterised by those vehement passions or that inordinate ambition, which places some men on the roll of the distinguished, and perpetuates their memory to honour or to shame, it has been my lot to be connected with persons whose story has a more substantial claim on the curiosity of mankind. It is their adventures, and not my own, that I am about to relate.

Godwin, *Cloudesley: A Tale* (1830)

One commonplace objection to the doctrine of necessity is that it promotes apathy. If people thought themselves powerless to affect the world, so the argument goes, they would cease to care about it. In *An Enquiry Concerning Political Justice* (1793), William Godwin pooh-poohs this logic.[1] He explains that "nothing could be more unreasonable than that the sentiment of necessity should produce in me a spirit of neutrality and indifference. The more certain is the connexion

[1]Thinkers like Thomas Reid (responding to Hume) and Samuel Clarke (responding to Leibniz) saw fatalism as the natural consequence of necessitarianism, but Godwin was careful to distinguish causal necessity from the kind of determinism associated with Calvinism or the Spinozists. Bernard Williams usefully reminds us that "'Fatalistic' does not pick out a class of arguments, but a class of beliefs or attitudes: those that involve the idea that action, choice, and so forth have no effects; that everything will be the same whatever you do" in Williams, *Making Sense of Humanity: And Other Philosophical Papers, 1982-1993* (Cambridge: Cambridge University Press, 1995, 11.See Jerome B. Schneewind, *The Invention of Autonomy: A History of Modern Moral Philosophy* (Cambridge: Cambridge University Press, 1998) for a good overview of the debate. For a discussion of anti-Stoicism in the eighteenth century, see for example Henry W. Sams, "Anti-Stoicism in Seventeenth- and Early Eighteenth-Century England." *Studies in Philology* 41, no. 1 (1944): 65-78. The most thorough literary treatment of these ideas is, of course, Denis Diderot's *Jacques le fataliste et son maître* (1796).

between effects and causes, the more chearfulness should I feel in yielding to painful and laborious employments" (PJ, 169). Far from discouraging us from pursuing our aims, our faith in necessary connection spurs us on when the going gets tough. Great deeds require great perseverance, and it is only because we know that specific actions will yield specific results that we can be happy even in hardship. Thus for Godwin, the necessarian is "superior to the tumult of passion" not because he is indifferent to worldly events—the opposite is true—but because he eschews the feelings of regret and resentment that go hand in hand with the doctrine of free will (PJ, 171). Whereas the libertarian unjustly blames himself and others when his plans go awry, the necessarian sees that each action is "a link in the great chain of events which could not have been otherwise than it is" (PJ, 170). For Godwin, recognizing the principle of necessity does not make us sluggish or unfeeling. Rather, it is by acknowledging necessity that we can be happy in a world that we affect but cannot direct.

Given the promise of "chearfulness" in *Political Justice,* it is striking that one of the clearest articulations of necessity in Godwin's second novel, *St. Leon: A Tale of the Sixteenth Century* (1799), comes from the mouth of Bethlem Gabor, the brooding misanthrope who imprisons St. Leon in an underground vault in order to stop him from executing his philanthropic projects. Although Gabor himself justifies his war on humankind with an explicit appeal to the logic of necessity, critics have tended to follow St. Leon in reading Gabor's misanthropy as a disorder of the passions. To hear St. Leon tell it, Gabor is "a great and admirable man" whose "mighty soul" has been corrupted by suffering: "all the unsocial propensities that animated him, were the offspring of love, were the sentiments of a lioness bereaved of her young."[2] According

[2] William Godwin, *St. Leon: A Tale of the Sixteenth Century,* ed. William D. Brewer (Peterborough: Broadview Press, 2006), 386. Hereafter cited parenthetically in the text as SL.

to St. Leon then, Gabor's misanthropy is, as one critic puts it, "a perversion of the inborn human craving for love."[3] Yet Gabor resists this diagnosis. He himself describes his hatred as a matter of understanding rather than a species of that "unreflecting feeling" Godwin speaks of so disparagingly in *Political Justice.* His misanthropy is born out of the realization that human action is mechanical, which is also to say universal. As he explains to St. Leon, "My revenge is not causeless; this [the slaughter of my family] was not the act of individuals. All men, in the place of these murderers, would have done as they did…. All men are excited by the same motives, urged by the same temptations, influenced by the same inducements. Why should I attempt a futile distinction, when nature had made none?" (398). For Gabor, everyone is a potential assassin and the question of who actually commits a given crime is merely academic. Here he echoes Godwin's claim in *Political Justice* that necessity will teach us to detach actions from their authors: "we shall… no more be disposed to repent of our own faults than of the faults of others" (PJ, 170). However, whereas in *Political Justice* the break between doer and deed renders concepts like culpability and responsibility obsolete—Godwin writes that "under the system of necessity the ideas of guilt, crime, desert and accountableness have no place" (PJ, 170)—Gabor reaches the opposite conclusion. For Gabor, necessity means that everyone is equally blameworthy. Far from making him "superior to the tumults of passion," subscribing to the principle of necessity transforms the malice Gabor bears towards the murderers of his family into full-blown misanthropy.

I begin with Gabor because his necessitarianism-turned-misanthropy exemplifies Godwin's efforts to rethink the account of human action laid out in *Political Justice*. Although commentators have often discussed the extent to which *St. Leon* modifies (or, less frequently,

[3] B. Sprague Allen, "Godwin as Sentimentalist" *PMLA* 33, no. 1 (1918): 22.

reaffirms) the ideas put forth in Godwin's earlier work, the novel's innovative treatment of the free will problem has been largely overlooked. Scholars who see *St. Leon* as deeply revisionist tend to focus on two aspects of the text: its genre and its representation of the domestic affections. This emphasis makes sense insofar as it addresses what are perhaps the two most glaring differences between *Political Justice* and *St. Leon.* In writing a picaresque about a man who possesses the philosopher's stone and the elixir of life, Godwin appears to abandon his mission to show "things as they are." As Christopher Lake Moody, one of the first to review the novel, complains, "Why imagine incredible situations and absolute impossibilities, in order to work on our feelings, passions, and convictions? …Of what use can such idle imaginations be to man in the actual state of his existence? We would put these questions to Mr. Godwin, because he would have it supposed that he has a moral and philosophical purpose to answer by his writings."[4] For many critics, the very form of *St. Leon* seems to undermine Godwin's aims in *Political Justice* and *Caleb Williams* by indulging useless flights of fancy rather than guiding the reader to a "clearness of perception" regarding "the moral concerns of mankind" (PJ, 171). The novel's positive portrayal of the passions distances it from *Political Justice* even further, for *St. Leon* seems to refute the anti-affectional rationalism promoted in the earlier text. Moody takes a standard line when he argues that the novel is proof of "a complete revolution in [Godwin's] opinion respecting the affections and charities of private life; so that, though these qualities met with no indulgence in the 'Enquiry concerning Political Justice,' they are in the present work treated with respect, and honoured with an impassioned eulogy."[5] There is then a long tradition

[4] Christopher Lake Moody, review of *St. Leon,* by William Godwin, *The Monthly Review* 33 (1800) in appendix C of Godwin, *St. Leon,* ed. Brewer, 471.
[5] Moody, review of *St. Leon,* 469-70. In his Preface, Godwin describes this change as a revision, but most readers have tended to see it, following Moody, as "a complete revolution in his

of reading *St. Leon* as a recantation of the realism and strict rationalism Godwin celebrates in his earlier work. The novel marks a turn in Godwinian thought, an attempt to find a place for the imagination and the passions in his theory of justice.

The greater philosophical stakes of this turn remain underexplored. This is partly because the logical link between the novel's fictional mode and its recovery of the passions is not obvious. Although Moody briefly hints at a connection—he regrets that the impossible romance appeals to "our feelings, passions, and convictions" rather than to our understanding—he otherwise addresses Godwin's seeming abuse of poetic license and his change of heart regarding the private affections as two distinct interpretative concerns. Most readers take a similar approach.[6] It is only relatively recently that critics have begun to suggest that Godwin's choice of genre was primarily determined by philosophical considerations.[7] Tilottama Rajan sees *St. Leon* as a challenge to "the realist novel… [which] depends on an equivalence of intra- and extradiegetic worlds that is essentially conservative." For Rajan, Godwin's romance is a fantasy that forces the reader to allow that even the most farfetched ideas in *Political Justice*—from

---

opinion." For an example of a less revisionist account of the novel see Allen, "Godwin as Sentimentalist."

[6] Since Godwin's contemporaries by and large applauded the recuperation of passion in *St. Leon,* on the one hand, and objected to its outlandish premise, on the other, they often presented the novel's style as an obstacle to be overcome in order to arrive at its "just observation and acute remark." The review in the *Monthly Magazine, and British Register* is typical on this point: "if it be possible for the reader to forget the impossibility of the principal events on which the history hinges, he will find parts of this novel of the most striking and interesting kind" (Unsigned review of *St. Leon,* by William Godwin, *Monthly Magazine, and British Register* [20 Jan. 1800] in appendix C of Godwin, *St. Leon,* ed. Brewer, 474).

[7] Of course, recent critics have had the benefit of reading Godwin's posthumously published essay, "Of History and Romance" (written in 1797, not published until 1988). There Godwin argues that the romance writer is the true historian because he shows how man's passions produce actions instead of just recording a series of events. See Godwin, "Of History and Romance." Accessed October 30, 2016, http://www.english.upenn.edu/~mgamer/Etexts/godwin.history.html

man's potential immortality to a future without government—*might* be true.[8] Along the same lines, Andrea Charise maintains that "the speculative mode is essential to constructing and asserting the fictionality of old age, and helps us to understand how Godwin's philosophical speculative strategy ultimately demands a specifically novelistic intervention to correct the error at the basis of his claims in the Appendix [of *Political Justice*]."[9] *St. Leon'*s unrealism is, according to Charise, "a literary method" that enables the reader to speculate upon—that is, to evaluate—Godwin's philosophical speculations, particularly those concerning the elimination of old age and death. For Rajan and Charise, it is the very implausibility of *St. Leon* that makes the reader seriously consider the possibility of its premise, that is, the possibility of the improbable claims set forth in *Political Justice.* [10]

Both Rajan and Charise focus on the question of perpetual life as, in Charise's words, the theme that "most directly suggest[s] [*Political Justice*'s] intellectual linkage with *St. Leon.*"[11] While this focus seems right, it makes it easy to overlook another equally immediate "intellectual

---

[8] Tilottama Rajan, *Romantic Narrative: Shelley, Hays, Godwin, Wollstonecraft* (Baltimore: Johns Hopkins University Press, 2010), 171-72.

[9] Andrea Charise, "'The Tyranny of Age': Godwin's *St. Leon* and the Nineteenth-Century Longevity Narrative." *English Literary History* 79, no. 4 (2012): 914. Charise is anticipated by Moody and others—but they take the opposite view. They think fictional case studies are cop-outs (because of course the case will support the theory if it is invented precisely for that purpose). Evan Radcliffe argues, on the other hand, that narrative—as the primary way in which we construe experience—is the only possible test for theory. See Radcliffe, "Revolutionary Writing, Moral Philosophy, and Universal Benevolence in the Eighteenth Century."

[10] Interestingly, Charise and Rajan read the novel's representation of the passions very differently. Charise takes the domestic affections turn for granted. For her part, Rajan argues that the novel is critical of the domestic affections as propping up the institution of the family, which is inherently conservative. What looks like cold-heartedness on St. Leon's part is a recognition of "the extent to which family responsibilities, by limiting the spirit of adventure, dampen speculation and keep things as they are" (*Romantic Narrative,* 170). In other words, St. Leon's detachment makes him a free agent.

[11] Charise, "The Tyranny of Age," 906. In *Romantic Narrative,* Rajan casts a wider net, maintaining that the main "operative metaphors" of the text are gambling and alchemy. Rajan agrees, however, that immortality is integral to St. Leon's being a free agent (144).

linkage" between the two texts: their shared concern with the problem of free will.[12] The novel's sheer implausibility is, I argue, part of a larger critique of the doctrine of necessity. Importantly, it is not just the opening premise of *St. Leon* that is hard to swallow. As the reviewer for the *Monthly Mirror* grumbled, "This, however, is not one of those bold fictions, which, though founded on a supposition irreconcilable with the originary course of human events, discovers such a felicity of arrangement, and consistency of character,—such a 'dependency of thing on thing,' that the reader, by taking the outset for granted can readily account for all the subsequent stages of the narrative."[13] Rajan makes a similar point when she argues that *St. Leon* is an adventure "organized by a logic of escape" that tests the reader's credulity more and more as the story wears on.[14] For the anonymous reviewer above, however, the novel persistently violates causal logic not only at the level of plot, as Rajan points out, but also at the level of character.[15] Thus the character of Gabor, for example, is unrecognizable as a human being: "[Gabor] looked, and spoke, and thought, as never human being thought, looked, or spoke, and… has a motive for his cruelty which it never did, nor can, enter into the heart of man to conceive" (477). According to the *Monthly Mirror* review, because *St. Leon* fails to establish necessary connection between

[12]Rajan and Charise both hold that the text's most immediate "specifically novelistic intervention" is to force the reader to entertain the idea of human immortality (and its consequences) as something more than the enthusiastic ravings of a notorious radical. After all, the novel's "speculative mode" hinges on the fiction of the *elixir vitae*, which speaks directly to Godwin's musings in the Appendix of *Political Justice.* Yet this focus sidelines the domestic affections from the discussion of genre. In doing so, it excludes what has been one the most enduring issues in the secondary literature on the novel.

[13] Unsigned review of *St. Leon,* by William Godwin, *The Monthly Mirror: Reflecting Men and Manners,* January 1800, in appendix C of Godwin, *St. Leon,* ed. Brewer, 476.

[14] Rajan's point here, put crudely, is that each time the reader chooses to continue reading in the face of St. Leon's incredible escapes etc., she is tacitly agreeing that the incredible is possible (Rajan, *Romantice Narrative,* 172).

[15] Godwin talks about necessity both in terms of the "great chain" and the connection between passion and action, but he is not always clear about how they relate to one another. The implicit claim seems to be that men's actions are necessary and part of the chain.

motive and action, its characters are as improbable as its plot. I will return to the question of whether or not this is a wholly accurate characterization of the novel. For now, however, I want to think about why Godwin might question necessary connection at precisely that moment when he seems to be taking steps to put it on more solid footing. Why does Godwin recover the passions as the driving force of human behavior only to undermine this account by denying a necessary connection between motive and action?

That *St. Leon'*s recuperation of the domestic affections goes hand in hand with a destabilization of necessary connection gestures towards the problem Godwin's turn to the passions poses for his theory of action. In *Political Justice,* passion is always self-interested and, as a result, it cannot be a motive for just action. Justice means putting aside our private affections and acting based on rational calculation, not sentiment. The trouble for Godwin is then that once he accepts that passion rather than reason is the driving force of human behaviour, he has no way to explain disinterested action. *St. Leon* offers a unique solution to this quandary. When St. Leon finally escapes Gabor's clutches near the end of the novel, he decides to abandon his efforts at large-scale philanthropy and to devote himself exclusively to the happiness of his son. St. Leon insists, however, that this partiality is not a "selfish propensity." To the contrary, he insists that his dedication to Charles allows him to "forget and trample upon every personal concern" (413). Thus while in *Political Justice* family fondness implied self-love, *St. Leon* associates the domestic affections with a radical disinterestedness ("self-oblivion"). At first glance, it might look like Godwin is simply abandoning universal benevolence.[16] I will argue, however, that St. Leon is not so much detaching disinterestedness from universal benevolence as bringing

[16] This would, of course, be in line with Hume's position: "In general, it may be affirm'd, that there is no such passion in human minds, as the love of mankind, merely as such, independent of personal qualities, of services, or of relation to ourself" (*Treatise,* 481).

universal benevolence and personal feeling into alignment. The novel reconciles what were, in *Political Justice,* two conflicting impulses by redefining the general love of humankind as a passionate affection for a person of "superior usefulness and worth." In *St. Leon,* universal benevolence is not a matter of *judging* impartially, but rather of *feeling* partial towards those who deserve it.

