DARWIN ON THE “MORE INTERESTING” AND “MOST INTRICATE” PROBLEMS FOR NATURAL SELECTION

A Thesis

Submitted to the Graduate School of the University of Notre Dame in Partial Fulfillment of the Requirements for the Degree of Master of Arts

by

Erica Schiller Freeman

Phillip R. Sloan, Director

Graduate Program in History and Philosophy of Science Notre Dame, Indiana

July 2013
CONTENTS

Chapter 1: Introduction .................................................................................................................. 1
  1.1 A Very Brief History of Moral Philosophy from Aquinas to Utilitarianism ....3

Chapter 2: Tracing Two Sources of Tension in the Reception of the Descent of Man ....... 9

Chapter 3: The Formation of Darwin's Moral Theory ................................................................. 23
  3.1 Prologue: From Darwin's University Years to the Notebook Period ......................... 23
  3.2 Moral Sense: Darwin and David Hume's Philosophy ............................................... 28
  3.3 Mackintosh on Benthamite Utilitarianism ............................................................... 39
  3.4 John Stuart Mill: From Bentham to Coleridge and Goethe ...................................... 43

Chapter 4: A Comparison of the Moral Philosophies of Hume, J. S. Mill, and Darwin ... 51
  4.1 Prologue: Darwin's Reaction to J. S. Mill's Appraisal of Natural Selection ........... 51
  4.2 Hume, J. S. Mill, and Darwin on Moral Epistemology ........................................... 52
  4.3 Hume, J. S. Mill, and Darwin on Moral Metaphysics ............................................. 69
  4.4 Hume, J. S. Mill, and Darwin on Moral Obligation and Human Nature ............ 78

Chapter 5: Darwin's Positive Moral Theory in the Descent of Man ....................................... 81
  5.1 Prologue: Richards and Ruse on Darwin's Approach to Morality ......................... 81
  5.2 Adam Smith's Moral Psychology in Darwin's Positive Moral Theory .................. 86
  5.3 Darwin: The Notebook Period to the Descent of Man .......................................... 89
  5.4 The Tension in Darwin's Positive Moral Theory ...................................................... 93

Chapter 6: The Foundation of Darwin's Moral Theory: "Nature" and "Selection" ....... 103

Chapter 7: Conclusion ............................................................................................................... 113

Bibliography ............................................................................................................................. 121
CHAPTER 1:  
INTRODUCTION

This paper is a critique of the important contribution that Robert J. Richards has made to the interpretation of Charles Darwin’s account of the origin of ethics.¹ My aim is to evaluate the adequacy of Richards’s portrayal of Darwin as a moral sense theorist indebted to German Romanticism for the philosophical resources required to articulate a fully naturalistic account of the emergence of moral sense. I pursue my aim of critiquing this portrayal through an examination of Darwin’s approach to what he called, in the final chapter of the Descent of Man, the “more interesting” and “most intricate” problems for the theory of natural selection. The “more interesting” problem concerns the origin of

moral behavior and conscience. The “most intricate” problem refers to the concept of progress as it pertains to the continued evolution by natural selection of the human species. I show that human development toward moral perfection (the “more interesting” problem) involves the development of humanitarian culture, and yet it is precisely the development of humanitarian culture that gives rise to the conditions Darwin acknowledged as necessary for the continued action of natural selection on the members of any population. This point of tension within Darwin’s work is a valuable location from which to evaluate the work of Robert J. Richards and others who have admirably deepened our understanding of the polyphony of sources that jointly influenced Darwin’s view.

In Chapter 3, I clarify Darwin’s approach to the “more interesting” problem of the origin of moral sense through natural selection. I claim that Darwin’s argument in the Descent is put in relief when compared at length with the ethical theory of John Stuart Mill. Darwin’s moral psychology in the Descent of Man represents a mature attempt to demonstrate the claim from his early Notebooks that his principle of natural selection subsumes both moral sense and Utilitarian principles of morality. Rather than overturning either of these earlier theories, natural selection supposedly explains them – and, yet, it seems to do so only by significantly reframing central moral concepts.

In Chapter 4, I continue my comparison of David Hume, J. S. Mill, and Darwin on basic themes of moral philosophy, including ethical knowledge, obligation, motivation, and the relation of these to the general character of human nature. Finally, I closely examine the importance of three footnotes inserted into the second edition of The
Descent of Man,\textsuperscript{2} Chapter 4, that enables us to understand Darwin’s account of the origin of moral sense, and the relationship of this account to Utilitarianism.

In Chapter 5, I offer a summary and analysis of Darwin’s theory of the development of moral qualities in terms of his distinction between higher moral rules (based in social instinct) and lower moral rules (based in public opinion). I claim that Darwin’s attempt to account for moral progress in terms of both biological and moral perfectionism creates a significant tension within his moral theory. This tension informs Darwin’s response to the “most intricate” problem of human progress through natural selection. He offers conflicting responses to the question of what to do with unfit people.

In Chapter 6, I argue that multiple, non-equivalent layers can be discerned in the concept of nature Darwin invoked over the course of his career. Darwin gradually came to understand the action of natural selection across generations in terms of nature understood as “vital ground,” “selector,” and “a-teleological system of laws.” I describe the process through which his complex nature concept developed through Darwin’s innovative appropriation of the principles in the writings of Alexander von Humboldt and Thomas Malthus.

1.1 A Very Brief History of Moral Philosophy from Aquinas to Utilitarianism

The issues of moral normativity and human progress were well-established areas of philosophical inquiry long before Darwin decided to address these issues from the perspective of natural history. Darwin’s approach to these issues centers on a dialogue he

creates among moral sense ethics, qualitative Utilitarianism, and Darwin’s own principle of natural selection. My investigation of Darwin’s response to the “more interesting” and “most intricate” problems thus has implications for the proper placement of Darwin within the western tradition of moral philosophy.

Seventeenth- and eighteenth-century British philosophers were particularly concerned with moral normativity, whereas the twin issues of human progress and social reform were central for British philosophers writing in the late eighteenth and early nineteenth centuries. Most of these philosophers were in some way responding to Thomas Aquinas’s [1225-1274] Treatise on Law, which represented a synthesis of elements of Aristotle’s ethics, Stoicism, and the teachings of the Hebrews. Aquinas distinguished among four kinds of law: (1.) divine eternal law, (2.) natural law which is a reflection of eternal law in the natural world, (3.) divine positive law, and (4.) human positive law. Aquinas claimed that nature itself is moral: natural law mandates justice and equality insofar as it reflects the eternal law, the end of which is the common good. Aquinas allowed that practical reason can discover natural law, but he recognized that knowledge alone cannot move conscience. Thus, divine law specifies through revelation the precise character of the moral distinctions, which, if obeyed, are conducive to eternal happiness.

Moral sense philosophy in general departed from Aquinas’s moral theory with the claim that humans discover moral distinctions through sense rather than through a

---


combination of reason and revelation. Lord Shaftesbury, Anthony Ashley Cooper [1671-1713], inaugurated this departure through his anti-rationalist stance, denying that reason can motivate moral behavior; on his view, only affections and passion can motivate. An individual creature of any kind is “good” if it acts conducive to the good of its kind as a whole. In this respect Shaftesbury maintained a version of Aquinas’s teleological link between individual good (Aquinas’s divine positive law) and common good (Aquinas’s divine eternal law), a theme upon which qualitative Utilitarians likewise elaborated. Furthermore, Shaftesbury clarified that only humans among all creatures have “moral sense,” an innate reflective capacity for moral distinctions. Shaftesbury’s emphasis on moral epistemology and normativity thus differs from the British moral philosophers of the next century who would emphasize moral evaluation with an eye to addressing the issue of human progress.

Moral sense philosophers differed in terms of the precise explication of moral sense. Shaftesbury described moral sense as a “sense of right and wrong” that presupposes reflective awareness of first-order passions and affections. Shaftesbury’s version of moral sense consists at least of second-order affections: the moral sentiments arising from the distinct pleasure (or pain) felt on reflection toward a first-order affection

5 MacIntyre, *A Short History of Ethics*, 162.


7 Shaftesbury, “Inquiry Concerning Virtue or Merit,” 169.

8 Ibid., 178-79.

9 Ibid., 179-80.

10 Ibid., 172.
or passion. Francis Hutcheson [1694-1746] agreed with Shaftesbury that individual good is linked with the common good, but he differed from his predecessor in the precise relationship he perceived between the moral sentiments and the pleasure (or pain) associated with moral evaluation. Hutcheson affirmed that moral sentiments are primary and thus disinterested. Hutcheson’s revision of earlier moral sense theory more effectively protected this form of anti-rationalist moral philosophy from the threat of moral egoists such as Bernard Mandeville [1670-1733] and Thomas Hobbes [1588-1679].

Within the framework of Aquinas’s four-fold distinction of law, the moral sense tradition appears as a critical commentary on the relationship between natural law, divine eternal law, and divine positive law. Shaftesbury and Hutcheson endeavored to account for moral normativity otherwise than through a strong appeal to human reason, yet without sacrificing the justification for altruistic behavior. The Utilitarian tradition, in contrast, can be considered as a commentary on the relationship between natural law, human law, and moral law.

Early Utilitarians, including Richard Cumberland [1631-1718] and John Gay [1699-1745], were contemporaries of Shaftesbury and Hutcheson but responded to Aquinas’s moral theory differently. Gay claimed that God mandates individual moral obligation but also wills the common good.11 This theological version of Utilitarianism thus emphasized the link between eternal law and divine moral law without the characteristic stronger emphasis of later Utilitarians on social reform through human law.

Jeremy Bentham [1748-1832] represents the first articulation of later forms of Utilitarianism. Bentham departed more significantly from the theistic moral tradition than his predecessors. Aquinas had claimed that human law is an imperfect attempt to prevent harm and protect individual members of society. Divine moral law is required to supplement for the imperfections of human law and to provide the justification for defying unjust laws. The assumption that pleasure and pain “govern us in all we do, in all we say, in all we think” may explain immoral behavior and unjust laws; yet, for Bentham, the human capacity for moral evaluation enables recognition of the principle of utility according to which the utility of virtue for the common good can be recognized and promoted through effective legislation. Bentham’s emphasis on the ability to overcome the alleged egoism of human nature through the exercise of human reason places his moral theory at some distance from the moral sense tradition, in closer proximity to the moral rationalists.

As I discuss in Section 1D, John Stuart Mill [1806-1873] departed from Bentham by rejecting moral egoist psychology. Similar to the earlier moral sense theorists, the younger Mill affirmed a potentially benevolent human nature. Although he persisted with Bentham in explicating the moral good in terms of pleasure (or pain), J. S. Mill nevertheless qualitatively distinguished the many pleasures. On both points, J. S. Mill’s view of individual moral development is reminiscent of the link between individual and common good defended by Aquinas and maintained by Shaftesbury and Hutcheson.

---

CHAPTER 2:
TRACING TWO SOURCES OF TENSION IN THE RECEPTION OF THE DESCENT OF MAN

Darwin’s account of human origins in terms of natural selection was preceded in England by Thomas Henry Huxley [1825-1895], Charles Lyell [1797-1875], and Alfred Russel Wallace [1823-1913] and in German-speaking regions by Ernst Haeckel in his Anthropogeny.13 These contemporaries of Darwin articulated evolutionary accounts of the emergence of the physical properties of humans, but did not offer a naturalistic account of the full range of human properties. T. H. Huxley proposed that ethics and morality develop in opposition to biological evolution through a process of cultural evolution, whereas Lyell and Wallace denied that these aspects of human life could be explicated within any kind of evolutionary framework. Darwin’s own position as it was first comprehensively set forth in the first edition of the Descent of Man, which represented the first attempt within the British natural historical tradition to account for these most distinctive features of human beings in terms of his principle of natural selection. His argument centers on a demonstration of the high probability that, “any animal whatever, endowed with well-marked social instincts, the parental and filial affections being here included, would inevitably acquire a moral sense or conscience, as soon as its intellectual

powers had become as well, or nearly as well developed, as in man.”\textsuperscript{14} Thus Darwin set out to address the “development of the moral qualities”\textsuperscript{15} and the “advancement of the welfare of mankind,”\textsuperscript{16} the problems he considered the “more interesting” and “most intricate”\textsuperscript{17} for the theory of natural selection.

Darwin’s attempt to include all aspects of human life within the scope of natural history appeared to have implications for the tenability of traditional theological frameworks. The early reception of Darwin’s treatment of the origin of the specifically human form of life and moral behavior centered on the implications, for Christian theology and philosophy, of Darwin’s denial of any absolute distinction between humans and nonhuman animals. The debates between St. George Jackson Mivart [1827-1900] and T. H. Huxley in Britain and Ernst Heinrich Haeckel [1834-1919] and Erich Wasmann [1889-1931] in the German-speaking regions of continental Europe illustrate the tensions ensuing upon the publication of the Descent. Both pairs of disputants were likewise concerned with the justification for extending the concept of homology, central to Darwin’s argument in the Descent, from anatomy to internal features such as mind, consciousness, and emotions. A third common concern was whether methodological naturalism was even applicable to the question of human origins. None of these issues was amenable to empirical resolution; rather, the opposing perspectives emerged from

---

\textsuperscript{14} Darwin, \textit{Descent of Man (2\textsuperscript{nd} ed)}, 682.

\textsuperscript{15} Ibid., 1048.

\textsuperscript{16} Ibid., 1054.

\textsuperscript{17} Ibid., 1048, 1054.
differing metaphysical frameworks and were further inflamed by influential elements of the social contexts, such as church-state relations and nationalism.

Darwin provided his account of the origin of ethics within his broader account of human origins. He was once again joined by his British contemporaries in his interest in addressing this issue from the natural historical perspective. Herbert Spencer [1820-1903] was allegedly the first to articulate a version of “evolutionary ethics.”18 Moreover, what was for Darwin the “more interesting” problem for his own theory of natural selection was the animating question for Spencer’s philosophical inquiry.19 Spencer published Social Statics in 1851 at least in part as an attempt to identify and analyze the necessary conditions for social morality. Spencer’s account appropriated many elements of Utilitarianism but was critical of the greatest happiness principle.20 Spencer acknowledged a “selfish instinct” in humans,21 but he nevertheless promoted a progressive view of human evolution, according to which moral behavior would more closely achieve the standards of virtue as humans gradually became increasingly adapted in a quasi-Lamarckian manner to the social environment.22

Darwin’s own account of the origin of ethics, which I examine at length in Section IV, was subject to critical reception. T. H. Huxley, who generally supported Darwin’s theory of natural selection, denied that Darwin’s (or any other evolutionary)


19 Farber, The Temptations of Evolutionary Ethics, 39.

20 Ibid., 40.


22 Farber, The Temptations of Evolutionary Ethics, 41 and 49.
approach to ethics could adequately address both the instinctual and cultural aspects of human nature. A. R. Wallace departed more fundamentally from Darwin by allowing that natural selection affected humans only to a limited extent. In the late nineteenth century, the Utilitarian moral philosopher Henry Sidgwick [1838-1900] was perhaps the first to reject all forms of evolutionary ethics as logically unsound, but as Alisdair MacIntyre has noted, Sidgwick was never able to identify an alternative foundation for ethics. The Oxford philosopher Thomas Hill Green was one of the many professionals who supported Sidgwick’s claim that evolutionary ethics commits what would later be termed the “naturalistic fallacy.” G. E. Moore maintained this response to natural historical accounts of the origin of ethics into the twentieth century. His primary concern while writing his *Principia Ethica* was to defend a strict distinction between ethics and empirical science. Moore thus claimed that moral good cannot be explained in terms of more basic physical principles.

Following a few decades after the synthesis of population genetics with evolutionary biology in the 1930s and 1940s, a renewed effort to give an evolutionary account of ethics was formulated, known commonly as “sociobiology.” The premier representative of this new form of evolutionary ethics, Edward O. Wilson, defined


24 Ibid., 70; see also Alfred Russell Wallace, *Darwinism* (Whitefish, MT: Kessinger Publishing, 2004).


sociobiology as “the systematic study of the biological basis of all social behavior,” with the ultimate goal to “reformulate the foundations of the social sciences in a way that draws these subjects into the Modern Synthesis.” The general thesis of sociobiology can be summarized as the claim that human social behavior has a functional biological origin. In 1971, the sociobiologist Robert Trivers presented a genetic interpretation of altruism in a now-classic paper. Building on the work of Trivers and others, Oxford zoologist Richard Dawkins endorsed an extreme form of scientific sociobiology, characterizing humans as epiphenomenal machines created by selfish genes. Dawkins and his supporters claimed that the selfish gene theory does not commit the anthropomorphic fallacy; they asserted that terms like “selfish” and “altruistic” have been redefined in value-free language. Critics retorted that this theory in fact reduces genuine altruism to the familiar sense of selfishness, since Dawkins’s view assumes that the purpose of all behavior, including altruism, is “to help the gene survive.”

The sociobiology debate proceeded to address claims about the nature of moral discourse, the relationships among moral concepts, the relationships among moral thinking and other forms of practical thinking, and accounts of moral institutions within human societies. The sociobiology debate thus included controversy over “appetitive

utterances” and as well as renewed debates over the naturalistic fallacy, with some authors asserting that evolutionary ethics necessarily commits this fallacy, and others arguing that the flaws of evolutionary ethics lie elsewhere.

Sociobiologists have contended that their general approach is a development and extension of Darwin’s original theory. This claim has been considered debatable, which in part explains the emergence of a trend in the literature concerned with clarifying the character and implications of Darwin’s account of moral sense, his views on the development of moral properties, and his position concerning the conditions necessary for the continued progress of the human species. Anthropologists and sociologists as well as historians and philosophers of science have contributed to this trend.

Historian John Greene, for example, traced a divide in scholarship in this field between those who interpreted Darwin as a social Darwinist and even as a racist, and those who ascribed to Darwin an interactionist position on human progress. This strong interpretative divide persists to this day in the scholarship on Darwin’s position on


34 See, for example, Quinton, “Ethics and the Theory of Evolution.”


human progress through evolution by natural selection. The stronger Social Darwinist reading finds in Darwin the “belief that competitive struggle between individuals, tribes, nations, and races has been the chief engine of progress in social evolution.”38 Critic of evolutionary theory, Richard Weikart, has emerged as a prominent opponent to Darwinism, based on its allegedly eugenic implications.39 Yet Weikart’s interpretation has not avoided significant criticism by several Darwin scholars, including Robert J. Richards.40 The contrasting interpretation claims that Darwin recognized that learned behaviors had at a certain point in human evolution become more important than natural selection in driving progress. Adrian Desmond and James Moore represent this interpretation with their recent argument in favor of a view of Darwin as a humanitarian


Rather than investigate the implications of Darwin’s theory, Edward Manier instead researched the sources that Darwin drew upon as he developed his theory. Manier used Darwin’s Notebooks as the foundation for a pioneering investigation into the “cultural circle” that influenced Darwin during the years 1837-1844 in which he developed the first versions of what would become his theory of natural selection.