This new account of universal benevolence needs further tweaking. If justice is love directed towards a worthy object, Godwin needs some way to explain how we are to find our bearings. How is it that we come to love the person who most deserves it? The unity of universal benevolence and passion hinges on this question, and as the terms "worthy" and "deserving" imply, the answer requires a return to morality. But this creates a new problem for Godwin, for in bringing back ideas like desert and virtue, it also brings back the question of how morality can exist without liberty (whereas in *Political Justice,* Godwin was content to forfeit both). *St. Leon* uses narrative point of view to navigate this problem. One of the novel's chief insights, I argue, is that Hume had it backwards when he associated the "false sensation of liberty" with first-person experience. Godwin's claim in *St. Leon* is that though we rightly recognize ourselves as bound by necessity, we see others as free, and it is this mistaken belief that leads us to deem them virtuous (or not). If universal benevolence is a love we feel for those people we believe to be virtuous, it is our faith in the liberty of others that makes that love possible.

In what follows I show that although Godwin's turn to the passions was more damaging to the principle of justice than he liked to admit, it also enabled him to smooth over some of the internal inconsistencies of *Political Justice. St. Leon* reworks the theory of action described in *Political Justice* in order to reconcile universal benevolence with passion, on the one hand, and necessity with morality, on the other. Ultimately, I argue, Godwin's turn away from reason is a

turn away from the arithmetical evaluation he associates with justice towards a morality based on a kind of love that is passionate, partial, and yet radically disinterested.

I

On the face of it, *St. Leon* seems to endorse the doctrine of necessity laid out in *Political Justice*. The novel opens with the title character's boast that he can, thanks to his alchemical powers, dictate the actions of others by working on that passion everyone possesses: ambition (and greed, which is one of its guises). He brags that he "can command, to an extent almost inconceivable, the passions of men. What heart can withstand the assault of princely magnificence? What man is inaccessible to a bribe?" (53). There is no man who cannot be bought at the right price, and with unlimited riches, St. Leon can buy anyone. St. Leon repeats this claim throughout the novel, likening himself to an omnipotent and benevolent God. He argues that the alchemist "possesses that attribute which we are accustomed to ascribe to the Creator of the universe…. The whole world are his servants, and he, if his temper be noble and upright, will be the servant of the whole world. Nay, it cannot happen otherwise…. Weakness and want are the parents of vice. But he [the alchemist] possesses every thing; he cannot better his situation; no man can come into rivalship or competition with him" (187). Wickedness stems from poverty and, likewise, wealth—which St. Leon equates with power—gives rise to virtue, and unlimited wealth to perfect virtue.[17] God has no ambition, and because He wants nothing Himself, He is able to be supremely benevolent. Similarly, the elixir of life and the philosopher's stone set St. Leon above ordinary men. He can afford to be generous because his benevolence costs him nothing: "he may

[17] This sounds a lot like *Political Justice* in that it reduces morality to justice. There Godwin argues that ambition is born out of a sense of (in)justice, the belief that one is "out of his place" and "wish[es] to be in it." Milton's devil rebels because "he saw no sufficient reason for that extreme inequality of rank and power which the creator assumed" (PJ, 144).

give away the revenues of princes, and not be the poorer" (187). In short, St. Leon's unlimited resources enable him to transcend personal concerns and to turn his hand to the welfare of others.

As St. Leon's story progresses, however, it becomes clear that his questions—"What heart can withstand the assault of princely magnificence? What man is inaccessible to a bribe?"—are more than rhetorical flourishes. His belief that money is an infallible spur to human action is tested several times over the course of the novel. His first effort to tempt someone to action fails. Suspected of murdering and robbing the mysterious stranger Zampieri—who shares the secret of alchemy on the strict condition that St. Leon will never speak of him again nor reveal his final resting place—St. Leon is thrown in jail. His distress is tempered, however, by the knowledge that "the power of money I possessed would easily unlock to me the gates of my captivity. I believed that, upon the lowest calculation, personal liberty was clearly included among the gifts of the stranger" (246). He puts his escape plan into motion by singling out a likely accomplice. He settles on "a negro [Hector]… as the subject of my pecuniary experiment" based on the assumption that he will be the easiest to bribe:

> I judged, from the meanness of his rank and his apparent poverty, that I could not have chosen better…. I was desirous that my first experiment should be a successful one. Though not unaware of the power of gold, I conceived that, among persons of middling rank and easier circumstance, there might be varieties of disposition, and I might be mistaken in my choice…. These distinctions may seem an idle and superfluous refinement, when it is considered that I had the power of raising my bribe to the level of any man's honesty or pride, be it as great as it might; and it may be thought that my offer might be so increased as to be too dazzling for mortal firmness to resist. Be that as it will,

> I am merely stating the reflections that passed through my mind, not entering into their vindication." (246)

St. Leon harbours no doubts concerning the "power of gold," but he has "no inclination … to make ostentation of more than was necessary," and he would like to buy his freedom for the cheapest price (246). He reasons that the poorer the man, the lower the price. Yet St. Leon's desire "that [his] first experiment should be a successful one" and his careful (and supposedly "idle and superfluous") selection of "the subject of [his] pecuniary experiment" imply an underlying anxiety about whether people are really so easily manipulated. As it turns out, his anxiety is justified, for Hector stoutly refuses to assist him and promptly reports St. Leon's illicit offers to his master. In portraying Hector's unpredictability as an obstacle to St. Leon's physical liberty, the novel seems to suggest that St. Leon has it exactly backwards when he contrasts his omnipotence with the principle of necessity that governs the actions of others. Here it is Hector who is free and St. Leon who is—precisely because Hector is free—trapped.[18] There is, in other words, an inverse relationship between St. Leon's freedom and that of the persons he seeks to influence. The alchemist's power is limitless only so long as the law of necessity holds true; where others exercise free will, he is unable to achieve his most basic aims. In the character of St. Leon then, Godwin imagines a fantastic figure whose power exists only insofar as the rest of the human species is driven by passion in predictable ways.

[18] The jailor who eventually sets St. Leon free makes a mockery of the whole thing. He justifies his acceptance of St. Leon's bribe by way of the doctrine of necessity: "I do not wonder… that you preferred applying to one of my servants. Their honesty must be expected to be had at a cheaper market. But, for my part, I am determined that no man shall ever pass these walls, without my being the richer. If then your escape is a thing that must happen, let us see what you can afford to give me for it" (250).

The episode with Hector is only the first of a series of failed experiments. Imprisoned once again, this time by the authority of the Holy Inquisition, St. Leon again attempts to bribe his jailor to set him free. As in Hector's case, his plan fails. The mosca (an informer for the inquisition) rejects St. Leon's proposals out of hand and reports them to his superiors, convinced that the prisoner's promises of fabulous riches constitute proof of his diabolical practices. Nonetheless St. Leon's faith in the power of wealth remains unshaken. He regards both Hector's and the mosca's refusals as singular events. The mosca is enthralled to a religious zeal that defies logic, and he is so certain of St. Leon's guilt that he is convinced that the bribe is just another one of the devil's tricks: "What I offered indeed, however dazzling in the statement, had not in fact the nature of a temptation. He to whom I addressed it gave no credit to my assertions; he thought... that my money, when possessed, would soon change its figure, and from seeming pieces of solid coin be converted into pieces of horn or of shells" (330).[19] If the mosca is not moved by St. Leon's riches, it is only because he does not believe in them. Similarly, St. Leon attributes Hector's rejection of the bribe not to an unimpeachable integrity but rather to an animal-like ingenuousness. He concludes that "the power of money was subject to limitations…. here was a poor creature utterly exempt from its operation…. It cost [Hector] no effort to be honest, and he uttered sentiments that would have given lustre to the most heroic character, without any consciousness of their greatness… Kings might have confessed their inferiority to this man. But is he to be regarded as the model of what a human creature should be wished to

[19] As demonstrated by the mosca's belief that evil can work in the service of God. See St. Leon's description of what he learned from this episode: "Such was the event of my attempt to bribe the officers of the inquisition. In my first experiment I could not even obtain a hearing; in what followed, my proposals were rejected with all the transports of religious abhorrence…. I gained nothing favorable for my situation by the trial I had made, but I added a new chapter to my knowledge of human nature. I found, that to be a knave, it was not necessary to be an infidel" (330).

be? Oh no!" (249). Hector's virtue is not virtue but ignorance, and he owes his honesty to a lack of ambition and imagination. If Hector is above bribery, it is only because his dominant passion is not greed (or as St. Leon puts it, he is "destitute of… all those exquisite sensations that most distinguish man from brute" [249]). What Hector and the mosca have in common then is that neither is ruled by ambition. If they behave in unpredictable ways, it is not because they are possessed of free will, but because they are moved by powerful and unusual motives that are not immediately visible. These two failed experiments do not undermine the principle of necessary connection. They merely show that there is some variety in the passions of men.

St. Leon's interpretation of his experiences with the mosca and Hector reaffirms Godwin's claims about the inviolability of the principle of necessity in *Political Justice.* That people sometimes behave in unexpected ways does not prove the existence of chance or free will. Echoing Hume, Godwin argues in *Political Justice* that such anomalies, like those in the material world, are anomalies in name only. All they show is that we have "see[n] only part of the premises." Thus Godwin writes:

> A philosophical experiment, which has succeeded a hundred times, may altogether fail upon the next trial. But what does the philosopher conclude from this?... Not that the connexion between effects and causes is imperfect, and that part of the effect happens from no cause at all. But that there was some other cause concerned whose operation he did not perceive, but which a fresh investigation will probably lay open to him…. The case is exactly parallel with respect to mind. (PJ, 159)

The principle of necessity in the moral world is, as Hume explains in the *Treatise,* no different from the laws that govern the natural world. Unexpected behavior does not point to chance or free will, but rather to something unknown either in the agent's character or in the forces

working upon her—or to a flaw in the experiment. Contemplating the behaviour of Hector and the mosca, St. Leon realizes that his mistake was to assume that all men are governed by ambition and greed. Human beings are universally motivated by their passions, but those passions are diverse and specific to them. In line with *Political Justice, St. Leon* describes actions that seem to contradict necessity as miscalculations concerning the passions.[20]

St. Leon's description of his encounter with Gabor suggests that he has learned, in hindsight at least, from the incidents with Hector and the mosca. When Gabor imprisons him, St. Leon initially makes the same mistake of thinking that he can bribe his way out of his shackles. Gabor quickly shatters this illusion by telling him that he cannot purchase his release at any price: "One mistake however that I see you make respecting my purposes, I will remove. You seem to suppose that, if you were to communicate to me your secret, I would then set you at liberty. No, by heavens! This cavern is your abode for ever" (403). St. Leon's reaction to his captor's seemingly inexplicable behavior is different this time around. Rather than dismiss Gabor as a religious zealot or a simpleton, he admits that his understanding of Gabor is limited. The misanthrope's character is, he reflects, "contemplative and close" (386). Indeed, when Gabor finally lays open his heart, St. Leon experiences an epiphany:

> [Gabor] opened upon me a new world. I conceived not, till now, the faintest suspicion of what had been laboring in his bosom. Amidst all my experience of the varieties of human character, this was a species that had never fallen under my observation before. What a painful and mortifying occurrence is it in human life, when we have lived with a man from day to day, when we have conversed with him familiarly, and seen him in all the

[20] That is, St. Leon's failed experiments are represented as data that help him to refine his theory of necessity, not exceptions that undermine it.

> changes of circumstance, and when we flatter ourselves we have penetrated all the recesses of his heart, suddenly to start upon something portentous that brooded there, of which to that moment we had not the lightest suspicion! I am not the only individual to whom this even has occurred. (400)

At last St. Leon understands the difference between inference based on empirical observation and knowledge—the difference between constant conjunction and true causation, Hume might say.[21] Necessary connection is an inviolable principle, but like any causal connection, it cannot be perceived directly. What St. Leon comes to realize, in other words, is that when people act unexpectedly it is an epistemological issue, a matter of not having all the facts. This is true for everyone, not just for odd cases like Hector and the mosca. People are able to surprise us because our knowledge of their characters and circumstances is always limited.

Up to this point, *St. Leon* is more or less parroting—by way of *Political Justice*—Hume on necessity. However, writing a first-person retrospective narrative rather than a philosophical treatise allows Godwin to overcome an obstacle that is, in a Humean analysis, especially challenging: our belief in our own liberty.[22] In the *Treatise,* Hume observes that whatever our thoughts concerning free will, in our everyday interactions, we treat other people as being bound by necessity. As he explains, "the same kind of reasoning runs thro' politics, war, commerce,

---

[21]See *Treatise,* 400.

[22] Hume gives three reasons for the widespread success of the doctrine of liberty: the failure to differentiate the liberty of spontaneity and the liberty of indifference; the mistaken idea that necessity undermines religion (and morality); and the false sensation of the liberty of indifference (407-09). While he is confident that he has an irrefutable reply to each of these, it is worth noting (although Hume does not) that the third poses a special problem insofar as it is grounded in sensation rather than ideas. Godwin will later argue that it is impossible (and undesirable) to overcome our "delusive sense of liberty" in "Of the Liberty of Human Actions, *Thoughts on Man, his Nature, Productions, and Discoveries* (London: Effingham Wilson, Royal Exchange, 1831), 226-43.

economy, and indeed mixes itself so entirely in human life, that 'tis impossible to act of subsist a moment without having recourse to it. A prince, who imposes a tax upon his subjects, expects their compliance. A general, who conducts an army, makes account of a certain degree of courage…"[23] In other words, we make predictions about the behaviour of others based on necessary connection—just as we do in the natural world—and these predictions determine our conduct. When it comes to ourselves, however, necessity no longer seems to fit experience. Hume notes that "we may imagine we feel a liberty within ourselves" and that "though we confess we were influenced by particular views and motives; it is difficult for us to persuade ourselves we were governed by necessity, and that it was utterly impossible for us to have acted otherwise; the idea of necessity seeming to imply something of force, and violence, and constraint, of which we are not sensible" (407).

We associate necessity with coercion, and even where we recognize that our actions are determined by motives, we still see ourselves as autonomous beings that act independently of events and people in the outside world. But Hume argues that this independence we grant only to ourselves is "a false sensation or experience," the falseness of which is borne out by the fact that "a spectator can commonly infer our actions from our motives and character; and even where he cannot, he concludes in general, that he might, were he perfectly acquainted with every circumstance of our situation and temper, and the most secret springs of our complexion and disposition" (408). If we could see ourselves from the outside—if we could be spectators of ourselves, so to speak—we would recognize the operations of necessity in the actions we feel to

[23] The quotation continues: "A merchant looks for fidelity and skill in his factor or super-cargo. A man, who gives orders for his dinner, doubts not of the obedience of his servants. In short, as nothing more nearly interests us than our own actions and those of others, the greatest part of our reasonings is employ'd in judgments concerning them" (*Treatise,* 405).

be free. For Hume, the contradiction of believing in both liberty and necessity, two completely opposite ideas, boils down to point of view. We experience our actions and those of others very differently. The former instance involves a sensation of liberty; the latter is a perception of order. In short, we believe one thing and do another because we *feel* ourselves free, but recognize others to be governed by the logic of necessity.