Against the view that Darwin approached the species question without presuppositions, Manier argued that Darwin’s engagement with philosophy, literature, and political theory “had a central impact upon the language, form and content of the theory of natural selection.”\footnote{Edward Manier, \textit{The Young Darwin and His Cultural Circle} (Dordrecht, The Netherlands: Reidel Publishing Company, 1978).}

Moreover, Manier claimed that Darwin considered “all the major alternatives
with which he had significant contact and which he understood” precisely during the years of drafting. As a result of his investigation, Manier had evidence to suggest, for example, that “[the philosopher David] Hume…and Darwin, shared a complex critical sensibility which placed them well outside the orthodox mainstream of the period.” In a primarily favorable book review, Michael Ruse shared Manier’s own concern that further research was needed in order strongly to defend the Hume-Darwin connection. Robert J. Richards, for his part, praised Manier for paying close attention to the influence of the biology of behavior on the formation of Darwin’s drafts of the theory of natural selection. More recently, Phillip Sloan has noted that Manier’s research initiated “considerable commentary” on the role of metaphor in Darwin’s description of nature. Prior to the close examination by Manier and others of the manuscripts and notebooks for the sources of influence on Darwin’s early work, the standard view held that Darwin was exclusively “formed by British natural theology, utilitarian ethics, and British political economy.” From 1969 to 1973, intellectual historian and historian of science Robert M. Young published a series of articles arguing for this strongly “contextualist”

46 Manier, The Young Darwin.


52 Sloan, “The ‘Sense of Sublimity,’” 257.
interpretation of the origins of Darwin’s position on the origin of species. In particular, Young defended the view that “Darwin received a decisive impetus from [Thomas] Malthus when formulating his theory of natural selection.” Silvan S. Schweber has supported Young’s argument for the significance of British political economy on the formation of Darwin’s thought. Desmond and Moore’s biography of Darwin marks a subsequent contribution to this tendency to emphasize the role of the British context in the formation of Darwin’s thought. Contextualism in Darwin scholarship has since received critical review by Donald Winch of the University of Sussex.

Manier’s investigation of Darwin’s early, post-Beagle years inaugurated a distinct tendency in the literature to defend a link between natural selection and Romanticism. Sloan has noted that The Young Darwin “discusses significant aspects of Darwin’s encounter with early British romanticism in the 1837-1842 period…and emphasizes


Wordsworth’s influence.” David Kohn considered the earlier *Beagle* period and concluded that “Darwin as a young man is cut, at times, very much in the sublime romantic mode. In one sense this is the Humboldtian Darwin. But as a mature scientist – as we see him in the entangled bank imagery of the *Origin* – he has attained the Wordsworthian reconciliation of the sublime and the beautiful.” Kohn argued that “[t]here is an aesthetic-emotional ambition awakened on the *Beagle* that is later transformed into high scientific theory” through “the reconciliation of the sublime and the beautiful in [Darwin’s] two most striking metaphors.” Sloan has concluded, largely in agreement with Richards, that Alexander von Humboldt’s *Personal Narrative* informed at least one layer of Darwin’s concept of nature, the layer which was most prominent during the *Beagle* voyage.

Robert J. Richards’s research has continued the tendency in the literature to link Darwin to Romantic traditions, against Young’s British empiricist-Utilitarian-political economist contextualizing. Beginning in 1987, Richards endeavored to unsettle the conclusion that Darwin defended a form of Utilitarian ethics, arguing instead that

“Darwin’s early theory of moral behavior, as well as its mature expression in the Descent of Man, can best be understood…as a biologizing of [James] Mackintosh’s ethical system. … Darwin set out to give, in natural-historical and evolutionary terms, an interpretation of the faculties and relationships he found described in Mackintosh’s Dissertation.” In Chapter 3.3, I discuss Mackintosh’s moral philosophy in depth.

Richards’s thesis in Darwin and The Emergence of Evolutionary Theories of Mind and Behavior is in fact the claim that Darwin endorsed a form of moral sense philosophy. In order to reject Utilitarianism, Richards has argued that Darwin appealed to German philosophy and natural science. Beginning in 1999, Richards has defended the thesis that Darwin’s concept of nature, exemplified in the entangled bank image from the Origin of Species, was informed by Humboldt’s Romantic aesthetic. More recently, Richards has elaborated this thesis by proposing a lineage traced from nature concepts in the German Romantic and Naturphilosophie traditions to Darwin’s own mature view. This lineage replaces the conceptual relationships anthropologist Marshall Sahlins had drawn through Thomas Hobbes, Adam Smith, Darwin, and E. O. Wilson. Richards’s doctoral student, Alessandro Pajewski, has found evidence in support of Manier’s Hume-Darwin

63 Richards, Darwin and the Emergence, 115.
65 Richards, Romantic Conception of Life: Science and Philosophy in the Age of Goethe.
connection and has extended Richards’s Darwinian moral sense lineage to include Marcus Aurelius.\(^{67}\)

My review of the reception of Darwin’s theory of natural selection, particularly as he applied it to the case of human beings in the *Descent of Man*, has traced the emergence of two pairs of opposing theses in the recent literature. One thesis concerns the interpretation of Darwin’s approach to the “more interesting” problem. Young and his followers have argued that Darwin was contextually predisposed to endorse a naturalized form of quantitative Utilitarian ethics, which allegedly evolved by natural selection from the machine-like economy of nature. Richards, following Manier, has rejected this interpretation in favor of the view that Darwin perceived a German Romantic nature from which he believed a version of moral sense ethics evolved.

A second thesis, first identified by John C. Greene, concerns the interpretation of Darwin’s response to the “most intricate” problem. Weikart, following Gasman, has argued that Darwin endorsed a kind of natural selectionist perfectionism. On the intra-group level, to promote this kind of perfectionism implies approving of the suffering of the weaker members of society. On the inter-group level, this may be tantamount to a tacit endorsement of racism. Desmond and Moore, supported by most prominent Darwin scholars, have argued against this natural selectionist perfectionism interpretation by underscoring the humanitarian character of Darwin’s positive moral theory.

I claim that the tension in the literature reflects a tension within Darwin’s own view. In the following chapters, I explore this tension in depth with particular focus on

Richards’s interpretation. I evaluate his claims alongside other theses concerning the sources of Darwin’s account of the origin of ethics and his positive moral theory. As a first step, I set the narrative context for Darwin’s pursuit of the species question as well as his extensive foray into philosophy as he formed his distinct response to this question.
CHAPTER 3:
THE FORMATION OF DARWIN’S MORAL THEORY

3.1 Prologue: From Darwin’s University Years to the Notebook Period

In 1826-1827, his second year at the University of Edinburgh, Charles Darwin formed a mentoring relationship with one of his professors, the zoologist Robert Grant [1793-1874]. Under Grant’s tutelage, Darwin’s interest in natural history began to develop more deeply. The year 1827 was also when Darwin met his maternal uncle, the philosopher James Mackintosh, who would publish his Dissertation on the Progress of Ethical Philosophy a few years later. Mackintosh captured his attention despite the fact that Darwin was “as ignorant as a pig about [Mackintosh’s] subject of history, politics, and moral philosophy.”\(^6^8\) I discuss in Chapter 3.3 the history of moral philosophy and the version of moral sense philosophy Darwin found when he read the Dissertation.

During his two years at Edinburgh, Darwin began serious zoological and chemical studies. He also took Jameson’s course on natural history. His approach and interest in natural history in particular was deeply impacted by his engagement with Jameson, as well as Grant and Hope. Toward the end of his second year of studies, however, he decided to leave Edinburgh and returned to his family’s home. Shortly thereafter in 1828, he left home once again, this time for Cambridge to study for the clergy at his father’s

request. At Cambridge, Darwin befriended the botanist and geologist Rev. John Stevens Henslow [1796-1861] and the geologist Adam Sedgwick [1785-1873], both of whom built upon the foundation of Darwin’s knowledge of natural science that he had developed at Edinburgh and further encouraged his development in the field of natural history. His Cambridge years were also when Darwin read Humboldt’s *Personal Narrative* and British astronomer John Herschel’s [1792-1871] *Preliminary Discourse on the Study of Natural Philosophy*. The writings of Herschel and Humboldt influenced Darwin’s approach to empirical research and informed his perception of the natural world.69

In December 1831, the same year Mackintosh’s *Dissertation* was published, Darwin embarked on the *Beagle* voyage. He seized upon the opportunity to apply what he had learned from his mentors and reading lists while on this expedition. While at sea, Darwin worked through a new list of books, including Charles Lyell’s [1797-1875] *Principles of Geology* (1st edition). He eventually had access to all three volumes of this publication as the *Beagle* voyage continued. Most scholars agree that this text was profoundly influential on the development of Darwin’s theory of natural selection.70 Darwin was so thoroughly convinced by Lyell’s uniformitarianism and actualism that he rejected his earlier geological convictions which he had drawn from Cuvier, Jameson,

---


70 For more details concerning the influence of Charles Lyell’s geological theory on Darwin, see again Sloan, “The Making of a Philosophical Naturalist,” in the Hodge and Radick volume.
Buckland, and Owen. Yet he would ultimately depart from his predecessor’s view as he formed his own response to the question of species origin.

One of the more important episodes of the *Beagle* voyage for the formation of Darwin’s theory was the visit to the Galapagos Islands in late 1835. Darwin remarked in his *Beagle* diary on the unique forms of life he found on these islands; reflection on these curious discoveries in the subsequent years contributed to Darwin’s desire to investigate the species question. By early 1837 Darwin’s consideration of this question had begun to cast doubt on the natural history, articulated and defended by apologist and Utilitarian William Paley, which Darwin had been required to study during his years at Cambridge. Darwin’s exposure to Paley’s view while at university was his first sustained, direct encounter with the Utilitarian tradition. Paley’s *Natural Theology* developed the thesis that the natural world has arisen according to divine design. Each species originated through a special act of creation. The evidence in favor of the design thesis, Paley claimed, could be found in the fact that many flourishing forms of life exist. In the context of the human world, Paley emphasized the order of society and individual happiness as further manifestations of divine design. Thus Darwin’s first direct encounter with Utilitarianism was through an early and very popular theological version of this tradition.

Darwin collected unusual specimens while on the Galapagos, but the significance of these discoveries for natural history did not become apparent to him until after the conclusion of the *Beagle* voyage. Darwin’s collaboration with John Gould on the study of bird species was particularly influential in unsettling, for Darwin, the tenability that species have been specially created. This study with Gould suggested that the life forms
he had discovered on the remote Galapagos islands could support the notion of species transformation.71

Darwin began in 1837 to document his early considerations of the species question in a series of transmutation notebooks. The “B Notebook,” opened in July of 1837, contains his first general concerns about a possible theory of species transformation. The text and sketches in this notebook reveal Darwin grappling with whether he can claim that there is evidence of such transformation, and whether he could be justified in inferring a common ancestor of all forms of life. The B Notebook also contains Darwin’s speculations concerning the relationship between transformation and adaptation, as well as the significance of similarities in characteristics across species. The general idea of species transformation suggested that human beings might not in fact be the privileged center of creation as had been taught for centuries by the Judeo-Christian tradition. Perhaps humans also descended from the common ancestor considered in the B Notebook. Such implications illuminate Darwin’s particular focus in the notebook entries from the next year on the human species itself.

The year 1838 was apparently a watershed for Darwin’s emerging theory of natural selection, as he filled four extant notebooks with his ever-deepening and encompassing approach to the species question. In this year Darwin began to think about human origins in terms of descent by modification. Against the thesis that each human

race arose from a distinct historical origin, Darwin believed that environmental adaptation was a more reasonable explanation for the multiplicity of recognized human races. Moreover, his entries from this year disclose his conviction that descent by modification was a fruitful way to account for the origin of the distinctive features of human psychology as well as for moral characteristics. The “C Notebook,” for example, contains entries in which Darwin applied the idea of species transformation to behavioral adaptations. Darwin took bolder steps in the “M” and “N Notebooks,” where he specifically wrote about human origins and the various expressions of emotions.

First in the C Notebook and thereafter in a distinct set of notebooks, Darwin began to record the list of books he had read and those he intended to read. He occasionally included notations along with the book titles, such as the specific date when he finished reading the text or a brief remark about his overall impression of the book. Several entries in these reading notebooks consisted of works of moral philosophy, including publications in the moral sense and Utilitarian traditions. Alessandro Pajewski has provided some evidence that Darwin read from the philosophical and historical corpus of David Hume over the course of his career. The reading lists confirm that this engagement with Hume began early. In the years 1839-1851, Darwin read his *Moral and Political Essays*, *Natural History of Religion*, and the posthumously published

---


73 These years correspond to the first of Darwin’s two reading notebooks, [DAR 119].

74 [DAR 119: 11b]

75 [DAR 119: 3v]
Dialogues Concerning Natural Religion.\textsuperscript{76} Darwin also intended to read Hume’s Enquiry Concerning Human Understanding.\textsuperscript{77} Darwin documents within this notebook that he read his uncle Mackintosh’s Ethical Philosophy of 1836 around the same time he read Hume’s Essays. Later, in mid-1846, he read the three-volume edition of miscellaneous works by John Stuart Mill.\textsuperscript{78}

As I move into the next section of this chapter, I would like to draw particular attention to the fact that, during the formative months of 1838, Darwin read the Enquiry Concerning Human Understanding as well as other works by Hume.\textsuperscript{79} In what follows I examine Hume’s philosophy at length and consider its implications for the formation of Darwin’s theory of natural selection and his eventual application of this theory to the case of human beings.

3.2 Moral Sense: Darwin and David Hume’s Philosophy

Darwin’s response to the “more interesting” problem for natural selection has implications for his position within the taxonomy of moral philosophies available to him. Moreover, Darwin’s proximity to the British moral philosophies of his era presupposes an interpretation of these philosophies. Both Young’s British contextualist thesis and the Manier-Richards thesis involve significant claims about Utilitarian moral philosophy. David Hume’s [1711-1776] moral psychology and positive theory has also been

\textsuperscript{76} [DAR 119: 6a]
\textsuperscript{77} [DAR *119: 3v.]
\textsuperscript{78} [DAR 119: 17b].

\textsuperscript{79} Pajewski, “The Face of Nature,” 267-68; Pajewski references Richards, Darwin and the Emergence, 106, as well as Manier, The Young Darwin, 86, on this point.
appropriated in a specific way in defense of Richards’s interpretation. Are these assumptions legitimate?

In the remainder of this chapter, I defend my answer to this question. I claim that neither side of this tension in the literature takes into account the full complexity of Hume’s or J. S. Mill’s views to which Darwin was directly or indirectly responding when addressing the “more interesting” problem. I argue that the characterization of Darwin as a proponent of Benthamite Utilitarianism (Young thesis) and the opposing characterization of him as an anti-Utilitarian moral sense theorist (Richards’ version of Manier’s original thesis) are misleading because both characterizations oversimplify the philosophies to which Darwin was exposed. Common to both interpretations is a tendency to minimize the complex reception of Hume’s moral philosophy. The varied reception of Hume reflects a crucial ambiguity within his moral philosophy that frustrates any attempt at a simple interpretation of his position on moral issues. Moreover, both interpretations have almost entirely ignored or obscured the innovative aspects of J. S. Mill’s qualitative Utilitarianism.

I argue in the present Chapter 3.2 that a closer examination of the work of Hume suggests that his writings on moral philosophy left the status of the moral sense ambiguous. Hume’s version of the moral sense theory thus could be interpreted in at least two distinct ways. One interpretation has emphasized Hume’s conclusion that motivation arises from passions and affections (based in sense impressions) as evidence for Hume’s status as a moral sense philosopher. Another interpretation has emphasized Hume’s skeptical treatment of external objects and arguments concerning the social utility of artificial virtues as evidence that he rejected moral sense realism in favor of some form of
Utilitarianism. In part due to the mediating influence of the writings of Scottish jurist and moral philosopher James Mackintosh [1765-1832], Darwin interpreted Hume as a moral sense philosopher. I will examine at some length the treatment of Utilitarianism by Mackintosh in his Dissertation to clarify the nature and extent of Mackintosh’s disagreement with Bentham and James Mill.

Second, I will show that the vagueness of the term “Utilitarianism” has obscured the distinction between Bentham and J. S. Mill on central concepts such as utility.80 Darwin began to reflect on ethics and moral theory about a decade before J. S. Mill published his major contributions to moral philosophy, Principles of Political Economy (1848), On Liberty (1859), and Utilitarianism (1861). Yet Darwin read the younger Mill’s On Liberty the same year it was published. He marked in his reading notebook that this work was “very good.”81 Nevertheless, in the Descent of Man Darwin briefly contrasted his position on the moral sense to that of J. S. Mill. In Chapter 3.3 I begin to use these considerations as a framework for comparing and contrasting Darwin and Mill on elements of moral philosophy more rigorously than has typically been done in the literature.

In particular, I discuss in Chapter 3.3 that Darwin wrote the first edition with certain presuppositions about Utilitarianism, which John Morley criticized shortly after its publication in his 1871 review of the first edition of the Descent of Man. Morley’s Pall Mall Gazette review instigated a revealing exchange of letters between Darwin and


81 [DAR 128: 25]
his critic. Morley’s defense of J. S. Mill’s philosophy in these letters inspired Darwin to call upon the assistance of his son, William Erasmus, to consider Utilitarianism more carefully.82 His deeper engagement with the younger Mill’s views is preserved in revised and new footnotes appearing in the second edition version of Chapter 4 of the Descent.83 The footnotes suggest that Darwin attempted to preserve some elements of the younger Mill’s philosophy while also distancing his own view from Utilitarianism in general, especially its popularized forms. I contend that this complex and never fully resolved relationship to the Utilitarian tradition helps clarify the tension in Darwin’s response to the “more interesting” problem for his theory of natural selection.

A sustained review of Hume’s view on traditional moral issues is required in order to appreciate the ambiguity within his writings concerning his status as a moral sense philosopher. I proceed now to offer such a review of his view.