The fictional first-person history seems to dismantle the distinction between first-person and third-person perspective that, according to Hume, permits us to both believe and not believe in free will. In Godwin's novel, St. Leon's narrating self is a special kind of spectator with regard to his experiencing self.[24] There is no epistemological barrier here between the storyteller and the protagonist of the novel. St. Leon is, in Hume's words, "acquainted with every circumstance of [his] situation and temper, and the most secret springs of [his] complexion and disposition" (409). The narrator's third-person perspective vis-à-vis his younger self coupled with his privileged insight into his private thoughts and passions allows him to recognize himself as driven by necessity (409).[25] Thus in the diegesis, St. Leon sees his imprisonment by Gabor as a shocking change in circumstance: "What an incredible reverse was thus the creature of a moment!" (394). Reflecting upon the event from a temporal distance, however, he admits:

> Yet, truly considered, there was nothing abrupt in the reverse under which I was now suffering. The whole was a chain, every link of which was indissolubly connected from one end to the other. My attempt to rescue a people from the horrors of famine necessarily exposed me to unfavorable accidents and misconstruction. It inevitably led to

[24] I take these narratological terms from F.K. Stanzel, *A Theory of Narrative,* trans. Charlotte Goedsche (Cambridge: Cambridge University Press, 1984).
[25] This insight is made conspicuous by the fact that it is St. Leon's secret that makes society inaccessible to him. St. Leon is the only one who is privy to his secret and his only friend is himself.

> my application to the government for its aid. It could not fail to excite the alarms and jealousies of government as to the tendency of my proceedings…. When I became sensible of the precarious situation in which I stood towards the powers of the state, could I have fallen upon a more natural expedient, than the endeavour to cover myself with the shield of friendship and gratitude in the person of one of its nobles? But this expedient would almost infallibly lead to the placing myself sooner or later in the power of the man whose friendship I sought. I had done so, and this was the termination of my views and my projects! (394)

The simulataneous identity of and distance between experiencing self and narrating self allows St. Leon both to accurately know his motives and to acknowledge the inevitability of his actions and the way they fit into a "great chain" of necessity.[26] With *St. Leon,* Godwin uses the fictional autobiography to imagine how a person might overcome the delusions of sensation to fully embrace necessity.[27]

By the end of the novel, St. Leon is under no illusions regarding his lack of agency. Yet the novel stops short of fulfilling *Political Justice's* fantasy wherein everyone is a necessarian and debilitating moral concepts like blame and guilt are obsolete. St. Leon does not dismiss free will altogether. His realization that he is governed by necessity leads him to flip Hume's formulation upside down. Instead of falling into Gabor's blanket necessarianism—and the misanthropy it produces—St. Leon comes to attribute free will to others. Thus in St. Leon's eyes,

[26] This is a fiction no less far-fetched than alchemy and the elixir of life. Godwin is imagining that we can be both self and other at the same time—a perfect sympathy Adam Smith and others argued was impossible.

[27] In *Political Justice,* Godwin hopes that we will all, eventually, recognize necessity and thereby make "morality" obsolete. Part of the fiction of *St. Leon* is therefore the idea that a single narrative might be both history and romance.

the unpredictability of Gabor's actions, for example, is not finally reducible to epistemological uncertainty, but to an upheaval in his character. Gabor frequently credits his cruelty to a change in his very constitution. Although he has "no native obliquity of character," when he discovers the bodies of his kin, his "whole nature was changed in an instant" and he vows to "become like to [his] brethren" (398). This might sound similar to St. Leon's description of "a lioness bereaved of her young," yet perverted love cannot explain Gabor's retransformation. By Gabor's own account, his decision to release St. Leon is owing to a reversion to his real self: "In this solemn moment, my original character returns here (striking his heart) to take possession of its native home; a character, stern and serious, if you will; but not sanguinary, not cruel, not treacherous or unjust" (407). Gabor acts so thoroughly out of character that he forfeits it, but in his last moments, he becomes himself once again.[28]

St. Leon does not contest this explanation. On the contrary, at Gabor's "suddenly changing his character," St. Leon's "soul melt[s] within" him, and his final verdict concerning his captor is positive almost to the point of being pathological. Anxious to "do [Gabor] justice," St. Leon asserts that "I may be mistaken; but this appears to me to have been a great and admirable man. He had within him all the ingredients of sublimity; and surely the ingredients of sublimity are the materials of heroic virtue" (408). St. Leon believes that Gabor is, at heart, a "great and admirable man" and, as his descriptions make clear, that greatness has everything to do with the Hungarian's power. This power is not just potential ("ingredients"). Gabor is a model of determination and efficiency—the polar opposite of the hapless would-be philanthropist: "His courage, though cool and deliberate, almost mounted to a degree of desperate rashness; and the

[28] See *Political Justice* on the consistency of character and "some extraordinary event modifying his mind" as an explanation for a person's strange behavior (158).

fertility of his invention, and the variety of his stratagems did not fall short of his courage. The celerity of his measures was equally distinguished; distance was no bar to him; and he had no sooner conceived a project, however arduous, than it was executed" (382). Gabor seems to possess the limitless power that St. Leon boasts of at the beginning of the novel. The reason why St. Leon has so much trouble predicting Gabor's behavior is not simply that Gabor's character is "close and contemplative," but because his character does not constrain him. Gabor is not a true misanthrope, and yet he can behave like one. It is this malleability that gives Gabor the sheen of free will. This in turn gives him an ethical dimension that St. Leon lacks.[29]

My contention is that in Godwin's second novel, we see others as free but consider ourselves bound by necessity, which is also to say that virtue, along with all the other moral terms that go hand in hand with free will, belongs only to others. It is this insight—that others are ethical subjects whereas he is merely an instrument—that induces St. Leon to definitively abandon his philanthropic dreams. Thus after escaping Gabor's castle, St. Leon declares that "I had made a sufficient experiment of the philosopher's stone, and all my experiments had miscarried. My latest trials in attempting to be the benefactor of nations and mankind, not only had been themselves abortive, but contained in them shrewd indications that no similar plan could ever succeed" (413). St. Leon's supernatural powers only serve to highlight his failures and to call attention to his lack of agency. One might expect him to sink into dejection—indeed, St. Leon frequently regrets ever having acquired the philosopher's stone and the elixir of life. Yet unlike his predecessor Zampieri, who dies in silence and pain, St. Leon tells a story that ends

[29] That Gabor does not reject the doctrine of necessity upon resuming his "original character" shows that necessity does not always lead to misanthropy. In his dying moments, he tells St. Leon: "Between you and me there is a deadly antipathy; but you did not make yourself" (407-08).

in hope. In the final lines of the novel, St. Leon "exults" in the possibilities of this world, writing that "I am happy to close my eventful and somewhat melancholy story with so pleasing a termination. Whatever may have been the result of my personal experience of human life, I can never recollect the fate of Charles and Pandora without confessing with exultation, that this busy and anxious world of ours yet contains something in its stores that is worth living for" (449-50). St. Leon's joy comes from his new understanding about how the philosopher's stone can be successfully used. Earlier in the novel he tells Gabor that he "received [the philosopher's stone] as the instrument for my own happiness and the benefit of mankind" (402). What he comes to realize by the end of his narrative is that the philosopher's stone is meant for neither of these purposes. Properly speaking, it is the alchemist and not his stone that is the instrument. No freer than Godwin's famous metaphysical knife, St. Leon can only use his powers indirectly, to help some other person who *can* benefit mankind. The central question of the novel is: whom should St. Leon serve?

II

One common criticism of *St. Leon* among Godwin's contemporaries was that the "moral and philosophical purpose" of the novel—to show that limitless wealth and immortality brings misery rather than happiness—is so self-evident as to make the novel redundant. The abovementioned reviewer for the *Monthly Mirror* was more scathing still, claiming that even though there is nothing contentious in the book's intended lesson, perversely, Godwin still manages to cast doubt upon it:

> If it were the express intention of Mr. Godwin to prove that the possession of those advantages which are supposed to constitute the perfection of the *chemic art* is incompatible with the present condition of humanity, and would operate as a curse, rather

> than a blessing; that intention should have been borne out by correspondent illustration; but St. Leon's misfortunes arise, not from his possession of the philosophers' stone and the *elixir vitae*, but from circumstances totally disconnected with either…. His sufferings in the castle of Bethlem Gabor, are the consequence, not of his wealth, nor of his immortality, but of his own folly, in selecting a "demi-devil" for his friend…[30]

A successful version of the novel would establish a cause-effect relationship between St. Leon's supernatural powers and his misfortunes. As is, *St. Leon'*s consistent violation of necessary connection makes inference impossible. Thus the reviewer cites the Gabor episode as a prime example of how the novel's discontinuities undermine its purpose. The only moral to be gleaned from St. Leon's misadventure is that we should avoid selecting "demi-devils" for our friends—a lesson even less controversial than the warning against alchemy's false promises of happiness. According to the reviewer then, the novel's lack of an overarching, unifying logic means that insofar as we can draw conclusions from St. Leon's narrative, those conclusions will always be local. In subverting the principle of necessity, the novel destroys its ability to carry a message.

For the *Monthly Mirror* reviewer, the Gabor incident is a discrete narrative within a larger episodic fiction and its motto, to the extent that it has one, is particular to itself. Yet if the purpose of this chapter of the novel is indeed to demonstrate the dangers of choosing the wrong friend, it is a purpose that crops up at several key points in the novel. Each time St. Leon is incarcerated, for example, his freedom depends upon his choosing the jailor most likely to accept his bribe. Tellingly, St. Leon tends to present his palm-greasing propositions as offers of friendship. Thus his first words to the turnkey Hector are: "My good friend, are not you poor?"

[30] Unsigned review, *The Monthly Mirror* in Godwin, *St. Leon*, ed. Brewer, 477.

(246).[31] Moments like these explicitly illustrate the problem of judging character and of choosing one's friends. My larger claim, however, is that even before he is fully aware of it, St. Leon's philanthropic impulses mean that all his actions are, at bottom, decisions about whom to help (and indirectly, whom not to help).[32] St. Leon considers the philosopher's stone to be a "gift… for holy and beneficent purposes; to such it is consecrated," and each time he employs it, he is "studying the happiness" of some particular person or persons, whether it be his son Charles, the people of Hungary, or himself (402). As the possessor of the alchemical secrets, St. Leon is (to his own mind, at least) not so much an actor as an instrument to further the interests of others. How he chooses those others is, I argue, the central problem of the novel. The *Monthly Mirror* reviewer is right to identify St. Leon's mistake in "selecting a 'demi-devil' for his friend" as the take-home lesson of his dealings with Gabor, but he is wrong to say that this lesson has no application beyond this case. The question of how we should choose our friends—that is, the objects of our affection and kindnesses—is very much at the heart of *St. Leon,* and insofar as the novel has a moral purpose, that purpose is to demonstrate what counts as a good choice.

Those familiar with Godwin's philosophy will recognize this problem, for to read the novel this way is to read it as a sustained discussion of one of the most notorious passages in

[31] St. Leon's language here is telling. The entire conversation turns into a discussion of the nature of friendship. Hector insists that friendship must be grounded in moral admiration. His friends are those whom he loves, and he loves them precisely to the extent that they are "good men." St. Leon is "surprised at the propriety of [Hector's] answers" (246-49). Following this exchange, St. Leon comes to the grim realization is that: "No one was qualified to be my coadjutor, till he had proved himself unworthy of all just and honourable society. The friend I must seek, was a man whose very soul melted at a bribe" (249).

[32] Thus *St. Leon* presents a stark contrast to *Political Justice,* which criticizes the practice of philanthropy in no uncertain terms: "Philanthropy, as contradistinguished to justice, is rather an unreflecting feeling, than a rational principle. It leads to an absurd indulgence, which is frequently more injurious than beneficial even to the individual it proposes to favor. It leads to a blind partiality, inflicting calamity without remorse upon many perhaps, in order to promote the imagined interest of a few" (PJ, 143).

Political Justice.[33] In Godwin's clearest and most infamous articulation of "pure, unadulterated justice," he imagines a scenario in which the palace of the Archbishop François Fénelon catches fire "just at the moment when he was conceiving the project of his immortal [book] Telemachus" (PJ, 53). The dilemma *Political Justice* proposes is this: Fénelon and his chambermaid are both inside, but I can only rescue one of them—whom do I save? The question scarcely needs asking, says Godwin. Fénelon's life is worth more than the chambermaid's because his *Telemachus* will benefit thousands whereas her actions will touch only a few. I am therefore obliged to choose Fénelon, even if "the chambermaid had been my wife, my mother, or my benefactor" (PJ, 53-54). I must disregard the natural affection I have for my mother-the-chambermaid and assess her worth based solely on her capacity to benefit mankind, a calculation that cannot but come out in Fénelon's favor.

At first glance, the Gabor episode seems to support *Political Justice's* critique of passion. Although St. Leon initially describes his decision to befriend Gabor as a practical matter, he soon reveals that it was Gabor's passion that drew him in. He explains:

> It may seem strange that such a man as I have described should be the individual I selected out of the whole Hungarian nation to make my friend. It may seem that his qualities were better adapted to repel than attract. My choice would not appear strange, if

[33] The *Monthly Mirror* reviewer's complaint about the farfetchedness of *St. Leon* recalls Dr. Samuel Parr's objections to the Fénelon scenario in *Political Justice.* In his *Spital Sermon,* Parr writes: "Probably, if the appeal were made to the common sense and common experience of mankind, the circumstance, that they are mine, would, even in the case supposed by our philosopher, be of great consequence. But, what if a father were neither a fool nor profligate, would it *then* be of consequence that he was mine?" See Samuel Parr, *A Spital Sermon, preached at Christ Church, upon Easter Tuesday, April 15, 1800; to which are added Notes* (London: J. Mawman, 1801), 38. See Peter H. Marshall's excellent literary biography for an account of the falling out between Parr and Godwin; Marshall, *William Godwin* (New Haven: Yale University Press, 1984), 225-27.

> the reader could have conversed with him, as I did. …I could not help admiring him: his greatness excited my wonder and my reverence; and, while his manners awed and overwhelmed me, I felt an inexplicable attachment to his person still increasing in my bosom. (383-84)

As if anticipating the *Monthly Mirror's* criticism, St. Leon attempts to justify his feelings of kindness and tenderness towards a man who neither seeks them nor merits them. Yet St. Leon cannot mount a satisfactory defense of his "inexplicable attachment" to Gabor because his affection has no rational basis. In the end, St. Leon can only assure his audience that although he cannot put Gabor's allure into words, "my choice would not appear strange, if the reader could have conversed with him, as I did" (383). Here St. Leon explicitly invites the reader to exercise her imagination in order to join "the school of Bethlem Gabor" by "mingling groans"—that is, sympathizing—with him (and, by extension, with Gabor). The lesson of Gabor's treachery, however, seems to be that it is an invitation we should reject. St. Leon's mistake is to choose his friend based on passion, and we can read his story as a cautionary tale, a warning against the distortions of sentiment that will lead us to ally ourselves with the wrong people.

This critique of passion encourages us to take a more moderate view of the novel's paeans to the domestic affections, lending credence to Godwin's claims that the novel's praise of familial love does not undermine the principle of justice. In the Preface, Godwin explains that *St. Leon* "modifies" some of his views without "mak[ing] any change respecting the principle of justice, or any thing else fundamental to the system there delivered" (52). He notes that domestic attachments give us "many opportunities of conferring pleasure… without interfering with the purposes of general benevolence. Nay, by kindling his sensibility, and harmonizing his soul, they

may be expected, if he is endowed with a liberal and manly spirit, to render him more prompt in the service of strangers and the public" (52).

Passion need not get in the way of justice. In fact, it is more likely to quicken our appetite for benevolent action than to stifle it. Gone here is the sense that one might love one's kin too much. Instead, the novel anticipates Godwin's assertion in his final philosophical work, *Thoughts on Man* (1831), that "whatever besides has a tendency to illustrate and honour our nature, descends from these [the affections between parent and child], or is copied from these, grows out of them as the branches of a tree from the trunk, or is formed upon them as a model, and derives from them its shape, its character, and its soul."[34] (188). Far from limiting our benevolent sentiments, the "vast fund of love, attachment and sympathy" inherent to the parent-child relationship is the source of and blueprint for all our social (and religious) propensities.[35] Yet this resolution is no resolution at all, for it was precisely those cases where domestic attachments *do* interfere with the ends of universal benevolence that were at issue in *Political Justice.* In *St. Leon*'s Preface, by contrast, Godwin avoids this limit case, opting instead to concentrate on the "many opportunities" where personal affections and general benevolence complement one another. The Gabor episode can be read as a sort of supplement to the Preface then. It brings the Fénelon limit case back to the fore, showing that when push comes to shove, reason still trumps passion.

This reading, however, fails to acknowledge just how radical Godwin's recovery of the passions is. By Godwin's own account, it was his realization, following a rereading of Hume's *Treatise,* that human action is motivated by passion rather than reason that prompted his turn to

[34] Godwin, *Thoughts on Man,* 276.
[35] Godwin, *Thoughts on Man,* 283.

the domestic affections. As he explains in a 1798 diary entry outlining his future projects, his revisions to *Political Justice* aim to correct "the circumstance of not yielding a proper attention to the empire of feeling. The voluntary actions of men are under the direction of their feelings."[36] The trouble with this new position is that once Godwin admits, following Hume, that human behavior is dictated by feeling, he is in danger of aligning himself with egoists like Hobbes and Mandeville (or with writers like Burke who championed man's natural attachment to his "little platoon"). Hume himself escaped this fate by offsetting self-interest with sympathy. He argues in the *Treatise* "that tho' it be rare to meet with one [man], who loves any single person better than himself; yet 'tis as rare to meet with one, in whom all the kind affections, taken together, do not over-balance all the selfish" (487). According to Hume, we are animated by selfish impulses, but the sum total of our loves for those near to us—and here he is thinking mainly of family—checks those impulses. We might expect Godwin to cash in on this solution. To go along with Hume here would, however, thoroughly undercut his claim that *St. Leon* is a revision of his earlier views rather than a recantation of them. In *Political Justice*, Godwin expressly denounces family feeling (the "little platoon") as a species of self-love, a belief that there is some "magic in the pronoun 'my'" that could justify a preference grounded in something other than usefulness or merit (PJ, 54). Acting well means acting for the good of "the whole family of mankind" rather than for the good of oneself or one's kin. *St. Leon* cannot support this account of benevolent action. If feeling is the driving force of human behavior, St. Leon is hardly at fault for following

[36] Godwin, Diary of William Godwin, September 1798, in Marshall, *William Godwin*, 201. Don Locke suggests that Godwin's attitude towards the domestic affections changed as a result of the happiness he found with Mary Wollstonecraft (Don Locke, *A Fantasy of Reason: The Life and Thought of William Godwin* [London: Routledge, 1980]). Mark Philp, by contrast, argues that Godwin's reversal of opinion vis-à-vis feeling precedes his relationship with Wollstonecraft. See Philp, *Godwin's Political Justice* (London: Duckworth, 1986). Philp's book provides a thorough account of Godwin's literary influences.

his passions in choosing Gabor—indeed, the term "choosing" is misleading. Subject to the law of necessity, St. Leon could not have done otherwise.[37]

Godwin's ongoing respect for Hume's theory of motivation means that he must either come up with some way to reconcile universal benevolence and passion or forfeit the notion of disinterestedness all together. *St. Leon* fashions just such a reconciliation by imagining a species of love that is personal with regard to the object but impersonal relative to the subject. Thus in *Political Justice,* impartiality went hand in hand with pure justice whereas domestic affection implied self-interest, a belief that there is some "magic… in the pronoun 'my'" (PJ, 54). In *St. Leon,* Godwin turns the logic of *Political Justice* upside down by detaching what Godwin terms "self-oblivion" from general benevolence and aligning it with personal feeling. Indeed, under Godwin's new moral system, self-love simply does not exist. Conscious of our own lack of agency, we see that our actions deserve neither praise nor blame, and we can no more love ourselves than we could love a knife. In *St. Leon,* only other people are free and, as a result, only they are candidates for moral approbation and the love it gives rise to.

The theory of action I have described above enables Godwin to imagine a wholly new way to reconcile personal feeling with impartiality. The novel brings together the partiality of

[37] Indeed, the novel suggests that the source of conflict between Gabor and St. Leon is more the weakness of St. Leon's affection than its excess. The breaking point for Gabor is St. Leon's decision to extend his charity to Gabor's sworn enemy. Gabor reproaches St. Leon for his indiscriminate kindness: "There was one native of Hungary between whom and me there subsisted an open and eternal war…. With a refinement of cruelty and insult difficult to conceive, you chose that man for one of the objects of your beneficence. Would I consent to see my name joined in pension list with my mortal enemy?" (399). Similarly, looking back, St. Leon reflects that for all their "mingled groans," he and Gabor were never really friends: "I called Bethlem Gabor my friend; I persuaded myself that I had cogent reasons for calling him so. But there was little sympathy between us; he was wrapped up in his own contemplations; he was withered by his own calamities; our souls scarcely touched in a single point. No, no; this is not friendship" (424).

private affections and the disinterestedness of universal benevolence by suggesting that we might genuinely love those who are morally deserving of our love with all the ardour usually reserved for our nearest and dearest. In other words, whereas *Political Justice* asks me to choose Fénelon *despite* my private affections, *St. Leon* imagines that I choose Fénelon *because* of them. Godwin invites us to consider the possibility that I really can love Fénelon more than my mother. Thus in *St. Leon*, we do not automatically love those to whom we are attached through familial or social bonds. It is the other way around. Love precedes interpersonal bonds and it is love that creates individual attachments. This love is grounded in moral admiration and is, for precisely that reason, at once necessary and justified.

III

That St. Leon is continuously frustrated in his attempts to be the benefactor of mankind seems to be proof that universal benevolence is out of human reach. Domestic charity, on the other hand, is very much within our grasp. St. Leon's single successful action is to engineer his son's wished-for marriage to Pandora. The only obstacle to their union is Pandora's poverty—a problem to which St. Leon's gift is a perfect solution. Initially it seems that St. Leon's plan to dedicate himself to Charles's happiness will be every bit as unsuccessful as his other schemes. Enjoying Pandora's company a little too much, he nearly comes between the two lovers as Charles begins to imagine that D'Aubigny (St. Leon's assumed name) has won Pandora's affections. All ends well, however, for although the novel closes with Charles disowning his father for (unbeknownst to him) a second and final time, St. Leon can nonetheless take credit for his son's marital happiness. It is thanks to St. Leon's "beneficent fraud" that Charles and Pandora are "finally made happy in each other" and St. Leon rejoices in having been "the fortunate means of supplying to this excellent and incomparable creature [Pandora] the only

defect under which she laboured, a want of fortune" (448).[38] St. Leon has finally succeeded in producing happiness albeit on a small scale. So although he was originally convinced that he owed his failures to a lack of resolution and narrow views ("I had blamed myself in review, that, in my experiments at Constance, at Dresden, at Pisa and at Madrid, I had not commenced upon a sufficient ample scale, but had suffered myself to be frustrated by the ingloriousness of my precautions" [371]), St. Leon comes to realize that the reason for his disappointed efforts was quite the opposite: the philosopher's stone only works on a small scale. Earlier in the novel this smallness would have been a disappointment, as his goal was to be a "servant of the whole world," but by the end, it is enough for St. Leon to be a servant to his son.

It is easy to see why *St. Leon* poses a problem for universal benevolence. If St. Leon has dismissed the importance of his personal experience, his concerns remain personal insofar as they are directed towards those with whom he has personal ties. As I have shown above, Godwin's moral system would declare St. Leon's newfound benevolence to be self-interested. How then is Godwin able to maintain that the novel is a reconciliation project? How can St. Leon's matchmaking be a form of universal benevolence? Godwin manages it, but he does so by way of a singular scenario. In Charles, Godwin invents a character who is Fénelon and mother bundled into one, so to speak. Charles is St. Leon's son, but he is also a public figure whose actions reach beyond his intimate circle. St. Leon tells us that "The chevalier de Damville is considered in that country as the great bulwark of the Christian frontier" (448). Likewise

[38] There is a hint in the text that St. Leon may have committed a murder to bring about this happy conclusion. He reports that Cabriera, his front man in this beneficient fraud, dies at an opportune moment for all involved: "Cabriera, worn out with years and fatigues, but still grasping and avaricious to his latest hour, expired in my arms in the city of Triestel and by his death yielded me this contentment, that henceforth the only obvious means for detecting my beneficent fraud in securing the dower of Pandora was for ever removed" (448).

Charles's description of his doings encourages us to think of his actions as having universal consequences. He proclaims that "I shall never think I have lived in vain, after having contributed, in however humble a place, to arrest the career of insolence and impiety which, under the standard of the crescent, threatened to overrun the whole Christian world" (423). Charles is, in short, an epic hero, and in serving him, St. Leon is serving someone whose heroic actions have large-scale consequences.[39]

In the terms set out in *Political Justice,* Fénelon is the author of the *Telemachus* as well as a family member—so this solution answers the problem of the incompatibility of universal benevolence and domestic affection. Yet it is an unsatisfactory conclusion in much the same way as the Preface's assertion that "the man who lives in the midst of domestic relations will have many opportunities of conferring pleasure… without interfering with the purposes of general benevolence" (52). That is, it is an exceptional, even artificial scenario that draws attention away from the problem. Yet it is exactly this farfetchedness that forces us, as Rajan and Charise point out, to take Godwin seriously.[40] Universal benevolence requires a particular kind of relationship, and *St. Leon,* free from the constraints of realism, illustrates what this might look like.

---

[39] Charles's excoriation of his alter-ego, Chatillon, makes St. Leon more resolved than ever to keep his identity secret and to abandon the fame and gratitude he once hoped for. He writes that Charles's speech "had the effect of making me regard with a more complacent satisfaction the plan I had formed of devoting myself to his service. Here I pursued no delusive meteor of fame; the very essence of my project lay in its obscurity. Kings and prelates, armies and churches, would no longer find an interest in disputing about my measures; I should indulge the secret promptings of my soul, undisturbed alike by the censure of the world, and its applause. It was thus that, under every change of fortune, I continued to soothe my soul with delusive dreams" (415-16).

[40] Discussing the third edition of *Political Justice,* Marshall argues that Godwin uses the blood relation as a necessary stepping-stone towards genuine disinterestedness: "under Hume's influence… [Godwin] added at the end of the volume that it is only 'after having habituated ourselves to promote happiness of our child, our family, our country or our species [that] we are at length brought to approve and desire their happiness without retrospect to ourselves" (Marshall, *William Godwin,* 160). As in Godwin's Preface to *St. Leon,* furthering the happiness

The reunion of St. Leon and his son at the end of the novel suggests a sort of fatality and permanence to their connection. Charles rejects his father only for their relationship to take, many years later, another form: that of siblings. Charles tells D'Aubigny—St. Leon's young alter-ego—that he is "cut off for ever from those of my own lineage and blood. It is with inexpressible delight that I thus cheat my fate, and hold you to my bosom as if you were indeed my brother. I would not part with the fiction for the mines of Peru; and I know not whether I do not cultivate it more assiduously, and regard it with a sentiment of more anxiety and seal, because it is a fiction, than I should if it were a reality" (417). The irony is, of course, that the fiction *is* a reality and that it is not Charles who is cheating fate, but fate that is pulling the wool over Charles's eyes. Likewise, Charles does not realize how close he comes to the truth when he accuses D'Aubigny, now unmasked as the infamous alchemist and friend to the "Turkish infidels," of his father's faults. Charles declares: "How can I tell that the fraternal resemblance borne by your features to my own, and the sudden and ardent partiality that rose in my breast when I first saw you, have not been produced by the most detested arts? …This is to me the most painful of all subjects. I had a father whom I affectionately loved: he became the dupe of these infernal secrets" (446). Neither the fraternal resemblance nor the "sudden and ardent partiality"

---

of others begins with domestic charity. However, Marshall's discussion calls attention to something that the Preface does not argue, namely that the expanding circle of sympathy is not simply about developing a penchant for doing good (which will make us more likely to help strangers). Rather, it is a habit that makes us empty our relations of their personal nature. According to Godwin, once we pick up the habit of benevolence, we will be able to act in a moral way without thinking about who will benefit from our actions. Domestic benevolence does not, in other words, simply expand. Domestic benevolence becomes as impersonal as universal benevolence in that there are individuals, but no personal relations between them. It is this lack of personality that allows us to benefit others in a way that is congruent with universal benevolence. In this way, Godwin makes universal benevolence concrete without it becoming self-interested.

Charles experiences is a magic trick. Indeed, St. Leon feels the same way. His instant recognition that Damville is his son has something supernatural to it. He explains: "It may seem extraordinary that I should instantly have known him…. It was fifteen years since I had beheld [Charles's countenance]; he was then scarcely above half his present age, a mere stripling… he was now a leader of warlike bands, his complexion olived over with service, and his eye rendered steady with observation and thought. But I knew him; I knew him in a moment. My soul, with the rapidity of lightning, told me who he was" (411). The idea that blood will recognize blood was a familiar trope in the period. As Ruth Perry has argued, at a time when family was losing currency as the privileged foundation for society, "the *cri du sang* or the call of blood, signified a fictional instinct whose popularity apparently reassured society that consanguinity was still powerful."[41] Father and son are attracted to each other through an irrepressible bond of blood. *St. Leon* suggests that kinship cannot be disclaimed and that the fates of family members are intertwined.

Yet at least one reader discerned in *St. Leon* an extremely anti-social bent. As scathing as always, the reviewer for the *Monthly Mirror* insisted that Godwin's novel offered only a "*mock*-SENSIBILITY" that undermined both the rationalism (or "extreme apathy") of *Political Justice* as well as the domestic affections extolled in the novel's Preface. He writes:

> … we are not sure that Mr. Godwin's *Political Justice* is not preferable to his *scheme of morality;* for it is better to sacrifice individual connexions to the public good, than foster a *mock*-SENSIBILITY, which is to be considered as paramount to all the ties of affection and consanguinity. St. Leon's is a family of love, and not a member of it but overflows with sensibility: but mark the disunion of these "*commingling souls."* The

[41] Perry, 95.

> mother dies, with a solemn injunction that St. Leon should *snap the cord that binds him to his children;* and St. Leon adopts the resolution, instantly, though he has nothing to do, nor any where to go; but he feels himself *alone in the world,* and, *therefore,* he can do no less than abandon his female offspring to their fate. The young ladies, it must be confessed, part with their parents with equal *sang froid,* notwithstanding their *excessive susceptibility,* and Charles, who is the very *soul* of affection and honour, gives up father, mother, sisters, and all, because, forsooth, a French coxcomb had thought proper to be impertinent in a tennis-court. If such are the *mysterious sympathies of our nature,* heaven forbid that they should ever be unravelled.[42]

According to the *Monthly Mirror* reviewer, *St. Leon'*s paeans to sentiment and domestic affection are hypocritical. In truth, St. Leon's family of love is a group of individuals who are invested in neither the public good nor the welfare of their kin. Thus St. Leon romanticizes his solitude, which gives him license to "abandon his female offspring to their fate," and Charles disowns his relations in order to protect his honour, which has been besmirched by a youth whose insolence "is a subject of general remark" (208).[43] There are, in short, no true interpersonal connections in *St. Leon.* Even the most selfless acts are merely show—St. Leon admits that "there are persons who will regard this passage [cutting himself off from his daughters] in my history as culpable, and the testimony of a cold and unsusceptible heart. I contemplate it, even at this distance of time, as the noblest and most virtuous effort of my life"

[42] Unsigned review, *The Monthly Mirror* in Godwin, *St. Leon*, ed. Brewer, 477.
[43] See G.J. Barker-Benfield for a succinct account of how sentiment came to be associated with excess and insincerity. Barker-Benfield, *The Culture of Sensibility: Sex and Society in Eighteenth-Century Britain* (Chicago: Chicago University Press, 1992).