David Hume developed his moral psychology in the midst of his departure from the British moral rationalist and natural theology traditions. British empiricist John Locke’s [1632-28] moral philosophy is representative of the moral rationalist tradition and serves as an illuminating point of contrast in the review of Hume’s moral psychology and philosophy that follows. Darwin’s reception of Hume as a moral sense philosopher contributed to Darwin’s ability to address the “more interesting” problem of the origin of

82 In Letter 7048 dated approximately April 1871 from William Erasmus Darwin to Charles Darwin, [DAR 88: 76-7], the younger Darwin noted that J. S. Mill’s response to moral sense theory in Utilitarianism “appears muddled.”

83 In the revised footnote 5 in chapter 3 of the second edition of the Descent of Man Darwin maintains his son’s sense of confusion over J. S. Mill’s treatment of moral sense in Utilitarianism.
ethics according to exclusively natural principles without reducing all moral beliefs to mere convention.  

Whereas Darwin’s position on moral psychology is the centerpiece of his project to give a fully naturalistic account of the origin of human beings, Hume’s moral psychology was part of his own larger project in the *Treatise of Human Nature* to clarify the role and limitations of human reason. In Book 1 of the *Treatise*, Hume began his project by reducing the domain of rational knowledge to logic and certain branches of mathematics (arithmetic and algebra), and enlarging the domain of justified belief to include the natural sciences. In Book 2, Hume continued this project by establishing an allegedly absolute difference between reason and the passions within moral psychology. His goal in Book 2 was to demonstrate that reason cannot oppose a passion. Book 3 contained a more sustained attack on moral rationalist philosophy. Here Hume argued for his negative claim that a moral law cannot be the “first virtuous motive” of an action, because the first virtuous motive (which bestows merit on any action) can never be a regard to the morality of the action. The rejection of this central principle of moral rationalist philosophy created a need for something like a moral sense within a comprehensive account of moral psychology.

According to the moral rationalism of Locke, “*morally Good and Evil…is only the Conformity or Disagreement of our voluntary Actions to some Law,* whereby Good

---

84 Pajewski’s extended treatment of the Hume-Darwin link in “The Face of Nature” discusses many of the implications Hume’s philosophy had on the formation of Darwin’s position on traditional moral issues.
and Evil is drawn on us, from the Will and Power of the Law-maker.”\(^{85}\) Moreover, Locke believed that “to call a voluntary action morally good is to mark its conformity to a law which the lawmaker backs by attaching natural good to compliance and evil to disobedience, that is, by offers of rewards or threats of punishment.”\(^{86}\) Such moral determinations involve ascertaining the relevant relation that links an action to a moral law.

In *Treatise* 3.1.1.25, Hume used an example that addresses the fact that humans alone are considered moral beings, a fact that Darwin himself would address in the *Descent* in his account of the origin of ethics. In this portion of the *Treatise*, Hume contrasted moral evaluations of human incest and animal “incest” as a challenge to Locke and others who explained the human capacity to make moral distinctions in terms of knowledge of moral laws. Hume asked, “why incest in the human species is criminal, and why the very same action, and the same relations in animals have not the smallest moral turpitude and deformity?”\(^{87}\) Hume’s opponents answered questions like this by saying that we consider the human virtuous because we recognize that a human being is endowed with a faculty, reason, that “ought to restrain him to his duty,”\(^{88}\) namely, his duty not to commit that act. Thus, they conclude, when a person does commit incest, we

---


consider that act to be vicious since it indicates that the person shows no regard to the turpitude of his behavior.

In the case of animal incest, however, a philosopher like Locke would explain our neutrality towards such relations among animals by appealing to the fact that “they [the animals] have not reason sufficient to discover its [the act of incest’s] turpitude.” If the animal were rational, it would recognize the moral rule that eternally and immutably determines the moral quality of incest, and its behavior would be judged according to their regard for this rule. Hume rejected this explanation as logically circular. In order for reason to be able to discover the turpitude of incest, as moral rationalists theorists claimed, the turpitude must exist prior to its discovery. Consequently, these philosophers are committed to the view that the turpitude is “independent of the decisions of our reason, and is their object more properly than their effect.” Moreover, moral rationalists were, for the same reason, committed to the view that “every animal, must be susceptible of all the same virtues and vices, for which we ascribe praise and blame to human creatures.” In order to avoid this untoward conclusion, Hume claimed that his opponents would try to modify their position. Locke and others would assert that the distinctively moral character of human action is due to the presence and operation of human reason. But this, Hume claimed, is just to argue that the moral character of an action is, simultaneously, both prior to and the effect of reasoning.

89 Ibid.
90 Ibid.
91 Ibid.
92 Ibid.
Before moving on to Hume’s characterization of the moral sense, I wish to underscore once again the centrality of epistemological concerns within Hume’s philosophical project. It is notable that “Locke was a hedonist about motivation, holding that only prospects of pleasure and pain can motivate us,”\(^93\) since he held that “it would be utterly in vain, to suppose a Rule set to the free Actions of Man, without annexing to it some Enforcement of Good and Evil, to determine his Will,” and, therefore, that “we must, where-ever we suppose a Law, suppose also some Reward or Punishment annexed to that Law.”\(^94\) The reward or punishment, in the case of divine laws, is the prospect of procuring “happiness, or misery, from the hands of the ALMIGHTY.”\(^95\) Hume rejected elsewhere the possibility of having justified belief in, let alone knowledge of, Locke’s God. Thus Hume would also reject the possibility of feeling motivated by punishments or rewards to come in a future life. According to Hume, then, the adequacy of a rule-based moral theory (such as Locke’s) can only be evaluated in terms of its rationalistic (rather than its hedonistic) element, namely, our capacity to know moral laws and understand their demands. In this respect Hume thought Locke’s theory failed to account for pre-philosophical moral evaluations.

It is crucial to recognize that for Hume the merit-bearing motives of moral action are the passions, not the moral sense itself. Hume introduced the moral sense into his philosophy because he recognized that in order to make moral judgments human nature must include some means of access to other people’s (as well as our own) motives.


\(^{95}\) Ibid., 2.28.8.26.
Reason, on his view, cannot provide this access. This moral sense, then, which took over the role that reason played in moral rationalist philosophies, produces in us the unique moral sentiments we feel toward an agent’s motives when we observe her action. We *sense* (rather than reason) whether an action is virtuous or vicious through our awareness of an agent’s motives. Naturally virtuous actions have motives that elicit approbation from the moral sense, and naturally vicious actions have motives that elicit disapprobation from the moral sense.

So far I have shown that Hume treated moral sense as *if* it were a kind of perception, like the standard five senses. Through moral sense humans gain perceptual access to experience, which is the basis of the impressions that in turn give rise to motivating affections or passions. In the *Descent of Man* Darwin equated moral sense with conscience, which he then explicated in terms of the unique human feelings of guilt and remorse. Hence, Darwin did not sharply distinguish between affections (for Hume, the source of motivation) and moral sense (for Hume, a kind of perception). It will become clear in that section that Darwin treated moral sense as the motive of moral behavior. Richards has acknowledged this point with his claim that Darwin’s moral sense is the naturalized version of the “innate knowledge” James Mackintosh posited in his own moral theory in the *Dissertation*. Yet this interpretation of Darwin’s moral sense does not seem to support the strong link between Hume and Darwin that Pajewski has recently advanced, at least not if my interpretation of Hume is acceptable.

Describing and defending the moral sense posed a problem for Hume from the beginning. Moral sense realism stands in tension with Hume’s skeptical epistemology. Hume seems to have believed that this sense is real, but the arguments I have discussed
imply that it cannot be real in any standard manner. It cannot be characterized in terms of relations of objects, events, or actions, nor in terms of matter of fact. In response to this problem, Hume chose to compare moral sentiments to secondary qualities, so that the operation of the moral sense could be understood by analogy with the primary-secondary quality distinction.96

In the same section of the Treatise in which Hume argues that reason cannot make moral distinctions, for example, he acknowledges the importance of the primary-secondary quality distinction, especially in modern natural philosophy. Furthermore, he suggests that this distinction can be helpful in ethics to clarify our feeling of evil from the claim that evil exists independently of us in the world. “Vice and virtue, therefore,” Hume says, “may be compar’d to sounds, colours, heat and cold, which, according to modern philosophy, are not qualities in objects, but perceptions in the mind.”97 Even in the second Enquiry Concerning the Principles of Morals Hume continues to explain our distinctively moral feelings as analogous to secondary qualities. In that later work he directly contrasts reason, which is the capacity by which a person “discovers objects as they really stand in nature, without addition or diminution,” with taste, which “has a productive faculty, and gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises in a manner a new creation.”98


Hume’s “solution” to the problem of describing the moral sense and its operation introduced a new problem. In Book 1 of the *Treatise*, Hume explicitly rejected the primary-secondary quality distinction, effectively denouncing it as bad metaphysics. Any attempt to establish this distinction by argument, as Locke did,\(^9\) is logically “to run in a circle, and make one idea depend on another, while at the same time the latter depends on the former.”\(^10\) Furthermore, Hume’s theory in Book 1 denied the possibility genuinely knowing that what Locke and other early modern philosophers called “primary qualities” are actually features of the external, physical world. The distinction between primary and secondary qualities thus collapses. All qualities, for Hume, turn out to be so-called secondary qualities.

Hume’s description of the moral sense threatened to undermine the consistency of the theory of human nature that he developed throughout the entire *Treatise* and maintained in both *Enquiries*. It is precisely this point of ambiguity that opened up space within the subsequent reception of Hume’s moral theory for alternative interpretations that ultimately contributed both to the continuation of the moral sense tradition and to a naturalized form of Utilitarianism.\(^1\) Pajewski and Richards have aligned Darwin with the moral sense realist interpretation of Hume without much discussion of this ambiguity

---

\(^9\) See, for example, Locke’s manna argument in *Essay Concerning Human Understanding*, 2.8.18. Maintaining this distinction was very important to Locke and other early moderns. It was used as a way to preserve the existence of an external world, as well as source of causal explanations of our perceptions of material objects.

\(^10\) Hume, *Treatise on Human Nature*, 1.4.4.10.

– as if to say that Hume’s philosophy was unambiguous on the issue of the ontological status of the moral sense. This smooth alignment of Darwin and “Hume-the-moral-sense-realist” obscures the fact that Hume was also an influence on nineteenth-century Utilitarians. It likewise ignores the fact that a significant influence on the development of Darwin’s own moral philosophy, the writings of James Mackintosh, contained a praise of the Utilitarian element of Hume’s philosophy while also expressing strong criticism of the Utilitarian tradition as established by Jeremy Bentham and James Mill. In Section 3.3, I examine Mackintosh’s stance in somewhat greater depth in preparation for discussing Darwin’s relationship to Hume and J. S. Mill.

3.3 Mackintosh on Benthamite Utilitarianism

Robert Richards has attempted to distance Darwin from Utilitarianism by arguing for the significant influence of James Mackintosh on the formation of Darwin’s moral theory. Upon closer examination, however, Mackintosh appears strikingly ambivalent toward moral concepts of utility. On the one hand, as I discuss at length in this section, Mackintosh admonishes Jeremy Bentham and his followers concerning the greatest happiness principle as a moral motive. On the other hand, in the midst of developing his own eclectic philosophy synthesized from elements of philosophies summarized in the Dissertation, Mackintosh praised the “proof” in Hume’s Enquiry on the Principles of

102 See Rosen, Frederick, “Reading Hume Backwards,” 29-30 for the claim that Hume influenced Jeremy Bentham’s concept of utility.

Morals, “that all the qualities and actions of the mind which are generally approved by mankind agree in the circumstance of being useful to society.”

Mackintosh admired Hume for concluding that “benevolent affections and actions” tend to promote social cohesion. The claim that “general utility constitutes a uniform ground of moral distinctions” is, according to Mackintosh, “a part of Mr. Hume’s ethical theory which never can be impugned, until some example can be produced of a virtue generally pernicious, or of a vice generally beneficial.” Mackintosh dismissed the epistemological concerns that motivated Hume to support these assertions with argument, claiming that Hume’s effort to prove them is “scarcely necessary.” The value of Hume’s philosophy was thus to be found in its defense of the social utility of virtue.

Mackintosh believed, in agreement with Hume to an extent, that humans are “naturally disposed” to approbate benevolence and justice. Moreover, he claimed that we cultivate our natural moral sentiments because they “materially dispose others, as well as ourselves, to cultivate these two virtues” that maintain society.

This broadly Utilitarian interpretation of Hume was understandably generally appealing to a natural historical realist such as Darwin, who, unlike Hume, was not

---

104 James Mackintosh, *Dissertation on the Progress of Ethical Philosophy Chiefly During the Seventeenth and Eighteenth Centuries*, 3rd ed, (Edinburgh, Scotland: Adam and Charles Black, 1862), 177.


106 Ibid., 177.

107 Mackintosh, *Dissertation*, 182.
concerned with epistemological problems. Moreover, even for Hume, who conceived of human nature as constant across generations, the moral sense is to a significant extent dynamic and developmental. The moral sense is influenced and transformed over the course of an individual lifetime through education as well as through cultural context (which is itself subject to change).

The developmental character of Hume’s moral sense readily leant itself to reinterpretation according to a perspective like Darwin’s in which nature itself is dynamic and human psychology emerges from properties present in lower forms of life. Darwin also read a highly critical interpretation of traditional Benthamite Utilitarianism in Mackintosh’s Dissertation. This criticism informed his account of moral origins and affected the positive moral theory expounded in Chapter 4 of the Descent. Mackintosh accused Jeremy Bentham of failing to recognize the very distinction Hume taught: the difference between moral approbation (which is the domain of moral sense) and moral qualities (virtuous habits grounded in the passions). He concluded that this confusion on Bentham’s part “led him to assume, that because the principle of utility forms a necessary part of every moral theory, it ought therefore to be the chief motive of human conduct.”

Notably, Mackintosh did not challenge Bentham on the coherence of the claim that a “desire to promote the welfare of men in general” can be a motive for action. Instead, he challenged him on the effectiveness of this motive relative to other possible motivations.

---


109 Mackintosh, Dissertation, 244.
for “useful actions.”  Hume most likely would have criticized Bentham differently—probably from an epistemological angle—than Mackintosh actually did in this passage.

Mackintosh criticized Bentham in this portion of the *Dissertation* as part of his effort to defend the claim that general concerns are not maximally effective motives. The force of his argument hinges on two appeals to associationist psychology. Associationism holds that the more remote an object is from a person’s experience, the less vivid that person’s corresponding impression of that object will be—hence, the passion she experiences alongside this impression will be weaker and less capable of motivating her to act. Therefore, “even if it were the interest of every man to be bold, it is clear that so cold a consideration cannot prevail over the fear of danger. Where it seems to do so, it must be by the unseen power either of the fear of shame, or of some other powerful passion, to which it lends its name.” Moreover, Mackintosh explained the appeal of Bentham’s philosophy by observing that whereas virtuous moral qualities and their underlying merit-bearing passions are not easily distinguished by a spectator, “the outward advantages…, cold, uncertain, dependent and precarious as they are, yet stand out to the sense and to the memory…perfectly on a level with the general apprehension.” It is easier for a spectator to consider moral qualities against useful consequences when attempting rationally to decide upon the best form of moral education, for example, or how to maximize justice in society.

---

110 Ibid., 245.
111 Ibid.
112 Ibid., 246, original emphasis.
113 Mackintosh, *Dissertation*, 251.
Mackintosh allowed that “the pursuit of our own general welfare” (i.e., self-love) and “that of mankind at large” (i.e., benevolence) play crucial secondary roles in moral life. Yet “from their vagueness and coldness they are unfit habitual motives and unsafe ordinary guides of conduct,” that is, they should not be promoted to a primary role. Nevertheless, benevolence “enables us to look on all men as brethren, and overflows on every sentient being,” which is precisely the point that Darwin made while outlining his positive moral theory in the *Descent of Man*.

In the following two sections of this chapter I introduce J. S. Mill’s moral philosophy and begin to compare his view with that of Hume and Darwin. This comparison provides a valuable perspective from which to evaluate the legitimacy of Darwin’s application of Mackintosh’s criticism to the case of J. S. Mill’s distinctive version of Utilitarianism. It also has implications for the adequacy of Richards’s interpretation of the proximity of the moral theories articulated by Darwin and the younger Mill.

3.4 John Stuart Mill: From Bentham to Coleridge and Goethe

John Stuart Mill published his major discussion of moral theory, the *Utilitarianism* of 1863, after sustained, serious engagement with Romanticism and in the midst of reading the first edition of Darwin’s *Origin of Species*. Mill first appealed to the Romantic poets during the defining emotional crisis into which he entered not long after he completed the rigorous course of study prescribed by his father, the Benthamite philosopher James Mill. The *Autobiography* of J. S. Mill, written in 1865, provides

114 Ibid., 248.
insight into this critical influence. Mill’s positive relationship to the Romantic tradition calls into question Richards’s claim that the specifically Romantic foundation of Darwin’s moral theory “overturned” the Utilitarianism to which Darwin responded in the *Descent*.

Moreover, in Mill’s correspondence with his contemporary philosophers and naturalists he shared his reaction to Darwin’s first-edition statement of the theory of natural selection. He would later include this reaction in a footnote he added to the fifth edition of the *System of Logic*.\(^{115}\) Examination of these writings shows that Mill was in fact an enthusiastic supporter of the philosophical form of Darwin’s theory. I take this consideration as guidance in my attempt to specify more rigorously the differences between Darwin and Mill on the origin and progress of ethics.