(302). For the reviewer, St. Leon's self-congratulation is hypocrisy.[44] Godwin's novel discards universal benevolence and domestic affection in a single gesture. What remains is simply the individual who is alone and self-interested, connected to and loving no one. All interpersonal connections are dissolved and there is no benevolence, no fellow-feeling left at all.

Yet *St. Leon* itself challenges the above reading in a number of important ways. The novel ends with Charles's story, which culminates, as per the conventional domestic novel, in marriage. That it is the union of husband and wife—"the most accomplished couple…"—that wins out over what would have been St. Leon's happy ending—the union between father and son and, in its reconfiguration, brother and brother—shows the novel's commitment to personal connections based on something other than familial love. The nature of this other bond of affection is outlined in Charles's differentiation between the love he bears his father and that he bears his mother. He tells D'Aubigny that "I loved my father: I loved him because he was my father; I had great obligations to him; he once had virtues. But my mother,—if I could have found her in the wildest desert of Africa, and have known her virtues, a stranger to my blood, descended from the remotest tribe of the human race, I should have chosen her for my friend, my preceptress and my guide, beyond all that youth and beauty, with their most radiant charms,

[44]The reviewer is making a not-so-thinly-veiled reference to Jean-Jacques Rousseau when he dubs Godwin's style "*mock*-SENSIBILITY." Godwin more or less invites the comparison. St. Leon's justification for abandoning his children echoes Rousseau's defense in the *Confessions:* "by abandoning my children to public education for lack of power to bring them up myself… I believed I was performing an action of a Citizen and a father…. More than once since then, the regrets of my heart have taught me that I had deceived myself, but my reason has been far from giving me the same admonition; I have often thanked Heaven for having protected them from their father's fate" in Jean-Jacques Rousseau, *The Confessions and Correspondence, including the Letters to Malesherbes*, trans. Christopher Kelly (London: University Press of New England for Dartmouth College, 1995), 299. Rousseau was roundly chastised for leaving all five of his children at foundling homes. Given his treatment of his own offspring, many found his decision to publish an educational tract (*Émile)* surprising, to say the least.

could tender to my acceptance!" (417-18). Charles's love for his father is a familial love only whereas his love for his mother is more deeply rooted. Even if she were as foreign as foreign can be—"a stranger to his blood descended from the remotest tribe of the human race"—Charles would love her as if she were his mother and wife rolled into one. Her virtue would have made her dear to him, so much so that she would become everything to him: friend, mentor, and implicitly, a wife who would instantly command both his respect and his love. His admiration for Marguerite de Damville's virtue would have tied him to her with bonds stronger than those of blood.

St. Leon comes around to Charles's way of thinking. Despite his claim that he is the engineer of Charles's good fortune, he eventually admits that he is a mere cog in a machine that was already in motion toward that end. Wrapping up Charles's history, he thinks upon his role in the adventure:

> This last trial, to which I had been the undesigning means of subjecting him, was none of the least arduous. Love often entails imbecility on the noblest of mankind: but Charles surmounted the most perilous attacks of this all-conquering passion. When he thought Pandora unworthy, he tore himself from her, and would not admit a struggle…. I was the hero's father!—but no! I am not blinded by paternal partiality;—but no! he was indeed what I thought him, as near the climax of dignity and virtue as the frailty of our nature will admit. His virtue was at length crowned with the most enviable reward the earth has to boast,—the faithful attachment of a noble-minded and accomplished woman." (449)

Proud of his son, St. Leon tries to take the credit for Charles's virtue "I was the hero's father!" only to realize in the next sentence that a move like that would automatically nullify his claims in the eyes of the reader. St. Leon can only convince the reader of his son's bravery by breaking the

bond between them. Charles stands alone, and the best evidence of his virtue is his ability to limit his love to those who deserve it.

The relationship between St. Leon and his son is mirrored in the relationship between Charles and his mentor, the general Gian-Battista Castaldo. There the connection is portrayed as the mutual attraction of virtue, and if Castaldo becomes a father to Charles, it is because the two share a moral affinity. Charles becomes an instrument for the Christian cause, whose merit and necessity is proven by Castaldo's honour and that of his followers. Charles exclaims: "'And shall I,' added Charles, in a sally of generous enthusiasm, 'ever desert a cause which has been thus honoured?... No: such virtue as I have described never could have been conceived, but in the bosom of truth! Great as is the pious devotion I feel for [that] God… I own my weakness, if it be a weakness, his cause is scarecely less endeared to me by the sublime exertions of his heroic followers, than by his own adorable condescension and mercy'" (421). Charles takes up the Christian cause because he believes in the virtue of Castaldo and, by extension, Castaldo's followers. Castaldo is Charles's Fénelon, and Charles is St. Leon's, such that even though St. Leon despises the wretchedness of war and recognizes Charles's prejudice, heroism alone becomes worthy of reward.[45] Thus while St. Leon invests everything in Charles, Charles lives to serve Castaldo and, more broadly, the Christian cause. In other words, everyone is a philanthropist in relation to someone. Recognizing our lack of agency, all we can do is help other agents who deserve our help.

[45] St. Leon admits that "though I could not entirely enter into this sentiment of [Charles's], and indeed regarded it as an infatuation and delusion, I did not the less admire the grandeur of soul with which this heroic fable inspired him. There was no present propensity in my heart that led me to delight in deeds of blood and war;… but… I was sensible to the lustre which military zeal cast round the character of my son. Nor is this incredible or absurd; the qualities of a generous and enterprising champion are truly admirable, though the direction they have received should be worthy of eternal regret" (424).

IV

St. Leon helps us to see, then, that for Godwin, every action is an act of philanthropy, which is also to say that every action is a deferral. This sense of incompleteness is already present in *Political Justice*. Because discussions of the Fénelon problem have tended to revolve around the attack on the domestic affections, the problem itself has not been examined in close detail. As a result, critics have overlooked the fact that the scenario as Godwin describes it relies upon certain logical inconsistencies. Frances Ferguson points out that Godwin's is like other "garden-variety accounts of utilitarianism" in that:

> … they feel obliged to make individuals equipped with calculative reason take on board too much, to imagine scenarios in which they might face the possibility of rescuing one person or ten, saving the writer and educator Fénelon or his servant (in William Godwin's example), deciding for a more diffusive influence and against a less diffusive one. That is, the examples all build in an open-endedness that warps the description of the choices to be contemplated, because the examples revolve around a series of conjectures about the future that only have provisional stopping points, since the greatest good of the greatest number is always at risk of being replaced with another account of the quantity of good and a larger number of persons who benefit from it.[46]

Ferguson's point is not simply that Godwin's scenario requires an impossible knowledge of the future, but that there is no ultimate future to know. That is, there is no moment where one can say that an action has achieved the greatest good for the greatest number because the number and

[46] Frances Ferguson, "The Sublime and Education: Educational Rationalization/ Sublime Reason" *Romantic Circles*, October 25, 2016, http://www.rc.umd.edu/praxis/sublime_education/ferguson/ferguson.html. Ferguson also makes an interesting point about how the Fénelon problem is a state of emergency which makes calculation difficult.

the good are always in flux. What Ferguson's argument shows, however, is that virtue is not simply a conceptual problem (envisioning the whole and the best course of action), but an ontological one. The object of benevolence is itself unstable. There is no permanent greatest number and, by extension, there is no way to produce an action that will produce the greatest good in anything more than momentary way.[47]

Ferguson's discussion is helpful in that it reveals the infinite deferral of actions implicit in many versions of utilitarianism, including Godwin's. She explains that under such theories, "the decisions that one made, that is, never really could achieve the status of independent actions. They were, instead, merely gestures that assisted other, more remote actions."[48] Because there is no once-and-for-all future where the greatest number and greatest good is established, actions are always merely steps toward other better, more accurate, more benevolent actions. This seems right, yet it seems important that in the Fénelon example, we are choosing between two people rather than two actions. My action leads to another future action, but that action belongs to Fénelon and not to me. In other words, Fénelon's writing of the *Telemachus* is portrayed as a final action, and my rescue of him is a means to that end. This is significant firstly because it reintroduces the element of desire. Speaking about persons means speaking about partiality and

[47]Ferguson's argument is similar to Godwin's in "Of the Connexion between Understanding and Virtue," where he states that "vicious conduct is always the result of narrow views" (PJ, 142). Vice stems from error, a mistaken idea of what counts as "the whole" and how best to serve the whole. Godwin outlines the sheer scale of the understanding that is a prerequisite for choosing and performing virtuous actions: "It must begin with a collective idea of the human species. It must discriminate, among all the different causes that produce a pleasurable state of mind, that which produces the most exquisite and durable pleasure…. God, according to the ideas usually conceived of that being, is more benevolent than man, because he has a constant and clear perception of the nature of that end which his providence pursues" (PJ, 140).
For Godwin, virtue is an epistemological problem. Much of our bad behavior can be attributed to our necessarily limited understandings, "the present imperfection of human nature [that] may serve as an apology for my error, but can never turn error into truth" (55).
[48] Ferguson, "The Sublime and Education."

emotion in a way that is more central than what we do in comparing actions. Secondly, and more importantly, it is an act of delegation. In choosing Fénelon, I am choosing someone who will perform an action that I recognize as good, but which I am myself not equipped to perform. Godwin writes:

> I am not to be blamed, if I employ a builder to construct me a house, or a mechanic to sink me a well; nor should I be liable to blame, if I worked in person under their direction. In this case, not having opportunity or ability to acquire the science myself, I trust to the science of another. I choose from the deliberation of my own judgment the end to be pursued; I am convinced that the end is good and commendable; and, having done this, I commit the selection of means to a person whose qualifications are superior to my own. The confidence reposed in this instance is precisely of the nature of delegation in general. No term surely can be more unapt than that of obedience, to express our duty towards the overseer we have appointed in our affairs. (PJ, 98)

Of the three forms of authority Godwin recognizes, confidence in another person more expert than myself occupies a middle ground between private judgment and coercion. I cannot know every science, therefore it makes sense that I might sometimes call on others to help me reach an end that I have decided is good but which I do not know how to achieve.

Godwin is quick to add, however, that delegation should be kept to a minimum, for like coercion, confidence dispenses with individual understanding. Reposing confidence in another means suspending the exercise of private judgment, relinquishing the responsibility of figuring out how to best serve the public good (therefore only doing half the work that goes into a benevolent action, as I show above). It is, as a result, equally dangerous, for "Man, when he surrenders his reason, and becomes the partisan of implicit faith and passive obedience, is the

most mischievous of all animals…. He is, in the instant of submission, the blind instrument of every nefarious purpose of his principal; and, when left to himself, is open to the seduction of injustice, cruelty, and profligacy" (PJ, 99). Given Godwin's criticism of delegation and his commitment to the idea that virtuous men must use their own reason both to determine right ends and to figure out how to reach those ends, the Fénelon dilemma seems strange. Why does Godwin's archetypal example of the exercise of judgment involve an abdication of that judgment in one and the same gesture?

To answer this question, we must first recognize that the delegation implicit in the Fénelon problem—and explicit in *St. Leon*—is even more problematic, from a Godwinian perspective, than the delegation Godwin focuses on in his discussion of authority. There Godwin portrays delegation as a matter of using someone else's skills as a means to an already defined end. I have decided that I want to build a certain kind of house in a certain place and I engage a person who builds houses to fulfill my plans. Thus Godwin writes that "delegation is not… the act of one man committing to another a function, which strictly speaking it became him to exercise for himself…. The individuals to whom the delegation is made, are either more likely from talents or leisure to perform the function in the most eligible manner, or at least there is some public interest requiring that it should be performed by one or a few persons, rather than by every individual for himself" (PJ, 92). Where delegation is in line with justice, it is done because someone else is better able to perform the action, to realize it "in the most eligible manner." Godwin studiously ignores the other kind of delegation—that in which I defer to another not only on the subject of how to do something, but of what to do. Critic Evan Radcliffe points out that this is omission leads to the conflation of decisions and actions. He explains that there is a difference between:

> persons fashioning actions... and persons evaluating or choosing among actions that have already been performed or devised…. Fénelon's potential rescuer is only a chooser faced with two clearly defined actions…. But Fénelon himself figures as a fashioner of actions, since in writing a book one does not choose among books already written. Fashioning an action involves someone's particular traits and thus cannot be detached from his or her particular character. But choosing an action can be more impersonal. Godwin ignores fashioning and the importance of character to it, in favor of choosing and evaluating.[49]

There is a qualitative difference between my action (rescuing Fénelon) and Fénelon's (writing the *Telemachus).* In the first instance, I am selecting from a set of options, in the second, Fénelon is creating a new option. Making this distinction allows us to see that the delegation in the Fénelon example inverts the kind of delegation Godwin describes elsewhere. In the latter case, I delegate the how—I use the skills of others to realize my intention. In the former, I have become Fénelon's instrument. My action enables his, and I am the means for his being able to benefit the whole.

For Radcliffe, Godwin's failure to differentiate fashioned actions from decision-actions means that he overlooks the importance of personal character—and the individual agent—to action. Indeed, by focusing on choosing actions, Godwin throws out the notion of creativity and originality all together. These two elements seem to play no part in his theory of action. However Godwin might counter that this is exactly his point. His utilitarianism is founded on the premise that persons are passive instruments rather than agents. Thus in his discussion of free will, he famously compares a man to a knife:

---

[49] Evan Radcliffe, "From 'Metaphysician' to 'Novelist': 'Political Justice', 'Caleb Williams', and the Tension between Philosophical Argument and Narrative." *Modern Philology* 97, no.2 (2000): 538.

> Virtue, if we would speak accurately, ought to be considered by us in the first instance objectively, rather than as modifying any particular beings. It is a system of general advantage, in their aptitude or inaptitude to which lies the value or worthlessness of all particular existences…. [But] all approbation or preference is relative to utility or general good. A knife is as capable as a man of being employed in the purposes of virtue, and the one is no more free than the other as to its employment. The mode in which a knife is made subservient to these purposes is by material impulse. The mode in which a man is made subservient is by inducement and persuasion. But both are equally the affair of necessity. (PJ, 167)

Human beings are merely functions in "a system of general advantage." They are not virtuous in the sense of possessing virtuous traits—courage, prudence, humility, etc. Rather their virtue lies in their utility, the degree to which they facilitate the operations of the system that generates general happiness. Insofar as we are choosers of persons, we are instruments that facilitate the actions of those persons.