In his 1872 biographical sketch, Justin McCarthy remarked on the younger Mill’s distinctive combination of logical rigor and emphasis on sentiment:

> There is something fascinating, moreover, about the singular blending of the emotional, and even the romantic, with the keen, vigorous, logical intellect, which is to be observed in [John Stuart] Mill. Even political economy, in Mill’s mind, is strangely guided and governed by mere feeling.\(^{116}\)

The “logical” aspect of J. S. Mill’s philosophy can be traced to the influence of the Benthamite moral calculus transmitted to him through his father. Undoubtedly, his infamously demanding and thorough education also contributed to his sharp reasoning


skills. The Romantic aspect of the younger Mill’s philosophical approach, however, has an altogether different origin.\textsuperscript{117}

In his \textit{Autobiography} the younger Mill specified the multiple layers of influence the English Romantics exerted upon him during his period of emotional darkness. First, he found in the poet Samuel Taylor Coleridge the words to describe his feelings when “the whole foundation on which my life was constructed fell down.”\textsuperscript{118}

\begin{flushright}
A grief without a pang, void, dark and drear,\\
A drowsy, stifled, unimpassioned grief,\\
Which finds no natural outlet or relief\\
In word, or sigh, or tear.\textsuperscript{119}
\end{flushright}

The “whole foundation” which Mill felt he lost, and to which Coleridge’s verse spoke, was his passion “to be a reformer of the world” in accordance with the philosophy of Bentham he first read in 1821.\textsuperscript{120} This reformist philosophy rested on the foundation of associationism, which asserted that all mental and moral feelings and qualities, whether of a good or of a bad kind, were the results of association; that we love one thing, and hate another, take pleasure in one sort of action or contemplation, and pain in another sort, through the clinging of pleasurable or painful ideas to those things, from the effect of education or of experience. As a corollary from this, I had always heard it maintained from my father, and was myself convinced, that the object of

\begin{footnotesize}


\textsuperscript{120} J. S. Mill, “A Crisis in My Mental History.”
\end{footnotesize}
education should be to form the strongest possible associations of the salutary class; associations of pleasure with all things beneficial to the great whole, and of pain with all things hurtful to it.\textsuperscript{121}

This foundation crumbled beneath him because his education “had failed to create these feelings in sufficient strength to resist the dissolving influence of analysis, while the whole course of my intellectual cultivation had made precocious and premature analysis the inveterate habit of his mind.”\textsuperscript{122}

Second, William Wordsworth provided Mill with a poetic image of the human possibility for renewal and growth:

What made Wordsworth’s poems a medicine for my state of mind, was that they expressed, not mere outward beauty, but states of feeling, and of thought coloured by feeling, under the excitement of beauty. They seemed to be the very culture of the feelings, which I was in quest of. In them I seemed to draw from a Source of inward joy, of sympathetic and imaginative pleasure, which could be shared in by all human beings; which had no connexion with struggle and imperfection, but would be made richer by every improvement in the physical and social condition of mankind. From them I seemed to learn what would be the perennial sources of happiness, when all the greater ills of life shall have been removed. And I felt myself at once better and happier as I came under their influence.\textsuperscript{123}

In this respect, the Romanticism of Wordsworth contributed to Mill’s ability to emerge from his emotional crisis.

Moreover, third, these kind of Romantic intimations of human possibility facilitated Mill’s innovative movement away from the quantitative Utilitarianism of

\textsuperscript{121} J. S. Mill, “A Crisis in My Mental History.”
\textsuperscript{122} Ibid.
\textsuperscript{123} Ibid.
Bentham and the elder Mill. On the one hand, the younger Mill’s deep engagement with poetry convinced him that Benthamite Utilitarianism was grounded in principles that were too narrow. In particular, Bentham and James Mill failed to appreciate the significance of art for moral character and political life. Due to J. S. Mill’s openness to a fuller range of human experience,

[the influences of European, that is to say Continental, thought, and especially those of the reaction of the nineteenth against the eighteenth, were now streaming in upon me. They came from various quarters: from the writings of Coleridge, which I had begun to read with interest even before the change in my opinions; from the Coleridgians with whom I was in personal intercourse; from what I had read of Goethe; from Carlyle’s early articles in the Edinburgh and Foreign Reviews, though for a long time I saw nothing in these (as my father saw nothing in them to the last) but insane rhapsody.]

On the other hand, the quantitative Utilitarian tradition in which he had been educated in his youth was not entirely false. It could be rehabilitated and transformed. J. S. Mill’s aim was thus “to shew that there was a Radical philosophy, better and more complete than Bentham’s, while recognizing and incorporating all of Bentham’s which is permanently valuable.”

Among the “permanently valuable” elements of Bentham’s philosophy was the recognition that “happiness is the test of all rules of conduct, and the end of life.”

124 J. S. Mill, “A Crisis in My Mental History.”

125 Ibid.


127 J. S. Mill, “A Crisis in My Mental History.”
is the greatest happiness principle as the moral criterion. Yet the younger Mill developed an understanding of “happiness” that was quite distinct from the orthodox Utilitarian concept:

I now thought that this end was only to be attained by not making it the direct end. Those only are happy (I thought) who have their minds fixed on some object other than their own happiness; on the happiness of others, on the improvement of mankind, even on some art or pursuit, followed not as a means, but as itself an ideal end.128

This externally-directed and even humanitarian sense of happiness was the foundation of J. S. Mill’s qualitative distinction among the pleasures. Mill gleaned this qualitative, aesthetic sense precisely from the Romantics that his forebears in the Utilitarian tradition excluded from consideration. His essays on Bentham and Coleridge are testament to the synthesis he endeavored to forge during the rest of his philosophical career.129

This distinctive blend of traditional Utilitarianism and Romanticism incubated in J. S. Mill for decades before he read Darwin’s first-edition Origin of Species within a year of its publication. These potent influences likely informed Mill’s reaction to Darwin’s theory; he most certainly judged it according to his own logically rigorous yet Romantic perspective. Moreover, just prior to reading the Origin, Mill had finished his draft of Utilitarianism. Thus it is notable that, upon reading the Origin, Mill wrote to Alexander Bain in July 1860, the year after the first edition of Darwin’s text was

128 J. S. Mill, “A Crisis in My Mental History.”

published, to share his very positive appraisal of the theory of natural selection. “I have read since my return here, several things which have interested me, above all Darwin’s book,” he wrote to Bain,

It far surpasses my expectation. Though he cannot be said to have proved the truth of his doctrine, he does seem to have proved that it may be true which I take to be as great a triumph as knowledge & ingenuity could possibly achieve on such a question. Certainly nothing can be at first sight more entirely implausible than his theory & yet after beginning by thinking it impossible, one arrives at something like an actual belief in it, & one certainly does not relapse into complete disbelief.130

The following year, in 1861, Mill included a similarly praiseworthy footnote on Darwin in the substantially revised fifth edition of his System of Logic. In this mature statement of his principles of inductive reasoning, the younger Mill reflected:

Mr. Darwin’s remarkable speculation on the Origin of Species is another unimpeachable example of a legitimate hypothesis. What he terms “natural selection” is not only a vera causa, but one proved to be capable of producing effects of the same kind with those which the hypothesis ascribes to it: the question of possibility is entirely one degree. It is unreasonable to accuse Mr. Darwin (as has been done) of violating the rules of Induction. The rules of Induction are concerned with the conditions of Proof. Mr. Darwin has never pretended that his doctrine was proved. He was not bound by the rules of Induction, but by those of Hypothesis. And these last have seldom been more completely fulfilled. He has opened a path of inquiry full of promise, the results of which none can foresee. And is it not a wonderful feat of scientific knowledge and ingenuity to have rendered so bold a suggestion, which the first

impulse of every one was to reject at once, admissible and discussable, even as a conjecture?\(^{131}\)

Thus, the form of Darwin’s theory of natural selection, which he would apply to the “more interesting” and “most intricate” problems a decade after this exchange of correspondence, was an exemplary instance of Mill’s principles of logic.

What about the content of Darwin’s application of his theory to the human species? Did Mill believe that Darwin’s moral theory similarly exemplified his own Romantic understanding of moral character and human possibility? In the following chapter I compare several elements of the moral philosophies of Darwin, Hume, and J. S. Mill in order to clarify the proximity or distance of their views from each other.

CHAPTER 4:
A COMPARISON OF THE MORAL PHILOSOPHIES OF HUME, J. S. MILL, AND DARWIN

4.1 Prologue: Darwin’s Reaction to J. S. Mill’s Appraisal of Natural Selection

The same year that J. S. Mill revised his *System of Logic* to include the highly praiseworthy footnote about the logical form of Darwin’s theory, the statesman and Cambridge political economist Henry Fawcett wrote a letter to Darwin. Fawcett himself had written an article in *Macmillan’s* praising Darwin for his logically rigorous scientific theorizing.¹³² He was also a friend of John Stuart Mill. In this letter to Darwin from mid-1861, Fawcett disclosed that Mill “considers your reasoning throughout is in the most exact accordance with the strict principles of Logic. He also says, the method of investigation you have followed is the only one proper to such a subject.”¹³³

A mere five days later, Darwin wrote a letter to American botanist and Harvard natural historian Asa Gray in which he mentioned Mill’s overwhelmingly positive appraisal of his own theory. This can be interpreted as a strategic move on Darwin’s part, for Gray himself had a complicated initial reaction to Darwin’s position on the species question. On the one hand, Gray was a proponent of natural selection. On the other hand,


¹³³ [DAR 98: B29-30].

50
he wished to preserve the more traditional concept of design from Paleyean natural theology. Darwin disagreed with Gray on the latter point; he mentioned Mill’s appraisal in the midst of addressing this disagreement, apparently in an effort to persuade Gray to abandon his allegiance to natural design. Darwin then reflected, “Is it not…very satisfactory to me?” again in reference to Mill’s praise.

Why was the philosopher J. S. Mill’s opinion so important to Darwin? The reading lists suggest that Darwin did not read the younger Mill’s works nearly as extensively as he did the works of Hume. Nevertheless, Darwin’s less thorough engagement with Mill was nevertheless apparently influential and strategically timed. Darwin read On Liberty the same year it was published, which was also the year the first edition of his own Origin of Species was published: 1859. Next to the title of Mill’s small book and the date on which he completed it, Darwin wrote in his reading notebook that this was a “very good” work of political philosophy.134

In the remainder of Chapter 4, I discuss the comparison of Hume, J. S. Mill, and Darwin on more fundamental concepts of moral philosophy.

4.2 Prologue: Darwin’s Reaction to J. S. Mill’s Appraisal of Natural Selection

My review of Hume’s moral epistemology in Chapter 3.2 showed that, for Hume, humans discover moral good and evil through the moral sense – a sense which is a kind of perception of motive (i.e., passion) that enables human emotional responsiveness (i.e., sentiments) manifesting as moral approbation and disapprobation. Moreover, I showed that Hume’s moral sense epistemology is best understood as a direct response to the

134 [DAR 128: 25]
moral rationalists of his era, e.g., Locke, who believed that humans discover moral good and evil through reason alone. These conclusions guide the examination of Hume’s moral philosophy that I resume in the present section.

Hume developed his arguments in order to establish that the passions perceived through the moral sense are the foundation of moral judgment, but he did permit reason a secondary role. Reason discovers situational facts that individuals take into consideration when deliberating on a future action or defending an emotional response to past actions. Through reason an individual can also discern the utility, in the long run, for society if many individuals engage in certain forms of behavior. Later in the paper I discuss the way in which Darwin was especially concerned with this latter role of reason in his own notes as well as his mature moral theory in the *Descent of Man*.

Within the framework of Hume’s philosophy, actions are external manifestations of the inner, motivating passion of an agent. An agent’s behavior and character are rightfully evaluated on the basis of the moral sentiments of a spectator, for which such passion is an object. These sentiments are not only *immediate*, insofar as they are felt prior to any exercise of reason; they are also *universal*. Hume thus believed that the utility of an action is only a secondary concern for spectators, whereas this evaluative criterion may take precedence in reflective judgment. Indeed, Hume attempted to justify the artificial virtues in terms of social utility, as Mackintosh praised in his *Dissertation*.

J. S. Mill emphasized the utility of moral action in *Utilitarianism* because he was concerned with the lack of uniformity of emotional responsiveness across individuals. This lack of uniformity was an obstacle for creating the optimal social conditions that the younger Mill wished to create through his philosophy. The central task for moral
philosophy, therefore, was not to clarify the source of moral knowledge and the process of moral judgment, but rather to determine how to achieve the most cultivated form of emotional responsiveness in all members of a society. Hume’s philosophy, which centered on an appeal to moral sense and which explicated moral judgment in terms of second-order passions, did not adequately address these concerns:

The difficulty is not avoided by having recourse to the popular theory of a natural faculty, a sense or instinct, informing us of right and wrong. For – besides that the existence of such a moral instinct is itself one of the matters in dispute – those believers in it who have any pretentions to philosophy, have been obliged to abandon the idea that it discerns what is right or wrong in the particular case in hand, as our other senses discern the sight or sound actually present. Our moral faculty, according to all those of its interpreters who are entitled to the name of thinkers, supplies us only with the general principles of moral judgments; it is a branch of our reason, not of our sensitive faculty; and must be looked to for the abstract doctrines of morality, not for the perception of it in the concrete.135

Mill traced the lack of uniformity in moral judgment to upbringing and education, as well as to the framework of a society. Like his forebears in the Utilitarian tradition, he believed that only a rational principle could adjudicate in disagreements about the moral valence of an action. Appealing to a rational principle as a criterion for evaluation should not be conflated with claiming that moral behavior is motivated by reason. Nevertheless, critics have consistently misinterpreted Utilitarians on this issue throughout the history of this tradition of philosophy.

In fact, J. S. Mill was about as concerned with undermining moral rationalism as was David Hume. As I discussed in Section IB, moral rationalism in Hume’s day consisted of an appeal to some kind of moral law discoverable and applied by reason. At the time of J. S. Mill’s work, however, British moral rationalists appealed to the transcendental philosophy of Immanuel Kant. Kant’s philosophy was under discussion in nineteenth-century Britain through the exposition of authors such as Henry Mansel and Sir William Hamilton. The British sympathizers of Kant were especially interested in his particular approach to defending the claim that universal reason recognizes the moral law. Kant’s practical philosophy grounds moral judgment in intuition. The Scottish “intuitionists” appropriated Kantian intuition as an additional element of moral sense philosophy, ultimately with an aim to support the claim that humans have innate knowledge of morality. The claim that humans have innate moral knowledge was, for J. S. Mill, tantamount to a defense of moral primitivism, the view that humans are naturally morally good.

Mill rejected primitivism because it failed to acknowledge the need to create the conditions for human progress through social and legislative reforms. Thus he inclined even further than Hume toward an individually (as opposed to historically) developmental account of moral sense, specifically in opposition to the intuitionists. In this way, moral character and its development becomes central to Mill’s philosophy, as opposed to the claims of moral sense ethicists and their exposition of the process of moral judgment. I discuss in the following section the extent to which Hume and J. S. Mill differed in their estimation of the capacity for this development. Darwin’s moral theory is a distinct third way of accounting for the source of moral knowledge as well as the
development of moral character; due to its grounding in natural historical perspective, Darwin’s view cannot be reduced to the accounts defended by Hume or Mill.

J. S. Mill’s sharp disagreement with the moral rationalists (“intuitionists”) had significant implications for his use of words such as “natural” and “innate” in his account of the development of moral character:

…if, as is my own belief, the moral feelings are not innate, but acquired, they are not for that reason less natural. It is natural to man to speak, to reason, to build cities, to cultivate the ground, though these are acquired faculties. The moral feelings are not indeed a part of our nature, in the sense of being any perceptible degree present in all of us… Like the other acquired capacities above referred to, the moral faculty, if not a part of our nature, is a natural outgrowth of it; capable, like them, in a certain small degree, of springing up spontaneously; and susceptible of being brought by cultivation to a high degree of development. Unhappily it is also susceptible, by a sufficient use of the external sanctions and of the force of early impressions, of being cultivated in almost any direction: so that there is hardly anything so absurd or so mischievous that it may not, by means of these influences, be made to act on the human mind with all the authority of conscience.136

The concept of “innate” refers to the so-called “intuitionist school” of morality. Against this concept of innate, Mill emphasized that humans must first be taught to discern the moral value of actions and qualities that collectively we uphold as good or evil. Only on the basis of such moral education do we become genuinely moral and subject to moral judgment. Apart from particular forms of education and upbringing, humans are non-moral. Nevertheless, the fact that we can be taught to practice and appreciate virtue means, for Mill, that humans have a natural capacity for moral behavior. The collection

of moral principles inculcated through education and upbringing is, in turn, preserved and augmented *culturally* through oral and written tradition. Notably, Mill did not entertain the possibility that moral development could proceed *biologically* across generations, despite his slightly earlier and very positive encounter with Darwin’s theory of natural selection through his reading of the *Origin*.

When Darwin eventually applied his theory of natural selection to human origins in the *Descent*, his account included the claim that morality develops historically, across generations. According to this account, an individual human traces, in development, the history of the evolution of moral psychological properties in the human species. Thus, the principle “ontology recapitulates phylogeny” applies even to the moral aspect of human beings. This principle simultaneously accounted for the origin of ethics and the individual development of moral character. Yet this principle could only give a limited account of individual moral character development. Only those moral habits that had already been preserved as instinct through natural selection could be explicated as innately recapitulated. Beyond the range of moral instinct, however, character must be deliberately cultivated through education and enculturation. Darwin thus explained the origin of ethics in terms of *biological* preservation and augmentation across generations of moral psychology and moral habit.

In contrast, Mill claimed only that humans have an innate capacity for moral character development. He left the origin of this innate capacity unexplained. Perhaps he did not appeal to Darwin’s theory of natural selection (or to some other theory) as an explanation because he denied that there is any inherent directionality to character development. There is a direction of moral development that is more conducive to human
happiness, but no innate tendency for humans to develop in that direction. Hence the
great importance of culture, society, and individual upbringing in Mill’s philosophy:
these factors alone mould character over the course of development. Darwin’s mention of
Mill in footnote five of the third chapter in the first edition of the Descent connects to
their disagreement on this issue.

In the midst of his review of the first edition of the Descent in the Pall Mall
Gazette, John Morley criticized Darwin for his brief yet critical treatment of Mill’s view
on the moral sense. Darwin wrote a personal letter in response to Morley in which he
suggested that he was more perplexed by J. S. Mill’s moral philosophy than strictly
opposed to it. Darwin placed the responsibility on Mill, whom he claimed simply lapsed
incoherent when denying that “natural” feelings are “innate”:

With respect to Mr. Mill, nothing would have pleased me
more than to have relied on his great authority with respect
to the social instincts, but the sentence which I quote at p.
71 [of the first edition of the Descent of Man] (“if, as is my
own belief, the moral feelings are not innate but acquired,
they are not for that reason less natural”) seem to me
somewhat contradictory with the other words which I quote
[i.e., Mill’s characterization in Utilitarianism of the social
feelings as a “powerful natural sentiment,” and as “the
natural basis of sentiment for utilitarian morality”], so that I
did not know what to think; more especially as he says so
very little about the social instincts.137

Darwin’s reaction to Mill in the first edition of the Descent of Man is somewhat
unsurprising, as the other Utilitarian authors (such as Bain and Sidgwick) that he more
typically read affirmed simplistic, and, in Darwin’s eyes, “extremely improbable” views.