V

St. Leon thus shows us how human beings might be at once instruments--merely choosers of persons—and individual characters, that is, moral agents like Fénelon. In other words, the impossible fiction of the philosopher's stone and the elixir of life allows us to imagine what it might really mean to think of people as knives. Granted immortality, St. Leon forfeits a stable identity. With the elixir of life, he is St. Leon, Chatillon, and D'Aubigny all in one. His rejuvenations cut his ties not just to the world, but to his sensing and feeling self: "From this hour I had no passions, no interests, no affections; my heart has never expanded with one natural emotion" (349). What St. Leon learns is that in the absence of a sensing and acting self, it is the

other that becomes important. His lesson is that his personal experience is nothing and that all he can do is devote himself to another. He writes: "I entered into a solemn engagement with myself that I would forget and trample on every personal concern, and be the victim of sacrifice, if need were, of the happiness of my child" (413). This embrace of self-oblivion in the novel helps us to understand why Godwin might have chosen Fénelon of all people to illustrate his famous claim about "pure, unadulterated justice." Fénelon was a proponent of pure love—a love for God so complete that one should wish one's soul in hell if He wishes it.[50] There is no sacrifice too much. The disinterestedness of morality entails loving something so much that the self does not exist. Individual affections exist, but without the self. And it is in that sense that love can be impersonal while still being an individual connection between two beings. Love is a tie between two unequals—and this is exactly the point. One of them is an agent and the other is his servant. Justice is what other people do. All we can do is serve them in our modest way.

*St. Leon'*s representation of liberty and love anticipates many of the ideas Godwin would come to embrace in his later years. In the essay "Of Love and Friendship" in *Thoughts on Man,* Godwin argues that "the great secret in moral nature is that love is between unequals."[51] That inequality has its roots in a belief in the virtue of the other, which is a prerequisite for love. We are partial to particular people because we judge them to be virtuous, and we think them virtuous because we believe them to be free. "Belief" is very much the point here, for Godwin makes it clear that there is no such thing as rational partiality. He explains that "there can be no passion,

---

[50] See François Fénelon, *A Dissertation on Pure Love* (Dublin: Isaac Jackson, 1739) and "Maxims of the Saints," in *Fénelon: Selected Writings,* ed. Chad Helms (New York: Paulist Press, 2006), 207-97. For a good overview of Fénelon's concept of pure love, see Robert Merrihew Adams, "Pure Love," *The Journal of Religious Ethics* 8, no.1 (1980): 83-99.
[51] Godwin, *Thoughts on Man,* 187. See Alex Gold Jr. for a discussion of asymmetrical love in *Caleb Williams*. In Gold, "It's Only Love: The Politics of Passion in Godwin's *Caleb Williams." Texas Studies in Literature and Language* 19, no. 2 (1977): 135-60.

and by consequence no love, where there is not imagination. In cases where every thing is understood, and measured, and reduced to rule, love is out of the question. Whenever this sentiment prevails, I must have my attention fixed more on the absent than the present, more upon what I do not see than on what I do see. My thoughts will be taken up with the future or the past, with what is to come or what has been. Of the present there is necessarily no image. Sentiment is nothing, till you have arrived at a mystery and a veil…"[52] Love requires imagination, an attention to the absent, the future and the past rather than the present. It follows that if love produces partiality, which is the basis of all human action, imagination is the starting point for moral action. Read under this light, *St. Leon'*s spectacular failure to present "things as they are" can be seen as proof of Godwin's commitment to love and the fiction of liberty it depends upon.

Although some critics have argued that the open-endedness of the novel's ending is merely arbitrary, St. Leon's dismissal of his "personal experience of human life" is an invitation to think about what kind of story he is offering here. St. Leon finishes his narrative with his son's biography. By his own admission, this move is intended to create a sympathetic bond between the reader and himself. He writes: "That the reader may enter the more fully into my sentiment of congratulation upon the happiness of my son, and rise from the perusal of my narrative with a more soothing and pleasurable sensation, I will here shortly recapitulate the good qualities that had been unfolded in this truly extraordinary young man from his earliest infancy" (448-49). St. Leon insists that his narrative is not a tragedy, and if we read it like that, it is only because we are confused about who the protagonist really is. St. Leon's personal experience is sad, but his is not the tale that matters. Instead St. Leon encourages us to focus on Charles's story and its happy

[52] Godwin, *Thoughts on Man,* 187.

ending. The novel finishes on a high note because St. Leon is able to share his insight that stories only really matter when told from the third person. In turning the Humean account of necessity on its head, the novel imagines a world in which first-person experience is reduced to mechanics. Because we are sensitive to our own lack of agency, we can only see our personal experiences as one long series of events—a chronological record rather than a life. It is only when we talk about other persons, persons who feel and act, that we can talk about heroes and villains.

Coda

"Ourself… is in reality nothing": Character, Self, and the Limits of Realism

Near the end of Jane Austen's *Persuasion* (1818), Anne Elliot expresses a certain satisfaction that her mistrust of Mr. Elliot—her cousin, but also an unwanted suitor—is justified. When she learns of his "ingratitude and inhumanity" from her old friend Mrs. Smith, Anne rejoices that "there was no longer anything of tenderness due to him. He stood as opposed to Captain Wentworth, in all his own unwelcome obtrusiveness; and the evil of his attentions last night, the irremediable mischief he might have done, was considered with sensations unqualified, unperplexed. Pity for him was all over."[1] Anne's trademark "sweetness of character" fails her here, and if we excuse her for feeling relief rather than sadness at the news that her kinsman is not so good as he seems, it is because this feeling has its source in her affection for Captain Wentworth. For if truth be told, up until this point, Anne's misgivings concerning Mr. Elliot's character have been baseless. Contemplating the reasons why she could never marry him, she complains that:

> Though they had now been acquainted a month, she could not be satisfied that she really knew [Mr. Elliot's] character. That he was a sensible man, an agreeable man, —that he talked well, professed good opinions, seemed to judge properly and as a man of principle, —this was all clear enough. He certainly knew what was right, nor could she fix on any one article of moral duty evidently transgressed; but yet she would have been afraid to answer for his conduct. She distrusted the past, if not the present. … though he might

[1] Jane Austen, *Persuasion* (London: J.M. Dent, 1895), 218. Hereafter cited parenthetically in the text.

> now think very differently, who could answer for the true sentiments of a clever, cautious man, grown old enough to appreciate a fair character? How could it ever be ascertained that his mind was truly cleansed? (164)

Anne's final indictment of Mr. Elliot as "a clever, cautious man" who has come to value "a fair character" borders on the perverse. These traits are not obviously faults. After all, Anne generally promotes the virtues of prudence and caution; moreover, her wariness towards Mr. Elliot is itself a kind of cautiousness. Elsewhere, Anne claims that her "idea of good company… is the company of clever, well-informed people," and although the novel ridicules Sir Walter's preoccupation with rank and name, there is no doubt that the baronet's excessive attentiveness to reputation is still preferable to the indifference and carelessness shown by the young Mr. Elliot. Surely if this is the worst that can be said of Mr. Elliot, he cannot be so very bad.

That Anne turns out to be right about Mr. Elliot should not distract us from the faultiness of her reasoning. Indeed, although the discovery of Mr. Elliot's past seems like a vindication of Anne's suspicions and a testament to her perceptiveness, the novel resists this reading. Seeing Mr. Elliot for the first time since hearing Mrs. Smith's story, Anne experiences a visceral reaction of revulsion: "She had been used before to feel that he could not be always quite sincere, but now she saw insincerity in everything. …she could hardly bear the sight of his present smiles and mildness, or the sound of his artificial good sentiments" (220). Anne pointedly ignores Mrs. Smith's reassurances that Mr. Elliot "is no hypocrite now. His present attentions to your family are very sincere, quite from the heart" (210). To her eyes and ears, Mr. Elliot's every word, every gesture is a lie. Instead of enabling her to see his "true sentiments," the truth addles Anne's senses so completely that she sees the person in front of her not as a person, but as an instantiation of her idea of him. In other words, the effect of Mrs. Smith's revelation on Anne is

not to lay bare Mr. Elliot's character, but to conceal it. Susan Morgan has explored the logic of this seeming paradox. She explains that:

> The problem with revelations is that they contradict what has previously been thought. Revelations also have a quality of finality. They don't allow for the possibility of getting to know someone. Revelations are endings. They deny time. Alternative forms of perceiving and presenting character include recognizing that character is mixed, that perception is based in part on the emotional perspective of the viewer and that people have the power to change. Understanding is not the revelation of a moment, nor is truth a matter of facts. People are more interesting than the boxes we put them in.[2]

To use Mr. Elliot's checkered past as a key to understanding him in the present is to buy into the idea that character is internally coherent, that character never changes, and that when someone's behavior takes us by surprise, it is only because we were mistaken about what kind of person she *really* is (but now we know better). Given this attitude, Anne's question—"How could it ever be ascertained that his mind was truly cleansed?"—cannot but be disingenuous. Mrs. Smith has scarcely begun her story when Anne decides that she has heard enough. The notion that Mr. Elliot "might now think very differently" is banished from her mind. According to Anne's rather bleak view of human nature, every action a person performs is either a faithful disclosure of her real (and so unchanging) character or hypocrisy, pure and simple.[3] Moral reformation is

---

[2] Susan Morgan, "The Nature of Character in *Persuasion"* in *Jane Austen's Persuasion*, ed. Harold Bloom (Philadelphia: Chelsea House, 2004), 93.

[3] E.M. Dadlez maintains that "Anne Elliot never denies that people can change, but only wonders whether Elliot has in fact changed." See Dadlez, *Mirrors to One Another: Emotion and Value in Jane Austen and David Hume* (Chichester: Wiley-Blackwell, 2009), 74. It is worth noting, however, that Anne condemns Elliot *before* his "double game" with Mrs. Clay comes to light. Again, it seems that Anne's reasoning is faulty regardless of whether her conclusions turn out to be right.

impossible. Thus Mr. Elliot must be either "a designing, wary, cold-blooded being, who thinks only of himself" or "a man of principle." There is nothing in between.

Anne's understanding of character as an epistemological problem—an understanding which, as Deidre Lynch and Morgan have pointed out, the novel does not share—is a familiar one in literary studies.[4] However, Austen adds a twist to Anne's account that makes it worth revisiting. This twist concerns Anne's reaction to the truth about Mr. Elliot's past. If her response is out of keeping with her habitual "sweetness of character," the novel takes pains to show that Anne's uncharitable sentiment is fleeting. Despite her resolution to pay Mr. Elliot no further heed—"Pity for him was all over"—by the end of the novel, Anne has recovered her "cheerful or forbearing feelings for every creature around her," including Mr. Elliot: "she could pity him" (254). Without overstating the mean-spiritedness of taking pleasure in another's moral failings (and in adopting an attitude of unfeelingness), I want to suggest that Anne's brief moment of ill will is out of character for her and, furthermore, that Austen includes this detail precisely in order to call Anne's understanding of character into question. As I have argued throughout this dissertation, the idea that knowing a person's character might be, as Lynch puts it, "a matter of unmasking" is incompatible with the idea of acting out of character.[5] Such cases are, almost by definition, not revelations. Clarissa's refusal to marry Solmes is not symptomatic of some ingrained disobedient streak, nor is Valmont's rescue of the Vicomtesse a sign of his "innate sensibility." Similarly, no reader thinks the worse of Anne for being momentarily pleased that Mr. Elliot's former wickedness has let her off the hook, so to speak. The sentiment is so unlike her and so short-lived that it is easy enough to dismiss—or miss. Yet if we turn a blind eye to

[4] See Morgan, "The Nature of Character in *Persuasion"* and Deirdre Lynch, "Jane Austen and the Social Machine" in *The Economy of Character,* 210-53.
[5] Lynch, *Economy of Character*, 246.

this incident, we risk overlooking Austen's most scathing condemnation of Anne's theory of character "as a center to be revealed": namely, that it fails to explain the character of Anne herself.[6]

Austen's *Persuasion* is an instructive case for thinking about how the problem of acting out of character survived the "rise of the novel" and continued to haunt nineteenth-century representations of subjectivity in troubling ways. This persistence has been largely overlooked by traditional accounts of the history of the novel. For many literary critics, Austen marks a turning point in the genre's history. Most famously perhaps, Ian Watt argues that the novel's potential for rendering human experience is finally achieved in the work of Austen. According to Watt, it is the creator of Elizabeth Bennet and Emma Woodhouse who first masters the "formal realism" that turns literary characters into *people*.[7] Watt takes his cue here from E.M. Forster, who praised Austen's technique for exactly the same reason: "all her characters are round, or capable of rotundity."[8] Forster's analysis is of particular interest for the way it describes versatility as the round character's calling card. He assigns special significance to moments where characters act in unexpected ways, explaining that "the test of a round character is whether

[6] The theory of character as "a center to be revealed" is Morgan's formulation, but she attributes this way of thinking not to Anne, but to Mrs. Smith and Captain Wentworth. Morgan exempts Anne from the charge of "thinking of people in layers," arguing that the novel's protagonist is the only character who understands that character is not fixed, that love (and character) is "a repetition with differences" ("The Nature of Character in *Persuasion,"* 96). I disagree—at least so far as Mr. Elliot is concerned. For her part, Lynch argues that Mr. Elliot's exposure "is curiously irrelevant to the novel as a whole…. This irrelevance testifies to how Austen rejects the mode of apprehending another's 'real character" (*Economy of Character,* 246).

[7] Ian Watt famously writes that "Jane Austen's novels… must be seen as the most successful solutions of the two general narrative problems for which Richardson and Fielding had provided only partial answers. She was able to combine into a harmonious unity the advantages both of realism of presentation and realism of assessment, of the internal and of the external approaches to character" (*Rise of the Novel,* 297).

[8] E.M. Forster, *Aspects of the Novel,* ed. Frank Kermode (London: Penguin, 2005), 78

it is capable of surprising in a convincing way. If it never surprises, it is flat. If it does not convince, it is a flat pretending to be round. It has the incalculability of life about it—life within the pages of a book."[9] Forster's point is that real life is unpredictable without being implausible and that good fiction must strive to replicate this quality. If Forster is right, then literary characters only emerge as people when they are able to act out of character in believable ways. Thus the fact that we can read about Anne's brief flirtation with ill will without questioning either its truth or her good character is proof of her roundness. What Forster shows us, finally, is that the great achievement of realist characterization is to smooth over those blips in behaviour that narrative struggles to assimilate, but which exist happily enough in the real world.

If *Persuasion*'s secondary literature is anything to go by, Anne passes Forster's test for rotundity. Few readers have remarked upon her uncharacteristic unkindness towards Mr. Elliot. Indeed, although critics have often noted Mrs. Smith's unreliability as a narrator, they have not been particularly interested in whether or not Mr. Elliot is treated unfairly. More often than not, they describe Mr. Elliot in strictly formal terms, complaining that as a villain, he is as bland as he is toothless. Thus Marilyn Butler grumbles that Mr. Elliot "seems curiously inexpressive and featureless. Jane Austen's unease in dealing with him is reflected in the inferior writing he inspires…. Failure to define the tempter-figure is surely the most significant of the failures of *Persuasion."*[10] Unlike the unsuitable suitors in Austen's other novels (Wickham, Frank

[9] Forster, *Aspects of the Novel,* 78.
[10] Marilyn Butler, *Jane Austen and the War of Ideas* (Oxford: Clarendon Press, 1987), 280. Similarly, Morgan describes Mr. Elliot as "a conscious fictional cliché" who poses no real threat to Anne's happiness ("The Nature of Character in *Persuasion,"* 89). Barbara K. Seeber takes a very different position, arguing that "narrative cameos" in the Austen novel "articulate what the main narrative tries to elide and are thus integral to the design of the novels." For Seeber, Mr. Elliot gestures to "the violence that is buried in the text"—the real and potential exploitation of Anne Elliott by her family and others—and is very much a destabilizing force in the novel See

Churchill), Mr. Elliot never achieves the fleshiness (Butler says "embodied self-sufficiency") necessary to unsettle the heroine.[11] Butler's comments offer a possible explanation for why readers tend to overlook the mean-spiritedness implicit in Anne's response to Mrs. Smith's story. Like Anne, we do not pity Mr. Elliot. Mr. Elliot is simply not round enough to register as a person and as a result, he makes no claim on our sympathies.[12]

Yet however willing Austen's readers are to overlook Anne's momentary uncharitableness and the callousness it engenders, the novel is, I think, less forgiving. There is something distinctly sinister in the way Anne's freedom—having discovered Mr. Elliot's true character, she is no longer obliged to take his feelings into account as she seeks to dispel Captain Wentworth's jealousy—is juxtaposed with Mr. Elliot's flatness. As Austen describes the scene, Anne is complicit in the flattening of Mr. Elliot. Anne's condemnation of Mr. Elliot is at once a distortion—Morgan labels him "a conscious fictional cliché"—and a dismissal. It is no surprise that Mr. Elliot disappears from the book shortly after Mrs. Smith's revelation, for as critics have often noted, the world of the novel is dominated by Anne's consciousness. To be banished from the latter is to be banished from the former as well. This unequal relationship exemplifies the

---

Seeber, *General Consent in Jane Austen: A Study of Dialogism* (Montreal: McGill-Queen's University Press, 2000), 68; 84.