137Charles Darwin to John Morley, Down, April 14, 1871, in More Letters of Charles Darwin: A Record of
His Work in a Series of Hitherto Unpublished Letters, ed. Francis Darwin and A. C. Seward (London,
England: John Murray, 1903), 327.
For example, Bain held that “the moral sense is acquired by each individual in his lifetime,” and many others “write as if there must be a distinct motive for every action, and that this must be associated with some pleasure or displeasure.” Yet Darwin explicitly mentioned Mill as the exception to both of these egregious faults. Mill’s position was closer than other Utilitarians to Darwin’s own position, and yet Darwin was perplexed by Mill’s failure to biologize moral sense as he himself had done.

Morley quickly responded to Darwin with a letter that was on the whole respectful but nevertheless an incisive clarification of Mill’s position in *Utilitarianism*. He drew attention to a portion of the passage from Mill’s text that Darwin somehow overlooked in the midst of his perplexed yet critical remarks in footnote five of the *Descent*. He explained:

I don’t think Mr. Mill’s expression in pp. 45 and 46 [of *Utilitarianism*] points to any fundamental difference between him and yourself. He admits that the moral faculty is capable of springing up “spontaneously” in “a certain small degree” (p. 45 of *Utilitarianism*), and this is as much as you want, is it not? Perhaps he would hold that the great force of association over the principles of human nature come into effective bearing, at a somewhat earlier stage than you would be inclined to allow. But I see nothing in what you have written to make me suppose that you would deny the preponderant power of association in shaping the social instinct, granting that to be an ultimate and insoluble fact in human nature.138

Morley rejected Darwin’s move from the observation that moral behavior is frequently spontaneous to the conclusion that moral sense is innate. Morley emphasized that even Mill’s a-historical moral psychology acknowledged “a certain small degree” of moral spontaneity. Perhaps social instinct – which, according to Mill, is “an ultimate and

---

138 [DAR 88. 71-72].
insoluble fact in human nature” – could be explicated as simply the sympathetic awareness of the passions of other humans, combined with the tendency to live in groups. This general heritable constitution would then be given shape over the course of an individual lifetime by education and cultural influence; in other words, that “the great force of association” comes into effect in each human life “at a somewhat earlier stage” than Darwin himself allowed in the Descent. Associations influenced by cultural and familial context would, on this alternative view, begin to form so early in individual human lives that we might – especially from a natural historical perspective like Darwin’s – fall into the illusion that moral and cultural development is transmitted biologically. Moral and cultural developments might instead be transmitted through other naturally-arising yet non-biological means.

Darwin’s immediate response to Morley’s attempt to reconcile the notion of social instinct with J. S. Mill’s version of Utilitarian moral philosophy remains rather mysterious. Darwin did not continue corresponding with Morley on this point of contention. Nevertheless, the legacy of Darwin’s written exchange with Morley can be found in the revised version of footnote five in chapter four of the second edition of the Descent of Man. Darwin did not remove any of his earlier remarks in the revised footnote, but he did add two sentences.

One of the sentences specifically quotes the remark from Mill to which Morley had drawn attention in his second letter: “Again he [J. S. Mill] writes, ‘Like the other acquired capacities above referred to, the moral faculty, if not part of our nature, is a natural out-growth from it; capable like them, in a certain small degree of springing up spontaneously.’” He aligned this remark with Mill’s acknowledgement that social feelings
are “a powerful natural sentiment” and are “the natural basis of sentiment for utilitarian morality,” yet maintained that such a view is inconsistent with Mill’s additional claim that moral sense is not innate. Thus, with qualifications, Darwin continued to affirm in the second edition against Mill that human morality is innate and that moral development is biological. But the disagreement is complex, and the degree of separation between Mill’s and Darwin’s views on this issue is not nearly as sharp as much of the literature has suggested.

I mentioned in Chapter 2 that, decades before expressing his mature view in the Descent of Man, Darwin began sketching out his ideas about moral origins in his Notebooks. An illuminating entry from the N Notebook in late 1838 clarifies Darwin’s views on the general Utilitarian philosophy prior to reading any of Mill’s works:

Two classes of moralists: one says that our rule of life is what will produce the greatest happiness. – The other says we have a moral sense. – But my view unites both /& shows them to be almost identical /& what has produced the greatest good /or rather what was necessary for good at all /is the /instinctive /moral sense: (&this alone explains why our moral sense points is to revenge). In judging the rule of happiness we must look far forward /& to the general action /-- certainly because it is the result of what has generally been best for our good far back – (much further than we can look forward: hence our rule may sometimes be hard to tell).\(^\text{139}\)

Young Darwin believed that taking a natural historical approach to the question of moral foundations could reconcile Utilitarianism and moral sense theory. Thus, contrary to Richards, I contend that Darwin was not trying to “overturn” Utilitarianism when he speculated on the application of species transformation to the question of human origins.

\(^{139}\) Charles Darwin, “Entry 30 (October 2d. 1838),” in Old and Useless Notes about the Moral Sense & Some Metaphysical Points (1838-1840), ed. Paul H. Barrett, [DAR91. 4-55].
Darwin persisted in his endeavor to reconcile and subsume these two philosophies even into *Descent* period, despite the significant departures Mill had made from the Benthamite Utilitarianism to which Darwin was exposed in his younger years.

The key lies in the temporality of the two perspectives. Darwin attempted to draw attention to the element of time by underscoring the tensed words in his notebook entry. Moral sense philosophies like Hume’s tended to be a-historical, and were thus content to account for moral psychology as it was observed in the present. Hence, versions of this philosophy were expressed in the present tense. In contrast, versions of Utilitarianism tended to take a more temporal approach. Insofar as they were future-oriented in their concern for maximizing happiness, this philosophical tradition would endorse the use of reason on the basis of past experience to decide on the most beneficial courses of action to take in the future. Yet the very long-term consequences of potential courses of action on a collective level – such as legislative changes that affect human welfare – could only be evaluated properly if very remote history and its consequences were likewise taken seriously as a basis of comparison.

At this early stage in his career represented by the Notebooks, Darwin believed that a natural historical perspective could improve upon Utilitarianism by replacing the greatest happiness principle with the principle of natural selection. Natural selection was the superior principle because it offered a vastly lengthened line of sight into the past. Assuming that Utilitarianism must draw on past experience to make judgments about maximizing happiness in the future, a glimpse into the primordial past promised unprecedented potential for making determinations about consequences far in the future. Similarly, young Darwin speculated in this Notebook entry that natural selection
provided moral sense psychology with an empirical foundation. Moral sense had always been the necessary condition for moral good; prior to the emergence of moral sense there was no moral good.

According to Darwin, the philosophy of utility and the philosophy of moral sense were “almost identical” because the same principle – natural selection – is their common foundation. Experience is also a crucial link between these two moral philosophies. I have discussed J. S. Mill’s claim that the ability to adjudicate in moral disagreements and to make sophisticated decisions about future courses of action is based in the cultivation of moral character through education and upbringing. Darwin himself emphasized that these means of cultivation are forms of experience. Likewise, moral sense or conscience\textsuperscript{140} is an instinctive feeling, and “acquired instincts [are] analogous to /& replace/ to experience gained by man in lifetime.” Moral sense as it exists at any time is also capable of further cultivation through education and upbringing (i.e., additional formative experience), which, if advantageous, would eventually become heritable. In fact, even for the young Darwin, the transition from savage to civilized existence in humans was analogous to changes in animal instincts – in both cases, new experience is the key. Thus, moral sense, for Darwin, is part of our hereditary constitution that has arisen through experience gained over countless generations. It is not strictly uniform across all individual humans at any given time, partly because of differential capability for receiving impressions and partly because of differential exposure to education and socially preferable upbringing.

\textsuperscript{140} The two terms are synonymous for Darwin even in his mature theory.
Yet Darwin accomplished more than benign reconciliation with the introduction of natural historical perspective into the question of moral foundations. He also significantly *revised* the two existing philosophical accounts, specifically on the issue of the moral good. The “instinctive” moral sense, as the Notebook excerpt above indicates, is “what has produced the greatest good /or rather what was necessary for good at all,” but it also “alone explains” the tendency to revenge. Hence the reconciled moral sense and Utilitarianism should now agree on a common sense of justice. Traditionally, however, Hume’s philosophy and J.S. Mill’s Utilitarianism disagreed on this point. To take Darwin at his exact words suggests that the relevant concept of justice in the natural historical perspective is retributive, i.e., grounded on revenge. This concept of justice involves the attempt is to avenge oneself or another through retaliation in kind or degree, yet often significantly out of proportion to the original offense.

Revenge is the least abstract of the various forms of justice. It is instead linked to the felt sense of satisfaction that the offending person is suffering for having caused the original offense. It is not explicitly or directly concerned with experientially-distant concepts of prevention or fairness. In this respect it is the most primitive, and therefore apparently the *oldest* form of justice, as revealed by historical perspective. Darwin suggested that our present-day inclination toward revenge – however much it might be

---

141 Justice is notoriously difficult for Hume’s moral philosophy. At the very least justice is, on his view, sometimes opposed to benevolence. On the one hand, justice must ultimately be an artificial virtue, as our natural sympathy for those to whom we are more closely related may override our judgment of the fairness of a course of action. On the other hand, it is difficult for Hume to identify precisely how the duty to justice is inculcated in individual members of a society across generations. For Mill, principles of justice arise from prior practices of distributing goods within a society combined with some conception of individual rights. (see *Utilitarianism*, Ch. 5). Although he arrives at the conclusion in a significantly different way, Mill nevertheless maintains with Hume that justice and benevolence are not equivalent virtues. Individuals develop a sense of pleasure independent of the sense of utility around the satisfaction of these principles through repeated association.
restrained by institutional structures and the education of individual members of society – is the evolutionary result of our advantageous (and thus heritable) inclination to preserve ourselves and those with whom we are most closely related. Justice as fairness or prevention of harm arose through culture and in social institutions, but never apart from this more fundamental, biologically rooted sense of justice.

Yet qualitative Utilitarianism raises a problem for Darwin’s approach to justice: What happens when decisions must be for the good of society on the basis of extremely remote future consequences? Are the decisions recommended on the basis of considering the good for the society in the very long run consistent with the advanced state of morality that will have emerged by the time that our intellect has developed sufficiently to even recognize and deliberate from the perspective of natural history? I will return to this question later in the paper.

According to Darwin’s moral theory, the category of morally good behaviors includes all behaviors directed away from the agent and benefitting some other living being (usually, though not exclusively, some other person). The most important subcategory of these behaviors (from his perspective) comprises those behaviors performed immediately (i.e., without deliberation), which are performed by humans as well as by animals with less advanced reasoning capacity. Darwin’s conclusion that moral motivation is separable from the exercise of reason bears superficial resemblance to Hume’s famous conclusion that reason alone (apart from a passion) cannot motivate. But key differences distinguish Hume’s conclusion from Darwin’s.

Darwin established his account on the assumption that the range of psychological complexity of moral behaviors (i.e., from immediate and nondeliberate action to very
refined behaviors) tracks the incremental historical development of morality.

Accordingly, the most basic form of moral behavior – the impulsive actions – is the oldest from Darwin’s natural historical perspective. The emergence of such immediate, other-directed forms of behavior traces part of the lineage of morality from analogous behaviors in lower animals to genuinely moral behaviors in humans. This kind of reasoning is foreign to Hume’s perspective, which was not informed by natural historical considerations.

The claim that humans can execute genuinely moral action apart from deliberation allegedly warrants the confident assertion that the same actions – though not yet called “moral” – were performed by lower animals and preserved by natural selection. Consider the implications for Darwin’s view of the other-directedness of actions. Though necessary, this is not a sufficient condition for an action to be considered “moral.” If the execution alone of such actions were sufficient for morality, then even lower animals would be moral. Darwin rejected this possibility:

A moral being is one who is capable of comparing his past and future actions or motives, and of approving or disapproving of them. We have no reason to suppose that any of the lower animals have this capacity; therefore, when a Newfoundland dog drags a child out of the water, or a monkey faces danger to rescue its comrade, or takes charge of an orphan monkey, we do not call its conduct moral. But in the case of man, who alone can with certainty be ranked as a moral being, actions of a certain class are called moral, whether performed deliberately, after a struggle with opposing motives, or impulsively through instinct, or from the effects of slowly gained habit.142

---

142 Darwin, *Descent of Man* (2nd ed.), 691.
Darwin thus explained human morality not in terms of heritable tendency to perform other-directed actions. Rather, he clarified that it is the capacity for reflective judgment that distinguishes genuinely moral behavior from analogous forms of behavior in other animals, even though in the moment of action the morally good behavior may not be deliberative. The human ability to experience the uniquely moral feelings of regret and shame (for example) depends upon such reflective judgment.

The concept of social instinct to which Darwin appeals to explain human sociability reflects his endeavor to reconcile the moral sense and Utilitarian philosophies. The moral sense aspect of this concept is manifest in Darwin’s claim that sociability involves sympathetic participation in the passions of others. Natural selection can explain the gradations of sympathy that a-historical moral psychologies (e.g., Adam Smith’s account of sympathy) left unexplained: “[W]ith all animals,” Darwin wrote, “sympathy is directed solely towards those members of the same community, and therefore towards known, and more or less beloved members, but not to all members of the same species.” Yet Darwin could only explain the preservation and augmentation of social instincts in exclusively naturalistic terms by positing that communities can be the unit of natural selection:

In however manner this feeling may have originated, as it is one of high importance to all those animals which aid and defend one another, it will have been increased through natural selection; for those communities, which included the greatest number of the most sympathetic members, would flourish best, and rear the greatest number of offspring.

---

143 Ibid., 688.

144 Darwin, *Descent of Man*, 688.
The concept of community selection as articulated by Darwin thus has a “utilitarian” emphasis to it. What is selected is what is best for the community and not for specific individuals.

Moreover, the limited range of social instincts is reflected in the heritable constitution that humans share with all other social animals. On the one hand, social instincts promote sociability and cooperation:

As man is a social animal, it is almost certain that he would inherit a tendency to be faithful to his comrades, and obedient to the leader of his tribe; for these qualities are common to most social animals. He would consequently possess some capacity for self-command. He would from an inherited tendency be willing to defend, in concert with others, his fellow-men; and would be ready to aid them in any way, which did not too greatly interfere with his own welfare or with his own strong desires.\(^{145}\)

On the other hand, there is an important ambivalence to the direction of human behavior on the basis of the hereditary constitution of humans alone:

At the moment of action, man will no doubt be apt to follow the stronger impulse; and though this may occasionally prompt him to the noblest deeds, it will more commonly lead him to gratify his own desires at the expense of other men.\(^{146}\)

Reflections on past behavior, apart from the pull of impulse, as well as conscience, guide future behavior. Education and habituation give direction and consistency to behavior, thereby eliminating much of this ambivalence. In this respect, moral development beyond the basic constitution of all social animals is like artificial

\(^{145}\) Ibid., 689.

\(^{146}\) Ibid., 693.
selection applied to humans by humans. Thus human moral development does not emerge as a necessary consequence of nature understood as a normative moral ground. Rather, ethics emerges first as a consequence of habits preserved in nature through selection that are recapitulated through the biological process of development. Behaviors that have proven to be advantageous to the community are preserved as such, and only afterward, upon reflection and association with feelings of guilt and remorse, are considered “moral.” The advancement of morality continues through culture and social norms that prescribe behaviors that may be selected and preserved across generations.

Now that the nature of the moral good has arisen for consideration, I will proceed with the comparison of Darwin, Hume, and J. S. Mill specifically concerning the metaphysics of moral obligation and moral judgment.

4.3 Hume, J. S. Mill, and Darwin on Moral Metaphysics

Earlier in Chapter 3, I showed that, according to Hume, the obligation to pursue the moral good can be traced to the particular structure of human psychology. Humans are social beings particularly disposed to feel pain and pleasure (i.e., the moral sentiments) in response to certain behaviors and interpersonal relations. What is considered “good” corresponds to this distinctive character of human moral psychology. Broadly speaking, and with some qualification, Hume rejected the more traditional belief that moral good exists objectively, e.g., as a property of external objects or in the mind of God. Rather, moral good is subjective, arising from associative psychological processes in the observer and primarily concerned with the psychology (i.e., interior condition) of
Consequently, Hume could not and did not attempt directly to ground the sense of “ought” (i.e., obligation and, to a certain extent, motivation) on naturalistic principles through some version of a “must” implies “ought” argument. In fact he cautiously and consistently remained at a remove from metaphysics, and argued instead for the more conservative proposition that it “ought to be the case” that “must” implies “ought.”

Hume’s moral psychology was developmental insofar as he acknowledged that individuals must be habituated and educated in order to recognize more readily and consistently respond to the universal moral sentiments. Although human nature could thus be cultivated, it is nevertheless significantly constant across generations. “Rather than seeing the effect of moral education as that of producing moral sentiments, it makes sense to say that moral education proceeds by working to strengthen and direct natural sentiments for morality that agents already possess to some degree or another.”

Artificial virtues are cultivated anew in each individual from birth, through unique combinations of the natural sentiments; such virtues are not grounded in new, basic sentiments but rather in compounds of basic sentiments. In this respect, Hume’s moral psychology seems to foreshadow that of J. S. Mill.

Like Hume, J.S. Mill believed that the obligation to pursue moral good ultimately lies in human psychology. Darwin would have learned from Mackintosh that the elder James Mill generally agreed with Hume that humans learn and behave on the basis of

---

147 Here I gloss over the intricacies of the ontological status of moral good in Hume’s philosophy.


psychological associations. The range of feelings that Hume called “positive moral sentiments” are, in Mill’s terminology, the pleasures. This terminological shift is significant because it signals that Mill’s human psychology traced moral experience to a more basic level, prior to the demarcation of moral from non-moral experience. Mill maintained from his predecessors in the Utilitarian tradition that humans learn to associate (and thus to behave in certain ways) due to the associations inevitably made between certain actions and the satisfaction of desire for the agent. Accordingly, humans must act in pursuit of pleasure (i.e., it is impossible for humans to act contrary to pleasure as an end): “pleasure and the freedom from pain are the only things desirable as ends; and…all desirable things…are desirable either for pleasure inherent in themselves or as a means to the promotion of pleasure and the prevention of pain.”\footnote{J. S. Mill, \textit{Utilitarianism}, 448a.} Therefore, individuals ought to maximize the pleasures inevitably sought.

J. S. Mill’s introduction of the \textit{qualitative} distinction among pleasures, as observed in the previous subsection, marked a significant departure from Bentham and James Mill.\footnote{Although Bentham and James Mill did not acknowledge such a qualitative ranking of pleasures, encouraging people to pursue the higher pleasures nevertheless satisfies the demands of the quantitative moral calculus of earlier Utilitarianism. Higher pleasures are typically conducive to producing more pleasure for more people, as compared with the lower pleasures.} For the younger Mill, the higher pleasures corresponded to what are ordinarily called moral ends. One implication of this qualitative innovation is that moral good is thus bestowed once again with a kind of \textit{objectivity} which renders Mill’s moral philosophy decidedly non-hedonistic and also distinct from Hume’s. He explained “higher” pleasures in terms of kinds of life, which serve as the context or result of the pursuit of such pleasures. Thus pleasure was no longer reducible to inner, psychological
states. Cultivation of virtue involved, for the younger Mill, learning to discern which pleasures are higher in matters of dispute or confusion, and to value them accordingly. In this way, pursuit of higher pleasures is conducive to others’ welfare as well as to the common good.

Toward the end of second edition of the Descent chapter four, Darwin inserted footnotes 41 and 42, signaling his second direct engagement with J. S. Mill regarding another criticism John Morley had raised in the Pall Mall review of the first edition. At this point in the chapter, Darwin explicitly mentioned Mill’s “greatest happiness principle” for the first time and directed readers in footnote 41 to a specific page in Utilitarianism for a definition of this term. Yet Darwin did not disclose what he believed Mill meant by “happiness.” Darwin’s failure to engage closely with Mill’s distinctive concept of happiness leaves room for his misunderstanding of the difference between Mill’s philosophy and that of the Benthamites before him. On my reading, John Morley’s review of the first edition provided the strongest critique of Darwin’s project of reconciling moral sense and Utilitarian philosophies through natural selection theory. Morley’s criticism underscored Darwin’s misunderstanding. Qualitative Utilitarianism may have been far more difficult to reduce to the theory of natural selection because it was a compelling alternative to this theory, operating at the same level of explanation. Darwin claimed that the earlier associationist psychologies were inadequate to account for the fact that some actions (namely, instinctive actions) are not attended by a coincident experience or anticipation of pleasure or pain. Nevertheless, Darwin noted,

---

152 In his April 14, 1871, letter to Morley, Darwin confessed his embarrassment over initially and erroneously attributing to all Utilitarians (including J.S. Mill) the claim that the greatest happiness principle is the proper motive for moral behavior. See [DAR146: 410].
when individuals are somehow prevented from performing such instinctive actions, the hedonistic character of our human psychology is revealed – the agent experiences a kind of pain (i.e., regret) in response to her inability to act as she would have done otherwise, had she not been prevented. Motivation is thus linked to instinct and habit; in the frequent case in which several motives compete simultaneously, the stronger motive prevails. Therefore, those authors who “write as if there must be a distinct motive for every action, and that this must be associated with some pleasure or displeasure,”¹⁵³ are mistaken.

As far as impulsive moral behavior is concerned, Darwin’s considered view was that Mill’s principle can at most be used as a means of reflectively evaluating actions already undertaken at some point in the past. Based on my review of Mill’s philosophy of utility earlier in this subsection, it is clear that on this point the two men in fact agreed, even if Darwin was not initially aware of this agreement. In their correspondence, Darwin and Morley seemed to have reached some sort of mutual recognition that Mill does not consider the greatest happiness principle a direct or even proper moral motive. Yet Darwin only obscurely acknowledges this recognition in the footnotes in the second edition Descent.¹⁵⁴ Richards mistakenly characterizes this point of dispute as “a consideration that distinguished his theory from that of the Utilitarians,”¹⁵⁵ as if Mill and Darwin drastically parted ways on this contentious point.

¹⁵³ Darwin, Descent (2nd ed.), 697.

¹⁵⁴ Darwin admitted in the footnote that follows that Mill’s philosophy on the whole cannot be fairly summarized by the “greatest happiness principle.” Here he contrasts J.S. Mill with other, anonymously referenced Utilitarians who seem exclusively to ground their views on a form of this statement.

¹⁵⁵ Richards, Darwin and the Emergence, 223.
Indeed, in Footnote 42 Darwin admitted that Mill included a role for habit within his account of moral experience. Darwin also conceded that to an extent the standard and motive of moral behavior are not wholly or strictly distinct. On the motive-versus-evaluation issue, therefore, Darwin’s distance from Mill’s position is not entirely clear. It seems likely that Mill would have agreed with Darwin that an adequate account of human morality must begin from the observation that both deliberate and impulsive actions are ordinarily referred to as “moral.” It is in fact only in the placement of emphasis that the two men would disagree – Mill considered reason and deliberation exemplary (because his aim was to explain how it is that individuals can disagree about moral ends in the first place, and how individuals can arrive at agreement about the progressive reform of society for the good of all), whereas Darwin privileged the impulsive (because his aim was to explain the biological origins of basic, transcultural moral experience).

As I have shown in this subsection, Darwin claimed that the obligation to pursue the moral good could be traced to the particular character of the psychological aspect of our “hereditary constitution.” He analyzed and explained the sense of duty and the experience of regret in terms of the felt experience of the conflicts that often arise between social instincts and instincts for self-preservation. When an individual feels the pull of the “ought” in upon reflection or deliberation, this just means, for Darwin, that this person is conscious of a rule for conduct that he or she ordinarily has the option to obey or violate.\footnote{156} Over the vast course of natural history, what has been \textit{better for one’s own survival and the survival of one’s community} (i.e., behavior in accordance with

\footnote{156} Darwin, \textit{Descent of Man} (2nd ed.), 694.
social instincts) translates upon reflection into the sorts of actions we recommend to ourselves and others as what we *ought to do*, and, moreover, what we praise as what is *right to do*. What is called “conscience” arose as the sense of “ought” was affirmed, strengthened, and elaborated through the influence of the “general point of view” and other factors, such as fear of divine punishment.

In the paragraphs that followed the insertion of footnotes 41 and 42, Darwin was still concerned with clarifying the proper conception of the moral good. The principle candidate at the time of Darwin’s writing was Mill’s very own notion of the maximization of happiness, which Mill explicated as the pursuit of higher pleasures. Darwin’s succinct response to Mill in this portion of the *Descent* was the claim that the concept of the “greatest happiness” arose from the express wishes of the community and has indeed become a conscious secondary influence on behavior. Darwin’s own (and more basic) concept of the “general good” was allegedly more directly linked to natural selection.

Earlier in this chapter of the *Descent*, Darwin made a brief appeal to community selection as a means of accounting for the emergence of human morality. As I discussed earlier in this subsection, he remarked that an “advancement in the standard of morality and an increase in the number of well-endowed men will certainly give an immense advantage to one tribe over another.” When Darwin returned to this tendency of natural selection in the paragraphs now under consideration, he termed it the “general good” and

---

157 Ibid., 683.
158 Ibid., 693-94.
159 Darwin, *Descent of Man* (2nd ed.), 689.
defined it as “the rearing of the greatest number of individuals in full vigor and health, with all their faculties perfect, under the conditions to which they are subjected.”

Darwin said in this same paragraph that it is also appropriate to substitute “general good” with the synonymous term “welfare of the community.”

What exactly does Darwin mean by “welfare,” and how is it related to the concept “general good”? Most basically, Darwin attempted to use the “general good” concept to explain what occurs, on the level of the community, as natural selection proceeds over time. He claims that those communities with the strongest, most courageous, most clever, most upright, and most loyal individuals tend to survive throughout conflicts and competition with other communities. Therefore, less vigorous, less healthy, and, in a specific sense, less moral individuals are a liability for the community as a whole with respect to the ongoing struggle among all communities.

Over time, these less fit individuals tend to be weeded out naturally – perhaps in the midst of war, but also through disease and other factors. Darwin himself seemed to acknowledge the significant implications of natural selection operating unchecked within a community, for he remarked that “this definition [i.e., of the general good] would require some limitation on account of political ethics.” Darwin may have realized that if this tendency of natural selection were to become the conscious will of the community, controversial consequences would result in an effort to promote the biological perfection of humans. This is an opening to the more unpleasant consequences of Darwin’s ethical

160 Ibid., 698.

161 Darwin, Descent of Man (2nd ed.), 698.
theory that many readers want to deny, and which scholars such as Richard Weikart almost exclusively emphasize.

Richards has not critically examined Darwin’s definition of the general good. Instead, he has merely echoed Darwin in claiming that this distinction between the general good and the greatest happiness defeats “the reproach of laying the foundation of the most noble part of our nature in the base principle of selfishness.”162 Richards has taken this conclusion to mean that Darwin had defeated J.S. Mill’s version of Utilitarianism. Yet more questionable is the fact that Richards seems to have misinterpreted Darwin’s conception of community selection and the greatest good. He has claimed that “community welfare stood as its [i.e., moral sense’s] ultimate motive and object; hence, altruistic behavior might be pure in the biological depths as well as at the surface of intention.”163

Later in the paper I argue that Darwin’s concept of the “general good” cannot adequately unravel the riddle of the origins of truly exemplary forms of altruism, such as benevolence toward the so-called “useless” members of community, which Darwin associated with the most highly developed state of morality.164 A fully comprehensive and consistent theory of moral origins would have to admit that moral development is not transmitted and augmented across generations through strictly biological means.

162 Ibid.
Before moving on to review and critique Darwin’s positive moral theory, it is worthwhile to treat at some length the issues of moral obligation and the general character of human nature as discussed by Hume, J. S. Mill, and Darwin. The two issues are intimately related. The extent to which moral obligation is natural or conventional typically tracks the extent to which human nature is basically self- or other-directed. Hume adopted an intermediate position on both issues. He believed that some moral obligations arise directly from human nature. These obligations correspond to the category of natural virtues. Other obligations, such as justice, arise from the conventions created in the context of society. These are the artificial virtues. In his bifurcation of the virtues, Hume thus departed from two established accounts of the character of the virtues in the British tradition: he was neither entirely a moral naturalist (i.e., like Hutcheson and Shaftesbury, the proponents of earlier versions of moral sense philosophy), nor entirely a conventionalist (i.e., like Mandeville and Hobbes, who espoused versions of moral egoism). His corresponding interpretation of human nature was similarly intermediate. He believed that humans are naturally neither wholly benevolent nor wholly selfish. Rather, humans are moderately selfish but also naturally disposed to benevolence toward others with whom we are closely affiliated. In this way, Hume rejected Romantic primitivism as well as Hobbes’s view of humans as naturally depraved.

J. S. Mill likewise took an intermediate stance on the twin issues of obligation and the character of human nature, although he ultimately conceived of humans as more active in their moral development and decision-making than did Hume. On the one hand, all social norms are conventional, but this can be taken as Mill’s positive advance on
Hume, for the emphasis Mill lays on this point is that the existing social framework is capable of peaceful revision.\textsuperscript{165} Most social frameworks – at least those that are just – include rules for revising the framework itself. On the other hand, Mill recognized that humans have a natural capacity for (though no inherent obligation to) self-development as individuals. This capacity checks the strongly deterministic flavor of associationist psychology and is proactive, whereas Hume’s philosophy maintained a generally passive view of human receptivity to experience and institutional structure.\textsuperscript{166} Thus humans must pursue happiness, but the resulting obligation is non-specific. In Mill’s own words, “among the works of man, which human life is rightly employed in perfecting and beautifying, the first in importance is surely man himself.”\textsuperscript{167} Humans have the freedom in self-development to choose the kind of life each individual believes will be most fulfilling. Nevertheless, there remains a general sense in which the extent of individual human development can be evaluated according to the isomorphic scale of qualitative pleasures.

J. S. Mill perceived human nature as potentially progressive over the course of individual human lives; humans also have the capacity to develop responsibility in self-mastery across generations through the advancement of culture and social forms. Thus, both on the individual and species level, human nature is dynamic. Mill expressed this fundamental character of human beings in organic terms that manifest the influence of

\begin{itemize}
\item \textsuperscript{165} Thus Mill has a generally optimistic view of individuals who become critical of existing social structures and norms. This is an important point of contrast between his view and that of moral egotists.
\item \textsuperscript{167} J.S. Mill, \textit{On Liberty}, 134.
\end{itemize}
“Germano-Colridgean” philosophy on his own view. Mill insisted that human nature is “not a machine to be built after a model, and set to do exactly the work prescribed for it, but a tree which requires to grow and develop itself on all sides, according to the tendency of the inward forces which make it a living thing.” This progressive development of human nature across generations occurs through culture and social institutions, which are methods of transmission and transformation that may not be reducible to the biological-ecological process of natural selection.

How does Darwin’s characterization of human nature compare with that of J. S. Mill and Hume on this element of moral philosophy? The answer to this question lies in Darwin’s positive moral theory as he presented it in Chapter 4 of the second edition of the Descent of Man. I will now proceed to this examination in Chapter 5.

---

168 See, for example, J. S. Mill’s essay on Coleridge for a sense of the extent of the influence of German Romanticism on Mill’s own philosophy.

5.1 Prologue: Richards and Ruse on Darwin’s Approach to Morality

In Chapter 4, I challenged two dominant trends in scholarship by undertaking a detailed comparison of Charles Darwin’s moral psychology with the moral psychologies of David Hume and J. S. Mill. I engaged Darwin, Hume, and the younger Mill specifically on their explanations of the source of our moral knowledge, the foundation of moral obligation and motivation, and the role and extent of convention in the development of morality. I traced Darwin’s mature account of moral origins back to his early notebooks, in which he declared that his principle of natural selection was capable of reconciling the theory of moral sense with Utilitarian philosophy. In contrast with Robert Richards, I claimed that Darwin maintained even into the mature phase of his career that his theory of natural selection subsumes under it these earlier moral philosophies in the British tradition, and that therefore he cannot be considered simply as a naturalized moral sense ethical theorist. But I also claimed that in explaining the moral sense and the greatest happiness principle in terms of natural selection, Darwin also crucially revised these two earlier philosophies. Darwin’s account of morality thus represents a third alternative to Mackintosh’s moral sense ethical theory and J. S. Mill’s Utilitarian moral theory. Darwin’s view cannot be reduced to either of these prior views,
and not simply because Darwin draws upon sources external to the British tradition when developing his theory of natural selection.

As I discuss in detail in Chapter 6, Richards has more recently augmented his case against the Benthamite reading of Darwin by placing heavy emphasis on Humboldt’s influence on Darwin. Sloan, in drawing attention to the equally important French influence, has attempted to temper Richards’s claims on this point while also generally supporting Richards against the mechanistic, Benthamite reading. To this conversation I add that Darwin never integrated these several and often contradictory influences into a fully coherent complex. The implications of this tension at the heart of Darwin’s work emerge most starkly in his treatment of the “most intricate” problem for natural selection.

Ultimately, my detailed examination in Chapter 4 raised the issue of the character of human nature itself. Is it fundamentally benevolent or selfish? Is this question perhaps based on debatable presuppositions? I showed that Hume and Mill both took a basically intermediate position with respect to the general character of human nature, but that they disagreed on the extent of human nature’s plasticity. They also differed as to whether the development of our nature can be transmitted and augmented across generations. Hume’s moral philosophy may not preclude an account of moral progress, but it was not central to his philosophical project to provide such an account; Mill, however, was concerned with precisely this issue of moral progress. My research revealed that Mill’s understanding of human nature was not strictly hedonistic and dispassionately mechanical, but in fact overtly organic. On this point I stand opposed to the assumptions of the two dominant trends in scholarship on Darwin’s relationship to Utilitarianism.
Historian of science and philosopher Michael Ruse has offered an interpretation of Darwin’s approach to human morality that on significant issues contrasts with Richards’s interpretation of Darwin as a moral sense theorist indebted to the German Romantic tradition for his pantheistic conception of nature. According to Ruse, “when it came to human morals, Darwin sometimes thought a kind of group selection would be needed,” but “[a]t other times…Darwin thought that a kind of enlightened self-interest, what modern thinkers call ‘reciprocal altruism,’ could cause morality… Ability to behave in this way obviously could be caused by individual selection.” Nevertheless, “[w]hether we should conclude that Darwin was a group selectionist or an individual selectionist with respect to morality is probably unanswerable.” Against Richards’s claim that Darwin drew upon the German Romantics to overturn the Utilitarian tradition, Ruse alleges that “Darwin’s roots are in the British society within which he was born, educated, and lived.”

If it is the case, as I contend Darwin seems to have believed, that his natural historical account of morality subsumed and reconciled moral sense theory and Utilitarianism, then it should be the case that on the whole, according to his own principles, Darwin would make the same moral endorsements and even predict the same moral outcomes as did Hume and J. S. Mill. Indeed, according to J. S. Mill, “to do as one would be done by, and to love one’s neighbor as oneself, constitute the ideal perfection of


171 Ruse, Darwinian Revolution, 246.
utilitarian morality,” which is likewise the culmination of moral development according to Darwin.173 But does Darwin’s moral theory agree with Mill’s on all endorsements and predictions? The answer, I believe, is no. My explanation of why this is the case will take us beyond chapter 4 of the second edition of the *Descent of Man* in the following sections of this paper, but it starts with the positive moral theory Darwin outlined in chapter 4. As the genuine difference between Darwin and J. S. Mill becomes clear, Darwin’s understanding of the general character of human nature will also come into focus.

Whereas J. S. Mill emphasized the role of rational consideration of the qualitative differences among pleasures in making moral decisions, Darwin gave prominence to the role of sympathy (which likewise played an important role in Utilitarian moral psychology) when giving his account of moral origins. We recall from Section I that sympathy, for Darwin, is best explained in terms of natural selection:

> In however complex a manner this feeling may have originated, as it is one of high importance to all those animals which aid and defend one another, it will have been increased through natural selection; for those communities, which included the greatest number of the most sympathetic members, would flourish best, and rear the greatest number of offspring.174

Moreover, Darwin assigned a central role to sympathy in the historical development of morality beyond the foundation of social instinct. Sympathy gives us access to the internal states of others in our community and compels us to act accordingly, in much the same way that spoken language, which expresses our emotional

---

173 “…I have so lately endeavored to shew that the social instincts – the prime principle of man’s moral constitution – with the aid of active intellectual powers and the effects of habit, naturally lead to the golden rule, ‘As ye would that men should do to you, do ye to them likewise;' and this lies at the foundation of morality.” (Darwin, *Descent of Man* (2nd ed.), 702).

condition outwardly and often publicly, is a means of engaging with others on a moral level:

[T]he social instincts, which must have been acquired by man in a very rude state, and probably even by his early ape-like progenitors, still give the impulse to some of his best actions; but his actions are in a higher degree determined by the expressed wishes and judgment of his fellow-men, and unfortunately very often by his own strong selfish desires. But as love, sympathy and self-command become strengthened by habit, and as the power of reasoning becomes clearer so that man can value justly the judgments of his fellows, he will feel himself impelled, apart from any transitory pleasure or pain, to certain lines of conduct. He might then declare – not that any barbarian or uncultivated man could thus think – I am the supreme judge of my own conduct, and in the words of Kant, I will not in my own person violate the dignity of humanity.\(^{175}\)

On these points, Darwin demonstrated his reliance on Adam Smith’s contribution to the British tradition of moral philosophy.