[11] Interestingly, Butler notes that Mr. Elliot's affair with Mrs. Clay is out of character: "The manoeuver by which Mr. Elliott is disposed of, his *affaire* with Mrs. Clay, seems decidedly undermotivated and inconsistent with the worldly wisdom which has hitherto been his leading characteristic" (*Jane Austen and the War of Ideas,* 228). Forster would say this is a case of a flat character pretending to be round.

[12] For discussions of the human propensity to treat fictional characters as real people (and the ethical consequences of doing so), see Wayne C. Booth's classic study, *The Rhetoric of Fiction* (Chicago: University of Chicago Press, 1961); Lisa Zunshine, *Why We Read Fiction: Theory of the Mind and the Novel* (Columbus: Ohio State University Press, 2006); Blakey Vermeule, *Why Do We Care about Literary Characters?* (Baltimore: Johns Hopkins University Press, 2010); Eileen John, "Caring about Characters," in *Fictional Characters, Real Problems: The Search for Ethical Content in Literature,* ed. Gary L. Hagberg (Oxford: Oxford University Press, 2016), 31-47.

dynamic between major and minor literary characters that Alex Woloch describes in *The One Vs. the Many*. Woloch argues that the "asymmetric structure of realist characterization… rounds out one or several characters while flattening, and distorting, a manifold assortment of characters."[13] According to Woloch, literary characters must compete for narrative space as each center of consciousness threatens to encroach upon every other such center. The protagonist who emerges from this fray as a psychologically fleshed-out *person* does so at the expense of the others: "the narrative price of her achieved interiority is the distortion of many other human figures."[14] It is only by stripping Mr. Elliot of his roundness (his realness, including his capacity for moral improvement) and his right to consideration (his personhood, that which makes him worthy of compassion) that Anne can pursue her desires unfettered by either moral duty or social decorum. And although we might be tempted to absolve Anne on the grounds that she is not the kind of person who injures others in the name of self-interest, this fact does nothing to mitigate the harm done to Mr. Elliot. As Forster reminds us, the ability to act out of character *is* Anne's roundness. This roundness precipitates and seemingly justifies the flattening of Mr. Elliot. In *Persuasion* then, acting out of character is synonymous with depth and individual liberty for the agent (Anne), and surface and subordination for others (Mr. Elliot).

I have been arguing that the realist novel's impulse to make its protagonists as round as possible is in part motivated by the need to incorporate the "incalculability" of real-world human action without undermining the plausibility of the fictional world.[15] After all, a sufficiently

[13] Woloch, *The One Vs. the Many,* 31.
[14] Ibid., 34.
[15] The contrast with the French classical theater's notion of *vraisemblance* is useful here. The founder of the *Mercure galant* (later the *Mercure de France),* Jean Donneau de Visé explains that whereas historians are obligated to record events as they occurred, no matter how incredible, poets must obey the laws of *vraisemblance* and "correct truth that is not believable" quoted in

capacious character should be able to accommodate any action.[16] The result, if Anne Elliot's case is anything to go by, is that we regard one-off actions in the realist novel with something like complacency. If the novel poses the question, "Why would a good (compassionate, charitable) person perform a bad deed (take pleasure in thinking the worst of someone, treat a family member as if his feelings were negligible)?" it does so only rhetorically. The answer is simply this: literary characters sometimes surprise us because people sometimes surprise us (and literary characters *are* people). Thus roundness—the psychological depth that motivates and justifies the analogy between fictional characters and real-life people—enables us to explain away actions that might otherwise require a Coleridgean-level effort of imagination to accept.

What this means, too, is that when it comes to acting out of character, we grant fictional characters the same moral latitude as we do living people.[17] This is why Anne's "sweetness of character" is, in the eyes of most readers, untainted by her passing hardheartedness towards Mr. Elliot. Yet Woloch's analysis helps us to see this incident for what it really is: the compression and elimination of a potentially round human being. In *Persuasion,* Anne's ability to act out of character allows her to reduce Mr. Elliot to a narrative device—that not terribly effective villain Butler complains about—and excuses this inhumanity in a single gesture. Or to put it another way, Anne's power to injure Mr. Elliot with impunity—her power to do injustice—is part and

---

Lennard J. Davis, *Factual Fictions: The Origins of the English Novel* (New York: Columbia University Press, 1983), 33.

[16] I take this to be Forster's point when he observes that Austen's "characters function all round, and even if her plot made greater demands on them than it does, they would still be adequate." (*Aspects of the Novel,* 79).

[17] Lord Shaftesbury's rule of thumb suggests that we give people a lot of latitude: "when the mistakes are either in their nature so gross, or so complicated and frequent, that a creature cannot well live in a natural state nor with due affections compatible with human society and civil life, then is the character of virtue forfeited" in Shaftesbury, *Characteristics,* 175.

parcel of her roundness. This is what Austen is flagging when she associates Anne's out-of-character unkindness with the distortion and disappearance of the character called Mr. Elliot. She reminds us that it is one thing to accept Anne's uncharacteristic hardness as true, but quite another to accept it as fair (or at least excusable). Far from being a formal weakness then, the Mr. Elliot subplot is central to *Persuasion'*s thinking on the nature of character. In pointing out Anne's prejudice, the novel disputes the idea that a person's character is an unchanging, one-dimensional truth about her that is revealed through action. More radically still, by presenting Anne's ability to act out of character as a liberty to distort and dismiss the character of Mr. Elliot, Austen forces us to recognize that roundness gives excuses as well as explanations—excuses that we are too quick to accept. Acting out of character is a problem in *Persuasion* not just because it is only available to major characters, but because it excuses them when they injure their minor counterparts. The novel calls attention to this injustice by showing how Anne flattens Mr. Elliot to nothing with a single out-of-character action. In doing so, *Persuasion* openly articulates the ethical stakes of its own narrative asymmetry.

***

I have ended this story with *Persuasion* partly in order to show why it makes sense to begin an investigation into the idea of acting out of character not with Austen—that master of realist characterization—but with the moral philosophy of David Hume. *Persuasion'*s clever critique of rotundity emerges in full only once we have recognized that acting out of character is, first and foremost, an ethical problem. Taking a Wolochian standpoint toward the novel, we see the flattening of literary characters as an inevitable consequence of trying to represent reality within a limited narrative space. The dehumanizing compression of minor characters like Mr. Elliot is a precondition for the existence of fleshy persons like Anne, for Austen has only so many words

with which to paint a world. Mr. Elliot is then a victim of the distributional matrix that underlies realist characterization and Anne's strange uncharitableness is both explained by and a proof of her roundness. In a word, we see acting out of character as a formal problem. The trouble with this reading, however, is that it obscures the moral dimension of (acting out of) round character. According to Woloch's analysis, Anne's acting out of character and the harm it inflicts are both issues of form. Yet by suggesting that Anne misreads Mr. Elliot by refusing the possibility of his having changed, and by making her directly responsible for Mr. Elliot's disappearance from the story, *Persuasion* invites us to point our chastising finger at a specific person: Anne Elliot. With this invitation, Austen's novel asks us, albeit sotto voce, to set questions of representation aside and to consider instead the real repercussions of her protagonist's behavior.

Realist characterization cannot be a solution to the problem of acting out of character. This is because acting out of character is not merely a psychological phenomenon or literary device that needs explaining. Rather, it is a limit case that raises important moral questions about agency and responsibility. As I have shown, philosophers and novelists writing against the backdrop of the eighteenth-century free will debate discussed acting out of character in precisely these terms. Guided by contemporary moral theories—associationist psychology, the concept of sympathy, the doctrine of necessity—these thinkers understood acting out of character as the result of and an attestation to the extra-personality of passions and actions. They recognized that just as emotions spread beyond the person who first feels them, actions are never completely our own; the outside world necessarily determines how they unfold. What this means is that the boundary between person and world, actions and events, is almost never clear. In the eighteenth century, this confusion concerning the limits of the person is, I have argued, often figured as various permutations of acting out of character. Each of these cases leads to some form of

"malice against the self": self-estrangement and self-harm for Hume and Richardson, self-doubt for Laclos, and self-abnegation for Godwin. For these authors, a person who acts out of character has abandoned herself. She has detached herself from her actions, passions, and even her moral character, and she has done so in order to plead the cause of some other person. Think of Clarissa, who would gladly incriminate herself to absolve her family of blame, or of St. Leon, who would promote his son's happiness at the cost of his own. That these narratives so often end in tragedy (Clarissa's death, St. Leon's loss of his son) is a reminder that these quintessential goods—fellow-feeling, free will, selflessness—have not always been so universally lauded. It reminds us that at bottom, acting for another, acting out of character, and acting against ourselves are all impulses to break with the self. The danger of such a break is that it renders "ourself… in reality nothing."

Bibliography

Adams, Robert Merrihew. "Pure Love." *The Journal of Religious Ethics* 8, no.1 (1980): 83-99.

Ainslie, Donald. "Character Traits and the Humean Approach to Ethics." In *Moral Psychology (Poznan Studies in the Philosophy of the Sciences and Humanities,* vol. 94). Edited by Sergio Tenenbaum. Amsterdam: Rodopi, 2007.

Allen, B. Sprague. "Godwin as Sentimentalist." *PMLA* 33, no. 1 (1918): 1-29.

Anonymous. Review of *St. Leon,* by William Godwin. *Monthly Magazine, and British Register* (1800). In appendix C of Godwin, *St. Leon*, edited by William D. Brewer. Peterborough: Broadview Press, 2006.

———. Review of *St. Leon,* by William Godwin, *The Monthly Mirror: Reflecting Men and Manners* (1800). In appendix C of Godwin, *St. Leon,* edited by William D. Brewer. Peterborough: Broadview Press, 2006.

Árdal, Páll S. *Passion and Value in Hume's Treatise.* Edinburgh: Edinburgh University Press, 1966.

Austen, Jane. *Persuasion.* London: J.M. Dent, 1895.

Baier, Annette C. *Death and Character: Further Reflections on Hume.* Cambridge: Harvard University Press, 2008.

———. "*Enquiry concerning the Principles of Morals:* Incomparably the Best?" In *A Companion to Hume.* Edited by Elizabeth S. Radcliffe. Singapore: Blackwell, 2008.

———. "Master Passions." In *Explaining Emotions.* Edited by Amelie O. Rorty. Berkeley: University of California Press, 1988.

———. *A Progress of Sentiments: Reflections on Hume's Treatise.* Cambridge: Harvard University Press, 1991.

Baillie, James. *Routledge Philosophy GuideBook to Hume on Morality.* New York: Routledge, 2000.

Barker-Benfield, G. J. *The Culture of Sensibility: Sex and Society in Eighteenth-Century Britain.* Chicago: Chicago University Press, 1992.

Baudelaire, Charles. Notes on *Les Liaisons dangereuses* (1856). In Laclos, *Les Liaisons Dangereuses.* Edited by Catriona Seth. Paris: Gallimard, 2011.

Beam, Craig. "Hume and Nietzsche: Naturalists, Ethicists, Anti-Christians." *Hume Studies* 22, no.2 (1996): 299-324.

Bellamy, Liz. *Commerce, Morality and the Eighteenth-Century Novel.* Cambridge: Cambridge University Press, 1998.

Binhammer, Katherine. "Knowing Love: The Epistemology of *Clarissa.*" *ELH* 74 (2007): 859-79.

Blackburn, Simon. "Hume on the Mezzanine Level." *Hume Studies* 19, no. 2 (1993): 273-88.

Booth, Wayne C. *The Rhetoric of Fiction.* Chicago: University of Chicago Press, 1961.

Bowers, Toni and John Richetti. Introduction to *Clarissa Or, The History of a Young Lady*, by Samuel Richardson. Peterborough: Broadview Press, 2011.

Bricke, John. "Hume's Conception of Character." *The Southwestern Journal of Philosophy* 5, no.1 (1974): 107-13.

Brooks, Peter. *The Novel of Worldliness: Crébillon, Marivaux, Laclos, Stendhal.* Princeton: Princeton University Press, 1969.

Bueler, Lois E. *Clarissa's Plots.* Newark: University of Delaware Press, 1994.

Butler, Marilyn. *Jane Austen and the War of Ideas.* Oxford: Clarendon Press, 1987.

Byrne, Patrick. *Les Liaisons dangereuses: A Study of Motive and Moral.* Glasgow: University of Glasgow, 1989.

Castle, Terry. *Clarissa's Ciphers: Meaning and Disruption in Richardson's Clarissa.* Ithaca: Cornell University Press, 1982.

Chandler, James. *An Archaeology of Sympathy: The Sentimental Mode in Literature and Cinema.* Chicago: University of Chicago Press, 2013.

Charise, Andrea. "'The Tyranny of Age': Godwin's *St. Leon* and the Nineteenth-Century Longevity Narrative." *ELH* 79, no. 4 (2012): 905-33.

Costelloe, Timothy M. "Beauty, Morals, and Hume's Conception of Character." *History of Philosophy Quarterly* 21, no.4 (2004): 397-415.

Crocker, Lester. "The Status of Evil in *Les Liaisons dangereuses."* In *Laclos: Critical Approaches to Les Liaisons* dangereuses. Edited by Lloyd R. Free. Madrid: José Porrúa Turanzas, D.L. 1978.

Dadlez, E.M. *Mirrors to One Another: Emotion and Value in Jane Austen and David Hume.* Chichester: Wiley-Blackwell, 2009.

Davis, Lennard J. *Factual Fictions: The Origins of the English Novel.* New York: Columbia University Press, 1983.

Damrosch, Leo. *God's Plot and Man's Stories: Studies in the Fictional Imagination from Milton to Fielding.* Chicago: University of Chicago Press, 1985.

Daniel, Georges. *Fatalité du secret et fatalité du bavardage: La Marquise de Merteuil, Jean-François Rameau.* Paris: A.-G Nizet, 1966.

Davidson, Donald. "Freedom to Act." Reprinted in *Essays on Actions and Events.* Oxford: Oxford University Press, 1980.

Diaconoff, Suellen. *Eros and Power in Les Liaisons dangereuses: A Study in Evil.* Paris: Librairie Droz, 1979.

Dunn, Susan. "Valmont, Actor and Spectator." *French Review* 58 (1984): 41-47.

Duranton, Henri. "Les Liaisons dangereuses ou le miroir ennemi." *Revue des sciences humaines* 153 (1974): 125-43.

Eagleton, Terry. *The Rape of Clarissa: Writing, Sexuality, and Class Struggle in Samuel Richardson.* Minneapolis: University of Minnesota Press, 1982.