In the following section of this chapter I examine the role of Smith’s psychology of sympathy in Darwin’s moral theory so that it may be considered alongside my earlier treatment of the role of Hume’s moral psychology in Darwin’s thought. Earlier, I argued that as a result of his own natural historical realism and the influence of Mackintosh, Darwin found in Hume a utilitarian justification for moral virtue. This concept of utility was a predecessor of similar concepts in the philosophies of Bentham and James Mill. In the following section, I begin the work of establishing the claim that the concept of sympathy that Darwin obtained from Adam Smith – and which was prominent in the earlier writings of Hume – has more in common with the moral sense realist tradition that began with Hutcheson and Shaftesbury. The importance of both Hume’s and Smith’s

\(^{175}\) Darwin, *Descent of Man* (2\(^{nd}\) ed.), 690.
moral psychologies thus forms the basis of the distinction – and the tension – between the elements of individual moral perfectionism, natural selectionist group perfectionism, and perfection of the social community relevant to Darwin’s moral theory. Perhaps, then, Darwin drew upon Hume and Adam Smith as part of his mature but ultimately incomplete effort to reconcile Utilitarianism and moral sense realism by subsuming both of these moral philosophies to what he recognized as the more general theory of natural selection.

5.2 Adam Smith’s Moral Psychology in Darwin’s Positive Moral Theory

Adam Smith [1723-1790] advanced perhaps the most mature account of the dynamics of sympathy at the foundation of the moral sense. Not only the stricter moral sense philosophers, but also J. S. Mill and his followers drew upon Smith’s rich analysis of sympathy to support their own claims. In the DAR 119 Reading Notebook entry for January 15th of 1843, Darwin recorded that he had “skimmed parts” of the *Theory of Moral Sentiments*. He was convinced by his survey of Smith’s moral psychology that it “ought to be studied for comparison of man and animals—derives all from sympathy—considers we approve or disapprove of ourselves by placing ourselves in person of another & thus viewing ourselves as we should view others—Children disprove this—Perhaps this origin of such instincts.”

Smith developed this analysis out of the moral psychology he inherited from his teacher Frances Hutcheson and his close friend David Hume. He clarified the role of the

\[176\] [DAR 119: 12a].

\[177\] [DAR 119: 12a].
spectator and the process of sympathy in moral evaluation. Equally so, he made a more significant departure from early moral sense theory than even Hume had made by emphasizing moral judgment over and above moral sense.

The psychological function of the imagination, on Smith’s view, plays a crucial part in our ability to take on the role of spectator. As spectator of another human’s suffering, for example, a person can imagine what it was like for her when she was in a similar situation. According to Smith, this process involves the spectator making a copy of the impression of sense stored in her memory from an occasion when she was in a similar situation. The imagination allows a spectator to “enter as it were into [the other person’s] body, and become in some measure the same person with the agent.”

Although Darwin drew significantly from Smith’s account of sympathy, he criticized Smith for failing to account for the gradations in sympathy that correspond to the strength of relationship an individual has to others whom he or she observes.

Like Hume, Smith endeavored to apply “experimental” methodology to moral philosophy. Smith believed moral rules are ultimately founded upon experience of what, in particular instances, our moral faculties, our natural sense of merit and propriety, approve, or disapprove of. We do not originally approve or condemn particular actions; because, upon examination, they appear to be agreeable or inconsistent with a certain general rule. The general rule, on the contrary, is formed, by finding from experience that all actions of a certain kind, or circumstanced in a certain manner, are approved or disapproved of.


Moral rules are thus discovered in much the same way that natural laws are discovered: through careful and cumulative observation. Just as it is possible to predict natural phenomena from natural laws, codified moral norms make prescriptions according to which members of an advanced society gradually conform their behavior. Moral rules are the result of a uniquely human reflective stance that we take toward our accumulated experience.

According to Smith, social cohesion can be explained by appeal to the shared perspective established through sympathy, rather than by appeal to the basic benevolence of human nature, and certainly not by appeal to the allegedly basic egotism of human nature. As human societies become progressively more “civilized,” individuals gradually begin to respond to motivations that are simultaneously self- and other-directed. Society continues to advance and human morality continues to develop through the critical role of the impartial spectator that forms in the faculty of imagination. The imagined spectator allows an individual reflectively to evaluate herself as well as her society, for the spectator contains both the collective perspective of society (which can criticize individual members of society) and the self-knowledge possessed by a society member but to which the society at large does not have access (and which can therefore be used to criticize the society). Mill considered the implications for human progress of this dimension of Smith’s philosophy in a more consistent, systematic manner. Darwin did not seem to appreciate these implications to the same extent that Mill did, which may help to explain the tension in Darwin’s own account of the possibility for biological human progress.
With these preliminary considerations of Adam Smith’s psychology of the spectator in mind, I proceed to careful review of the moral philosophy Darwin outlined in chapter 4 of the second edition of *Descent of Man*. I show that his theory represents his mature attempt to subsume both moral sense and Utilitarian moral philosophies under the broader explanatory principle of natural selection. But first I narrate the progression of Darwin’s work from the Notebook period through the publication of the *Descent of Man*.

5.3 Darwin: The Notebook Period to the *Descent of Man*

As early as 1842, Darwin had considered introducing his early transmutation theory to the public – indeed, he had even considered writing a “Treatise of Man” expanded from his Notebook entries on the question of human origins – but was dissuaded by several concerns prior to Joseph Dalton Hooker’s encouraging reaction. Darwin was not only concerned that his theory would be rejected because it departed too strongly from established natural science; he was also specifically distressed by the threat of atheism. Members of the religious community to which he still belonged at this time might accuse him of atheism. Perhaps more troubling was the worry that confirmed atheists would appropriate his theory as supporting rejection of more traditional theological beliefs. Hooker’s largely sympathetic reception of the 1844 draft of what would later be published as the *Origin of Species* encouraged Darwin to share his emerging theory more widely. Yet in early 1847, when Hooker read the 1844 draft while working as Darwin’s apprentice, Hooker admitted that he still had difficulty completely accepting Darwin’s transmutation thesis.
After a few years marked by personal tragedy and sustained work on his barnacle project, the next stage of Darwin’s investigation into the origin of the human species was initiated by his meeting T. H. Huxley in 1853. The next year, Darwin completed the barnacle project and resumed his concentration on species transformation. As is well-known, Huxley proved to be quite receptive toward the version of Darwin’s draft from the mid-1850s. In 1858, Darwin received A. R. Wallace’s letter putting forth Wallace’s own theory of species transformation. Wallace’s theory contained striking similarities to Darwin’s own theory but differed on points that would become decisive elements of Darwin’s mature response to the “more interesting” and “most intricate” problems for natural selection in the *Descent of Man*.

In 1862, Huxley published his own *Man’s Place in Nature*. Although Huxley disagreed with Darwin on certain points, he was a staunch supporter of the general thesis of “Darwinism,” which he himself named as such in 1860. This publication instigated a heated debate between Huxley and Richard Owen specifically over the question of human origins. Owen’s allegiance to a form of “teleological developmentalism” and his refusal to take a stand on the character of the process through which species transformation occurs confirmed his distance from Darwinism. A year later Darwin read Lyell’s new book, *Antiquity of Man*, which likewise addressed the origin of the

---

180 For a recent argument in favor of reading Richard Owen as a “teleological developmentalist,” rather than as the Paleyan design theorist that the Huxleyean tradition has portrayed him as, see Christopher E. Cosans, *Owen’s Ape and Darwin’s Bulldog: Beyond Darwinism and Creationism* (Indianapolis, IN: Indiana University Press, 2009).
human species. As compared with Owen, Lyell was far more receptive of Darwin’s basic position; nevertheless he took “excessive caution” toward it.  

A. R. Wallace departed from Darwin by claiming that natural selection could only have produced a brain size sufficient for the survival of primitive humans. He did not believe that this principle could explain the further development of the brain, in terms of size or functional capacity, beyond the basic needs of survival. According to Wallace, humans escaped natural selection in two ways in moving out of the primitive state.

…in two distinct ways has man escaped the influence of those laws which have produced unceasing change in the animal world. By his superior intellect he is enabled to provide himself with clothing and weapons, and by cultivating the soil to obtain a constant supply of congenial food. This renders it unnecessary for his body, like those of the lower animals, to be modified in accordance with changing conditions – to gain a warmer natural covering, to acquire more powerful teeth or claws, or to become adapted to obtain and digest new kinds of food, as circumstances may require. By his superior sympathetic and moral feelings, he becomes fitted for the social state; he ceases to plunder the weak and helpless of his tribe; he shares the game which he has caught with less active or less fortunate hunters, or exchanges it for weapons which even the sick or the deformed can fashion; he saves the sick and wounded from death; and thus the power which leads to the rigid destruction of all animals who cannot in every respect help themselves, is prevented from acting on him.  

---


Later in the 1860s, Darwin began to develop more extensively his concept of sexual selection. Wallace, who in many other respects paralleled Darwin’s view, altogether rejected this concept. In a work published nearly two decades after the first edition of the *Descent*, Wallace explained that he believed a “spiritual power” had intervened in human evolution:

The special faculties [mathematical, musical, and artistic faculties] we have been discussing clearly point to the existence in man of something which he has not derived from his animal progenitors – something which we may best refer to as being of a spiritual essence or nature, capable of progressive development under favorable conditions. On the hypothesis of this spiritual nature, superadded to the animal nature of man, we are able to understand much that is otherwise mysterious or unintelligible in regard to him, especially the enormous influence of ideas, principles, and beliefs over his whole life and actions…It will, no doubt, be urged that the admitted continuity of man’s progress from the brute does not admit of the introduction of new causes, and that we have no evidence of the sudden change of nature which such introduction would bring about. The fallacy as to new causes involving any breach of continuity, or any sudden or abrupt change, in the effects, has already been shown; but we will further point out that there are at least three stages in the development of the organic world when some new cause or power must necessarily have come into action.

Darwin’s approach to the “more interesting” and “most intricate” problems could be read as a response to Wallace’s rejection of this concept.

Nevertheless, sexual selection, along with concepts that can be traced to several entries from the M and N Notebooks, were the essential elements of Darwin’s *Descent of Man*.

---

183 Darwin first proffered the concept of sexual selection in the first edition of the *Origin of Species* as a secondary cause, but did not develop this concept until the *Descent of Man.*

Man, published in early 1871, a year prior to the publication of the sixth edition of the Origin of Species. In the first edition Descent, many of the M and N entries were weaved together as chapters 2 and 3, the “Comparison of the Mental Powers of Man and the Lower Animals,” in which Darwin presented his mature naturalistic account of the origin of human moral psychology, the origin of ethics, and his own positive moral theory. Here he responded to the limitation Wallace had put on natural selection in the case of humans. Darwin acknowledged that “civilized” humans challenge natural selection in a variety of ways (e.g., through military enlistment and the right of primogeniture) but selection is still at work on the whole. Darwin appealed to anthropological evidence for his claim that traces of barbarism remain in existing customs, beliefs, and language. This evidence suggested to Darwin that natural selection together with sexual selection had enabled humans to move out of a primitive state. These chapters also contained, as I discussed in Section II, negative claims about the versions of Utilitarian philosophy popular at that time, including the views promoted by Bain and Sidgwick.

Upon publication, the first edition was subject to critical reception, particularly over the account advanced in chapters 2 and 3. In Section III, I reviewed the politician John Morley’s criticism of Darwin in the Pall Mall Gazette for his brief but questionable treatment of J. S. Mill’s philosophy in these chapters. Darwin was particularly interested in Morley’s review and proceeded to engage with this defender of Mill through a series of letters. The challenge posed by Morley’s clarifications inspired Darwin to ask his son to read Mill’s Utilitarianism for himself and to explain it to his elder. The second edition of the Descent, prepared over several months, was published in 1874. What were chapters 2 and 3 in the first edition were now chapters 3 and 4. Some parts of these chapters were
revised as a result of the Morley-Darwin correspondence, but when compared against the letters, these revisions suggest Darwin never fully resolved his precise disagreement with J. S. Mill.

5.4 The Tension in Darwin’s Positive Moral Theory

The presupposition with which Darwin introduced his positive moral theory states that “man can generally and readily distinguish between the higher and lower moral rules.”\(^{185}\) The category of “higher moral rules” are “founded on the social instincts.”\(^{186}\) These behavioral patterns are typically universal because they emerged relatively early within human history. Such behaviors – which include filial affection, mutual defense, and aiding one’s peers – can be found, as Darwin notes earlier in the chapter in his discussion of sociability, even in tribal communities. Behavior in agreement with the higher moral rules is ultimately motivated by the general impulse of social instinct; the more proximal motivation is the language of praise and blame as well as the (actual or imagined) judgment of the spectator humans experience through sympathy. The object of “higher” moral behaviors is the “general good,” or the “welfare of the community,” which, as I have already explained, includes the biological perfection of the species, the survival advantage of the group relative to other communities, and the mutual advantage of agent and recipient of the action.

\(^{185}\) Darwin, *Descent of Man* (2\(^{nd}\) ed.), 699.

\(^{186}\) Ibid.
Distinct from the higher are the “lower moral rules.” Although this category is populated with rules that “relate chiefly to self,” it also includes rules that promote behaviors that are other-directed. Thus, Darwin admitted, “some of [these rules] when implying self-sacrifice hardly deserve to be called lower.” Clearly, this moral category is more complex than the “higher” category just described. What can be said of all the lower rules, whether self- or other-directed, is that they tend to vary across communities. As relatively recent developments within the history of evolution, they form the basis of the distinctions in Darwin’s day between civilized and uncivilized societies. Ultimately motivated by public opinion (which Darwin also refers to as the “will of the community”), the object of so-called lower moral behaviors is in many cases the “greatest happiness” of the individual or the group.

The specific characterization of the object of these lower behaviors (e.g., egalitarian or apparently irrational) depends upon the more proximate motivations causing the behavior. Darwin explained that the motivation provided by public opinion is (like social instinct) relatively general, apart from the specification and guidance provided by individual experience as well as sociocultural and familial upbringing. The unity of this otherwise diverse category is maintained by Darwin’s observation that “the judgment of the community will generally be guided by some rude experience of what is best in the long run for all members.”

187 Darwin, Descent of Man (2nd ed.), 699.
188 Ibid.
189 Darwin, Descent of Man (2nd ed.), 698.
“this judgment will not rarely err from ignorance and weak powers of reasoning.”

Thus we can distinguish two sub-categories of the lower moral rules: reason-based and non-reason-based.

In a civilized society such as Darwin’s, “virtuous tendencies” like “chastity, temperance, humanity to animals, etc.” are “first impressed on the mental organization through habit, instruction and example, continued during several generations in the same family” and perhaps, Darwin speculates, even become heritable. As the will of the community continues to develop through the exercise of reason, individuals are raised to act benevolently toward the “imbecile, maimed, and other useless members of society,”

as I have already identified. This “highest possible stage of moral culture,” based to some extent on the “greatest happiness principle,” is also signaled when individuals recognize that “we ought to control our thoughts, and ‘not even in inmost thought to think again the sins that made the past so pleasant to us’” because “whatever makes any bad action familiar to the mind, renders its performance by so much the easier.” These behaviors illustrate the reason-based sub-category I have suggested for the lower moral rules. Darwin inclines toward endorsing the heritability of such virtuous tendencies that mark civilized societies because they can be rationally understood in terms of the mutual advantaged conferred on all concerned.

190 Ibid.

191 Darwin himself does not explicitly define such a sub-categorization but, as I will indicate momentarily, the examples he provides to illustrate the various kinds of lower moral behaviors support this addition I have made to the organizational structure of Darwin’s moral theory.


194 Ibid., 699.
Non-reason-based behavioral guidelines dictate in matters of taste and etiquette in a given society, but they also explain to an extent the “strange customs” and superstitions found in other communities. These forms of behavior strained Darwin’s hypothesis that specific kinds of virtuous tendencies may become heritable, because such tendencies are “senseless.”\(^{195}\) Darwin acknowledged that even elements of his own culture did not seem to have a use or serve a purpose with respect to survival, and thus it would not be appropriate to attribute their preservation and transmission across generations to nature. As he clarified in the *Origin*, nature only selects on the community level in favor of those characteristics and habits that are mutually beneficial, not those that are merely customary. Darwin appealed to his controversial concept of sexual selection to account for this aspect of moral experience.

The tension in Darwin’s positive moral theory emerges most clearly several hundred pages forward in the *Descent*, in Chapter 20, in which Darwin applies his concept of sexual selection to the continued biological perfection of humans. In the section entitled, “The Manner and Action of Sexual Selection with Mankind,” Darwin hypothesized how a primeval community of uncivilized men may have reproduced so as to encourage the further advancement of the species. From his perspective, “it would be an inexplicable circumstance if the selection of the more attractive women by the more powerful men of each tribe, who would rear on an average a greater number of children, did not after the lapse of many generations somewhat modify the character of the tribe.”\(^{196}\)

---

\(^{195}\) Ibid., 700.

\(^{196}\) Darwin, *Descent of Man* (2\(^{nd}\) ed.), 1036-37.
What happens when humans procreate without foresight and deliberation? Darwin explained that “promiscuous intercourse” is among the “chief causes” that “prevent or check the action of sexual selection with savages.”¹⁹⁷ From his perspective, “[it] is obvious that as long as the pairing of man, or of any other animal, is left to mere chance, with no choice exerted by either sex, there can be no sexual selection; and no effect will be produced on the offspring by certain individuals having had an advantage over others in their courtship.”¹⁹⁸ The increasing biological perfection of humans across generations could be slowed, halted, or even reversed by unreflective reproduction.