Fénelon, François. *A Dissertation on Pure Love* (Dublin: Isaac Jackson, 1739).

———. "Maxims of the Saints." In *Fénelon: Selected Writings*. Edited by Chad Helms. New York: Paulist Press, 2006.

Ferguson, Frances. *Pornography, the Theory: What Utilitarianism Did to Action*. Chicago: Chicago University Press, 2004.

———. "Rape and the Rise of the Novel." *Representations* 20 (1987): 88-112.

———. "The Sublime and Education: Educational Rationalization/ Sublime Reason." *Romantic Circles.* Accessed October 25, 2016. http://www.rc.umd.edu/praxis/sublime_education/ferguson/ferguson.html

Fieser, James, Editor. *Early Responses to Hume's Metaphysical and Epistemological Writings I.* Bristol: Thoemmes Press, 2000.

Finn, Margot. *The Character of Credit: Personal Debt in English Culture, 1740-1914.* Cambridge: Cambridge University Press, 2003.

Fisher, Philip. *The Vehement Passions.* Princeton: Princeton University Press, 2009.

Foot, Philippa. “Freewill as Involving Determinism.” *Virtues and Vices: And Other Essays in Moral Philosophy.* Oxford: Oxford University Press, 2002.

Forster, E.M. *Aspects of the Novel.* Edited by Frank Kermode. London: Penguin, 2005.

Frykholm, Erin. “The Ontology of Character Traits in Hume,” supplement. *Canadian Journal of Philosophy* 42, no. S1 (2012): 82-97.

Fulton, Gordon. *Styles of Meaning and Meanings of Style in Richardson’s ‘Clarissa’.* Montreal: Mcgill-Queen’s University Press, 1999.

Gallagher, Catherine. *Nobody’s Story: The Vanishing Acts of Women Writers in the Marketplace,1670-1820.* Berkeley: University of California Press, 1995.

Godwin, William. *Caleb Williams.* Oxford: Oxford University Press, 2009.

———. *An Enquiry Concerning Political Justice.* Edited by Mark Philp. Oxford: Oxford University Press, 2013.

———. “Of History and Romance.” Written in 1797, published in 1988. Accessed October 30, 2016, http://www.english.upenn.edu/~mgamer/Etexts/godwin.history.html

———. *St. Leon: A Tale of the Sixteenth Century.* Edited by William D. Brewer. Peterborough: Broadview Press, 2006.

———. *Thoughts on Man, his Nature, Productions, and Discoveries.* London: Effingham Wilson, Royal Exchange, 1831.

Gold, Alex Jr. “It’s Only Love: The Politics of Passion in Godwin’s *Caleb Williams.” Texas Studies in Literature and Language* 19, no. 2 (1977): 135-60.

Handwerk, Gary and A.A. Markley. Introduction to *Caleb Williams,* by William Godwin. Peterborough: Broadview Press, 2000.

Hartley, David. *Observations on man, his frame, his duty, and his expectations.* London: J. Johnson, 1749.

Hinton, Laura. “The Heroine’s Subjection: Clarissa, Sadomasochism, and the Natural Law.” *Eighteenth-Century Studies* 32 (1999): 293-308.

Hobbes, Thomas. *Leviathan.* Oxford: Oxford World’s Classics, 2008.

Hume, David. *Dialogues Concerning Natural Religion: And Other Writings.* Oxford: Oxford University Press, 2008.

———. *An Enquiry concerning the Principles of Morals.* Edited by Tom L. Beauchamp.

Oxford: Clarendon Press, 1998.

———. *A Treatise of Human Nature.* Edited by L.A. Selby-Bigge. Oxford: Clarendon Press, 1978.

Jeffreys, George. Letter to William Duncombe, 28 February 1749. In *Clarissa: The Eighteenth-Century Response 1747-1804.* Edited by Lois E. Bueler. New York: AMS Press, 2010.

John, Eileen. "Caring about Characters." In *Fictional Characters, Real Problems: The Search for Ethical Content in Literature.* Edited by Gary L. Hagberg. Oxford: Oxford University Press, 2016.

Johnson, Samuel. "Recollections of Dr. Johnson by Miss Reynolds." In *Johnsonian Miscellanies*. Edited by George Birbeck Norman Hill. New York: Harper, 1897.

Jones, Wendy S. *Consensual Fictions: Women, Liberalism, and the English Novel.* Toronto: University of Toronto Press, 2005.

Kail, Peter. "Hume and Nietzsche." In *The Oxford Handbook of Hume.* Edited by Paul Russell. Oxford: Oxford University Press, 2016.

Kant, Immanuel. *Groundwork for the Metaphysics of Morals.* Translated by Mary Gregor. Cambridge Cambridge University Press, 2002.

Kavanaugh, Thomas. *Enlightened Pleasures: Eighteenth-Century France and the New Epicureanism.* New Haven: Yale University Press, 2010.

Keymer, Thomas. *Richardson's Clarissa and the Eighteenth-Century Reader.* Cambridge: Cambridge University Press, 1992.

Kinkead-Weekes, Mark. "*Clarissa* Restored?" *Review of English Studies* 10 (1959): 156-71.

Kupperman, Joel. "Character and Self-Knowledge." *Proceedings of the Aristotelian Society* 85 (1985): 219-238.

Kramnick, Jonathan. *Actions and Objects: from Hobbes to Richardson.* Stanford: Stanford University Press, 2010.

Laclos, Pierre-Ambroise Choderlos de. *Les Liaisons dangereuses.* Manuscript c. 1779-1781. Bibliothèque Nationale de France, Département des Manuscrits, Doc. 12845, fol. 35r.

———. *Les Liaisons dangereuses.* Edited by Catriona Seth. Paris: Gallimard, 2011.

———. *Les Liaisons dangereuses.* Translated by Richard Aldington. London: Routledge, 2010.

———. *Œuvres complètes*. Edited by Laurent Versini. Paris: Gallimard, 1979.

Lecaldano, Eugenio. “The Passions, Character, and the Self in Hume.” *Hume Studies* 28, no. 2 (2002): 175-193.

Locke, Don. *A Fantasy of Reason: The Life and Thought of William Godwin.* London: Routledge, 1980.

Lowrie, Joyce O. “The Prévan Cycle as Pre-Text in Laclos’s *Les Liaisons dangereuses.”* In *Sightings: Mirrors in Texts—Texts in Mirrors.* Amsterdam: Rodopi, 2008.

Lynch, Deirdre. *The Economy of Character: Novels, Culture and the Business of Inner Meaning.* Chicago: Chicago University Press, 1998.

Macpherson, Sandra. *Harm’s Way: Tragic Responsibility and the Novel Form.* Baltimore: Johns Hopkins University Press, 2009.

Mahoney, Robert Heath. “Hume’s Conception of Character.” PhD diss., University of Southampton, School of Humanities, 2009.

Maimonides, Moses. *The Guide of the Perplexed.* Translated by Shlomo Pines. Chicago: Chicago University Press, 1963.

Marshall, Peter. *William Godwin.* New Haven: Yale University Press, 1984.

Marter, Shirley van. “Richardson’s Revisions of *Clarissa* in the Third and Fourth Editions.” *Studies in Bibliography* 26 (1973): 107-32.

McCallam, David. *L’art de l’équivoque chez Laclos* .Genève: Droz, 2008.

McIntyre, Jane. “Character: A Humean Account.” *History of Philosophy Quarterly* 7 (1990): 193-205.

———. “Hume and the Problem of Personal Identity.” In *The Cambridge Companion to Hume.* Edited by David Fate Norton and Jacqueline Anne Taylor. Cambridge: Cambridge University Press, 2009.

Mercer, Philip. *Sympathy and Ethics.* Oxford: Clarendon Press, 1972.

Miller, Nancy K. *The Heroine's Text: Readings in the French and English Novel, 1722–1782*. New York: Columbia University Press, 1980.

Moody, Christopher Lake. Review of *St. Leon* (1800) in appendix C of Godwin, *St. Leon*, edited by William D. Brewer. Peterborough: Broadview Press, 2006.

Moody-Adams, Michele. “On the Old Saw that Character is Destiny.” In *Identity, Character,*

and Morality: Essays in Moral Psychology. Edited by Owen J. Flanagan and Amelie Oksenberg Rorty. Boston: Massachusetts Institute of Technology Press, 1997.

Morgan, Susan. "The Nature of Character in *Persuasion."* In *Jane Austen's Persuasion.* Edited by Harold Bloom. Philadelphia: Chelsea House, 2004.

Nagel, Thomas. "Moral Luck." In *Mortal Questions.* Cambridge: Cambridge University Press, 2000.

Nazar, Hina. "Judging Clarissa's Heart." In *Enlightened Sentiments: Judgment and Autonomy in the Age of Sensibility.* New York: Fordham University Press, 2012.

Parr, Samuel. *A Spital Sermon, preached at Christ Church, upon Easter Tuesday, April 15, 1800; to which are added Notes.* London: J. Mawman, 1801.

Perry, Ruth. *Novel Relations: The Transformation of Kinship in English Literature and Culture, 1748-1818.* Cambridge: Cambridge University Press, 2004.

Philp, Mark. *Godwin's Political Justice.* London: Duckworth, 1986.

Pinch, Adela. *Strange Fits of Passion: Epistemologies of Emotion, Hume to Austen.* Stanford: Stanford University Press, 1996.

Pitson, A.E. *Hume's Philosophy of the Self.* London: Routledge, 2002.

Postema, Gerald. "Cemented with Diseased Qualities: Sympathy and Comparison in Hume's Moral Psychology." *Hume Studies* 31, no. 2 (2005): 249-98.

Purviance, Susan M. "The Moral Self and the Indirect Passions." *Hume Studies* 23, no.2 (1997): 195-212.

Radcliffe, Elizabeth S. "Hume's Psychology of the Passions: The Literature and Future Directions." *Journal of the History of Philosophy* 53, no. 4 (2015): 574-81.

Radcliffe, Evan. "From 'Metaphysician' to 'Novelist': 'Political Justice', 'Caleb Williams', and the Tension between Philosophical Argument and Narrative." *Modern Philology* 97, no.2 (2000): 528-53.

———. "Revolutionary Writing, Moral Philosophy, and Universal Benevolence in the Eighteenth Century." *Journal of the History of Ideas* 54 (1993): 221-240.

Rajan, Tilottama. *Romantic Narrative: Shelley, Hays, Godwin, Wollstonecraft.* Baltimore: Johns Hopkins University Press, 2010.

Reid, Thomas. *An Inquiry into the Human Mind, on the Principles of Common Sense.* Edinburgh: A. Millar, 1764.

Richardson, Samuel. *Clarissa, or The History of a Young Lady.* Edited by Angus Ross. London: Penguin, 2004.

———. *Clarissa, Or, the History of a Young Lady.* Third Edition. London: Printed for S. Richardson, 1751.

Rosbottom, Robert. "Dangerous Connections." In *Laclos: Critical Approaches to Les Liaisons dangereuses*. Edited by Lloyd R. Free. Madrid: José Porrúa Turanzas, 1978.

Roulston, Christine. *Virtue, Gender, and the Authentic Self in Eighteenth-Century Fiction: Richardson, Rousseau, and Laclos.* Miami: University Press of Florida, 1998.

Rousseau, Jean-Jacques. *The Confessions and Correspondence, including the Letters to Malesherbes.* Translated by Christopher Kelly. London: University Press of New England for Dartmouth College, 1995.

———. *Emile, or On Education.* Translated by Allan Bloom. New York: Basic Books, 1979.

Sahlins, Marshall. *What Kinship Is—And Is Not.* Chicago: Chicago University Press, 2013.

Sams, Henry W. "Anti-Stoicism in Seventeenth- and Early Eighteenth-Century England." *Studies in Philology* 41, no.1 (1944): 65-78.

Schmidt, Claudia M. *David Hume: Reason in History.* University Park: Penn State Press, 2010.

Schneewind, Jerome B. *The Invention of Autonomy: A History of Modern Moral Philosophy.* Cambridge: Cambridge University Press, 1998.

Seeber, Barbara K. *General Consent in Jane Austen: A Study of Dialogism.* Montreal: McGill-Queen's University Press, 2000.

Seylaz, Jean-Luc. *Les Liaisons dangereuses et la création romanesque chez Laclos.* Paris: Minard, 1958.

Shaftesbury, Anthony Ashley Cooper. *Characteristics of Men, Opinions, Times.* Cambridge: Cambridge University Press, 1999.

Smith, Adam. *The Theory of Moral Sentiments.* Edited by D.D. Raphael and A.L. Macfie. Indianapolis: Liberty Fund, 1984.

Smith, Nick. *Justice through Apologies: Remorse, Reform, and Punishment.* Cambridge: Cambridge University Press, 2014.

Stanzel, F.K. *A Theory of Narrative.* Translated by Charlotte Goedsche. Cambridge: Cambridge University Press, 1984.

Stevenson, John Allen. "The Courtship of the Family: Clarissa and the Harlowes Once More." *ELH* 48 (1981): 757-77.

Stuber, Florian. "On Fathers and Authority in *Clarissa*." *Studies in English Literature, 1500-1900* 25 (1985): 557-74.

Taylor, Derek E. *Reason and Religion in 'Clarissa': Samuel Richardson and 'the famous Mr. Norris, of Bemerton.'* Farnham: Ashgate Press, 2009.

Taylor, Jacqueline. "The Idea of a Science of Human Nature." In *The Oxford Handbook of British Philosophy in the Eighteenth Century.* Edited by James A. Harris. Oxford: Oxford University Press, 2013.

Thelander, Dorothy. *Laclos and the Epistolary Novel.* Genève: Droz, 1963.

Todorov, Tzvetan. *Littérature et signification.* Paris: Larousse, 1967.

Vermeule, Blakey. *Why Do We Care about Literary Characters?* Baltimore: Johns Hopkins University Press, 2010.

Versini, Laurent. *Laclos et la tradition, essai sur les sources et la technique des Liaisons dangereuses.* Paris : Klincksieck, 1968.

———. *Le Roman le plus intelligent: Les Liaisons dangereuses de Laclos.* Paris: H. Champion, 1998.

Vila, Anne C. *Enlightenment and Pathology: Sensibility in the Literature and Medicine of Eighteenth-Century France.* Baltimore: Johns Hopkins University Press, 1998.

Warner, William. *Reading Clarissa: the Struggles of Interpretation.* New Haven: Yale University Press, 1979.

Watt, Ian. *The Rise of the Novel: Studies in Defoe, Richardson, and Fielding.* Berkeley: University of California Press, 1962.

Williams, Bernard. *Making Sense of Humanity: And Other Philosophical Papers, 1982-1993.* Cambridge: Cambridge University Press, 1995.

Wilt, Judith. "He Could Go No Farther: A Modest Proposal about Lovelace and Clarissa." *PMLA* 92 (1977): 19-32.

Winnett, Susan. *Terrible Sociability: the Text of Manners in Laclos, Goethe, and James.* Stanford: Stanford University Press, 1993.

Woloch, Alex. *The One Vs. the Many: Minor Characters and the Space of the Protagonist in the*

Novel. Princeton: Princeton University Press, 2003.

Zizek, Slavoj. *Metastases of Enjoyment: Six Essays on Women and Causality.* London: Verso, 1994.

Zunshine, Lisa. "Richardson's *Clarissa* and a Theory of Mind." In *The Work of Fiction: Cognition, Culture, and Complexity.* Edited by Alan Richardson and Ellen Spolsky. Burlington: Ashgate, 2004.

———. *Why We Read Fiction: Theory of the Mind and the Novel* (Columbus: Ohio State University Press, 2006