In the final chapter of the Descent, Darwin hearkened backs to his earlier discussion of artificial selection in the Origin when he complains that “[m]an scans with scrupulous care the character and pedigree of his horses, cattle, and dogs before he matches them; but when he comes to his own marriage he rarely, or never takes such care.”¹⁹⁹ It seemed to Darwin that even with the increase in reasoning power and the advancement of civilization, humans routinely countered the progressive action of sexual selection by pairing with unsuitable mates. With the exercise of prudence and wisdom, however, humans could improve the physical, mental, and moral character of future generations. Darwin even ventured a prescription: “Both sexes ought to refrain from marriage if they are in any marked degree inferior in body or mind.”²⁰⁰

¹⁹⁷ Ibid., 1030.
¹⁹⁸ Ibid., 1031.
¹⁹⁹ Darwin, Descent of Man (2nd ed.), 1053.
²⁰⁰ Ibid.
Darwin said that “such hopes are Utopian and will never be even partially realized until the laws of inheritance are thoroughly known,” because only the most prudent would make choices about reproduction that encouraged biological progress. The majority would simply continue to follow blind passion or other irrational inclinations. The only effective way to ensure the improvement of the species would be to enact legislation, based upon “principles of breeding and inheritance,” that restricted marriage or mandated birth reduction. And in this lies Darwin’s hope for the continued biological perfection of humans. “Everyone does good service,” he continues, “who aids toward this end” of discovering these principles. Until that distant dream is realized, “if he is to advance still higher, it is to be feared that he must remain subject to severe struggle.”

The tension within Darwin’s positive moral theory thus centers on the question, *What should we do with imperfect people?* Darwin offered at least two answers to this question in the *Descent*. The two dominant interpretations of Darwin’s approach to the “most intricate” problem emphasize one or the other of these two “sides” to his response to the issue of imperfection by drawing attention to the corresponding portions of the text.

On the one hand, Darwin’s morally perfected human has “tender and widely diffuse sympathy” toward “men of all races, to the imbecile, maimed, and other useless members of society, and finally to the lower animals.” This is the “side” of Darwin’s

---

201 Ibid.
202 Ibid., 1054.
positive moral theory that scholars such as Adrian Desmond and John Moore emphasize.\textsuperscript{204} Such a man would out of benevolence try to improve the quality of life for individual existing humans through charitable acts and other forms of self-sacrificial, other-directed behaviors. Such a person would be an advocate of human rights. He would promote a form of government that would redistribute wealth and other resources so as to give all members of the community equal opportunity for self-development.

On the other hand, Darwin’s intellectually advanced individual restrains her moral sympathies so as to permit the “severe struggle” among existing humans to continue, thereby avoiding the possibility that “the more gifted men would not be more successful in the battle of life than the less gifted.”\textsuperscript{205} This is the opposing “side” of Darwin’s positive moral theory bears some resemblance to the interpretation that more controversial scholars such as Richard Weikart have emphasized.\textsuperscript{206} Although Darwin’s intellectually advanced individual would have recognized, like Darwin, that natural selection works toward the “general good,” this individual’s society would not yet try to augment this work of nature through population control. The society would recognize, prior to discovering the “laws of inheritance,”\textsuperscript{207} that legislated population control would be unjustified – or, as Darwin wrote, “would perhaps require some limitation on account of political ethics.”\textsuperscript{208} The highly rational civilization to which such an intelligent

\textsuperscript{204} See Section I for a review of the Desmond-Moore interpretation and references to the relevant literature.

\textsuperscript{205} Darwin, \textit{Descent of Man} (2\textsuperscript{nd} ed.), 1054.

\textsuperscript{206} See Section I for a review of the Weikart interpretation and references to the relevant literature. Although there is some resemblance to Weikart in the passages I quote in this paragraph, it seems to me that Weikart unjustifiably exaggerates Darwin’s meaning.

\textsuperscript{207} Darwin, \textit{Descent of Man} (2\textsuperscript{nd} ed.), 1053.

\textsuperscript{208} Ibid., 698.
individual belongs would search for the as-yet-undiscovered (in Darwin’s day) laws of inheritance which alone could “justify” laws to control marriage rights or set limits on family size based on (for example) physical or intellectual criteria.

The general moral impulse – social instinct – depends upon reason to guide and specify the patterns of behavior it motivates. Thus, morality and intellect have evolved together. The morally perfected human and intellectually advanced individual I have just described would be, on Darwin’s account, one and the same person. But how can the man of tender and widely-diffuse sympathy actually endure the sight (or knowledge) of the abject suffering of individual people? How could he in some future time enforce limits on marriage and reproduction? Wouldn’t restrained moral sympathy atrophy over generations? If it is indeed “better” in the long run to allow the severe struggle to continue, wouldn’t the community, upon recognizing this, start to praise those who ignore the poor and useless? Such lack of compassion would, according to the influence exerted by the will of the community, become a new kind of cruel virtue. Herein lies the fundamental tension in Darwin’s positive moral theory.

Many scholars have tended to read Darwin’s moral theory “one-sidedly,” in the sense of emphasizing one or the other of Darwin’s responses to the question of what to do with unfit people. Those on the side of Desmond and Moore emphasize the parts of Darwin that uphold humanitarianism as an ideal that progressive morality approaches through a naturalistic process augmented by culture. Those on the side of Weikart have emphasized the parts of the Descent where Darwin is concerned with biological perfectionism through sexual selection. Unless it is acknowledged that Darwin can be read in at least these two ways on morality, the tension in his view will remain obscure.
I claim that this tension in Darwin’s view concerns two irreconcilable aspects of moral perfection and biological perfection. On the one hand, Darwin denied, in contrast with Wallace and to some extent J. S. Mill, that one or more processes of cultural evolution had been more powerfully effective in the origin and increasing sophistication of human morality. Instead, in order to account for the fully natural origin of the inherent moral sense, Darwin endorsed the view that progressively more sophisticated forms of human morality become heritable and are therefore preserved, transmitted, and augmented 

biologically across generations. Moral development, according to this aspect, seems to be natural selectionist perfectionism driven by survival and thereafter deliberately furthered by a desire to enhance human traits. On the other hand, Darwin described a moral theory in chapter 4 of the Descent that upholds humanitarian virtue as the pinnacle of moral development. Moral development, according to this aspect, seems to be more consistent with traditional moral perfectionism. This tension creates incoherence because Darwin never reconciles these two aspects of biological and moral perfectionism in his theory. Therefore, perhaps the most fundamental criticism of Darwin’s moral theory is that it is ultimately incoherent, and not that it eliminates genuine morality or commits a version of the naturalistic fallacy.
Robert Richards opposes strong selectionism as well as Weikart’s emphasis on the natural selectionist perfectionism aspect of Darwin’s moral theory. Richards would have us believe that the natural selectionist perfectionist reading of Darwin is untenable because “the good nature fostered, in Darwin’s conception, contributed to that ever-growing, morally progressive state that the history of evolution exemplified, and, ultimately, produced.”\(^{209}\) He has attempted to ground his opposition to the Weikart interpretation by drawing attention to a concept of “nature” in Darwin that is itself “moral.” Richards claims that Darwin obtained this moral concept of nature from German Romanticism, particularly Humboldt. He supports his reading with references to passages from the *Origin* such as the one in which Darwin remarks that “as natural selection works solely by and for the good of each being, all corporeal and mental endowments will tend to progress toward perfection.”

In the previous subsection, however, I showed that there are other passages in the *Descent* in which a sense of “good” as biological perfection emerges in unresolved (and perhaps unresolvable) tension to this opposing sense of a natural moral good. If Darwin’s moral theory contains a multi-layered concept of “good,” then perhaps his concept of

\(^{209}\) Darwin, *Descent of Man* (2\(^{nd}\) ed.), 537.
“nature” likewise consists of multiple layers that Richards does not fully acknowledge. In order to investigate this possibility and simultaneously evaluate Richards’s Romantic interpretation of Darwin’s nature concept, I turn in this section to a careful exegesis of Darwin’s treatment of nature across his career, beginning in the Notebook period. Richards lays the foundation for his claim that Darwin endorsed a “biologized” version of James Mackintosh’s moral sense theory by tracing a line of descent from the late eighteenth- early nineteenth-century German Romantics and naturphilosophen to Darwin’s concept of nature.\footnote{See Richards, \textit{Romantic Conception of Life}.} He contrasts this concept with nature understood “as if it were a Manchester spinning loom – a clanking, dispassionate machine.”\footnote{Robert J. Richards, “Michael Ruse’s Design for Living,” \textit{Journal of the History of Biology} 37 (2004): 25.} I claim that Richards’s interpretation of Darwin’s nature is only partially correct. The assumption that Darwin conceived of nature in any univocal way throughout his writings simply does not hold up under close scrutiny of the texts. Rather, at least three distinct nature concepts reappear across Darwin’s writings: (1.) nature as vital ground, (2.) nature as selective agency, and (3.) nature as a non-teleological system of laws.\footnote{Phillip R. Sloan, “‘It Might Be Called Reverence,’” in \textit{Darwinism and Philosophy}, ed. Vittorio Hosle and Christian Illies (Notre Dame, IN: University of Notre Dame Press, 2005): 143-65.} The first concept – nature as vital ground – is more characteristic of Darwin’s early writings of the Notebook period. Passages of his writings that evoke this conception indeed support Richards’s interpretation, but they do not capture the whole of Darwin’s multi-layered nature concept.
Commenting upon Darwin’s reflections from the period of his *Beagle* voyage, Richards claims that Alexander von Humboldt “helped forge” Darwin’s “romantic conception of nature,” which, on his view, served as the foundation of his allegedly staunch anti-utilitarian moral theory. While on the *Beagle*, Richards continues, “Darwin experienced the South American environment, the interconnectedness of its various aspects, the sublimity of its scenes, and the moral behavior of its peoples – all filtered through a Humboldtian discourse on these very subjects.”

Darwin was undoubtedly inspired by the “vivid depictions” of South America included in the *Personal Narrative*, for as Darwin reflected, they “far exceeded in merit anything I have read on the subject.”

During this formative period, ‘nature’ for Darwin “typically referred to in capitalized form, takes on many of the attributes of Humboldt’s Goethean conception of nature… Like Humboldt, Darwin replaces the theological language of ‘creation’ or a creator-God with that of Nature.”

Humboldt’s romantic view of nature helped Darwin begin to conceive of nature as dynamic and creative, as the generative ground of even the most complex aspect of the most advanced form of life: human morality.

This “vital ground” aspect of Darwin’s concept of nature is perhaps most clearly expressed in the C Notebook. Darwin wrote that “[i]t is very remarkable as shown by Carus how intermediate plants are between animal life & ‘inorganic life.’ Animals only live on matter already organized. – This paper might be work consulting if any

---


214 Ibid., 522.

215 Ibid., 526.

metaphysical speculations are entered on upon life, namely Carus.”\(^{217}\) And in the previous entry he remarked that

> [a]fter reading “Carus on the Kingdoms of Nature, their life & affinity”, in Scientific Memoirs I can see that perfection may be talked of with respect to life generally. – When unity constantly develops multiplicity (his definition “constant manifestation of unity through multiplicity”) this unity, - this distinctness of laws from rest of universe (which Carus considers big animal) become more developed in higher animals than in vegetables.\(^{218}\)

This sense of perfectibility of life may be the foundation of the humanitarian aspect of the moral theory Darwin articulated decades later in the *Descent of Man*.

If this Romantic interpretation of Darwin’s nature were fully comprehensive, then the full corpus of Darwin’s writings should include Humboldtian language. On the contrary, a second layer of his concept of nature – nature as selector – took precedence once Darwin read the sixth edition of Thomas Malthus’s *Essay on Population* in 1838 and once again in 1847.\(^{219}\) During his first encounter with Malthus, Darwin was developing his first speculations on the transformation of species. “Entering his discourse from this date are a new set of images – metaphors of physical force, the pounding of hammers on wedges into an unyielding surface, the notion of a universalized struggle for

---


\(^{218}\) Darwin, Second Notebook [C], 93.

existence.”220 Darwin universalized Malthus’s principle to hold for all forms of life, such that species (regarded as populations) actively keep each other in check. Hence the well-known passage from Darwin’s D Notebook: “One may say that there is a force like a hundred thousand wedges trying to force <into> every kind of adaptive structure into the gaps <of> in the oeconomy of Nature, or rather forming gaps by thrusting out weaker ones.”221

Darwin continued to employ this language in the third chapter of the Origin, where he observed that “there is no exception to the rule that every organic being naturally increases at so high a rate, that if not destroyed, the earth would soon be covered by the progeny of a single pair.”222 Here Darwin suggests that the growth of all populations is limited by the competitive relationship among species, despite the appearance of peaceful coexistence. The competition among species constitutes the “war of nature” (E Notebook), from which “a struggle for existence inevitably follows.”223 Inadequate food or space sets limits on population only in extreme circumstances.

The influence from Malthus in Darwin’s writings introduced the concept of “providence” along with the “selector” layer of his nature concept. Richards seems to give far less emphasis to this aspect of Darwin’s nature concept because he wishes to preserve the fact that Darwin saw “a cosmos…a vast web pulsating with life,”224 rather

223 Darwin, Origin of Species, 380.
224 Richards, Romantic Conception of Life, 525.
than a world “clanking along in the manner of a nineteenth-century steam engine” described in the “hard, mechanistic language of chemistry and physics.” Richards seems rightfully concerned about the strong emphasis many readers of Malthus place on his “Iron Law of Population” and its role in the history of economics.

Nevertheless, Malthus himself clarified that he wrote his Essay in dialogue with those philosophers who believed that perfect societies could be created. He wished “to account for much of the poverty and misery observable among the lower classes of people in every nation, and for those reiterated failures in the efforts of the higher classes to relieve them.” In anti-utopian fashion, Malthus thus intended to critique the notion of perfectibility. This intention is reflected in the full title of his text from the second edition onward: An Essay on the Principle of Population: Or, a View of Its Past and Present Effects on Human Happiness, with an Inquiry into our Prospects Respecting the Future Removal or Mitigation of the Evils which it Occasions. This critique of perfectibility in Malthus may have informed Darwin’s approach to the issue of competition and extinction among species, and most certainly could have been an influence on his discussion of human progress through population control. If this aspect of Darwin’s nature concept is not acknowledged, there is a risk of misreading his response to the “most intricate” problem for natural selection.

225 Richards, Romantic Conception of Life, 539.
226 Ibid., 534.
228 Gordon has noted that as of the second edition Essay, Malthus added “‘moral restraint’ (abstinence, prudential postponement of marriage) as a controlling factor – thus, in effect, making procreation a matter of voluntary choice rather than a biological imperative,” p. 442
In chapter 4 of the *Origin*, Darwin explained that nature’s selection is constrained by a principle of mutual advantage:

> Natural selection will modify the structure of the young in relation to the parent, and of the parent in relation to the young. In social animals it will adapt the structure of each individual for the benefit of the community; if each in consequence profits by the selected change. What natural selection cannot do, is to modify the structure of one species, without giving it any advantage, for the good of another species; and though statements to this effect may be found in works of natural history, I cannot find one case which will bear investigation.²²⁹

Richards does not explicitly address this portion of Darwin’s text even as he strains to portray Darwin’s concept of “nature as selective agency” as the legacy of Romantic teleology. According to Richards, Darwin’s nature benevolently selects for the good of the individual; over the span of countless generations, nature’s benevolence has come to be mirrored in the benevolent behavior among humans living in community. Yet in this passage Darwin sounds quite different than in the *Beagle* journal entries of his youth. The shift in Darwin’s language became especially pronounced with the third edition of the *Origin*, published in 1861.²³⁰ In a passage added to chapter four of this edition, for example, Darwin presented a terse definition of natural selection: “This preservation of

---

²²⁹ Darwin, *Origin of Species*, 391. I fail to see the Romanticism Richards’ ascribes to Darwin’s nature in this crucial passage.

favorable variations and the rejection of injurious variations, I call Natural Selection. Variations neither useful nor injurious would not be affected by natural selection.”

Since he acknowledged that variations arising from one generation to the next could be advantageous, disadvantageous, or irrelevant to survival, Darwin added that “[w]e shall best understand the probable course of natural selection by taking the case of a country undergoing some physical change, for instance, of climate.” Thus nature selects so as to accumulate those variations that render a species more well-adapted to a given environmental niche. Environmental changes, however, might mean that species traits which were once favorable no longer confer advantage; thus, in subsequent generations, nature would select in favor of traits that better equip the species to flourish in the changed environment. Darwin found in Malthus a template for a naturalistic explanatory framework capable of reinterpreting these shorter-term, more deeply troubling, apparently “negative” trends of extinction in the natural world (the suffering and death of countless individuals) as longer-term “positive” trends toward adaptation in the natural world (the creative emergence of novel variations which Humboldt’s writings first allowed Darwin to think about in naturalistic terms). The process of selecting that Darwin described in the third edition is thus characterized, with respect to the long-term adaptation of a species, by a sense of Malthusian providence.

On the basis of this passage from the third edition version of chapter four in the *Origin*, it might be concluded that in this more mature phase of his writing, Darwin was moving away from the tendency in his youth to discuss nature in personified terms as a

---


232 Ibid.
vital ground. Beginning in the third edition of the *Origin*, Darwin did in fact remove some of the language of personification in response to critical reception of the earlier editions. He claimed that those who had given the first two editions of the *Origin* a “teleological selector” interpretation had misunderstood his theory:

> Several writers have misapprehended or objected to the term Natural Selection. …It has been said that I speak of natural selection as an active power or Deity; but who objects to an author speaking of the attraction of gravity as ruling the movement of the planets? Everyone knows what is meant and implied by such metaphysical expressions; and they are almost necessary for brevity. So again it is difficult to avoid personifying the word Nature; but I mean by Nature, only the aggregate action and product of many natural laws, and by laws the sequence of events as ascertained by us. With a little familiarity such superficial objections will be forgotten.\(^{233}\)

Thus most clearly in this addition to the third edition version of Chapter 4 in the *Origin*, Darwin’s view of nature as a non-teleological system of laws was beginning to emerge. He began to qualify his earlier characterization of nature’s selection through his increasingly prominent use of the language of law in his discussion of species change. His claim that the term “natural preservation” may have been more apt to describe the selection of adaptive variations over time suggests that nature is more appropriately conceived as a “Grand Sorter” rather than as a “Grand Selector.”

The legacy of the “a-teleological system of laws” aspect of Darwin’s nature concept may lie in the biological perfectionism that is prominent in his response to the “most intricate” problem for natural selection. The sixth edition of the *Origin* was, after all, published within months of the first edition of the *Descent of Man*. As I discussed

\(^{233}\) Charles Darwin, *Origin of Species* (3rd ed.), 64. This passage remains virtually unchanged in subsequent editions of the *Origin*.  

110
earlier, the language of law is prominent in Darwin’s discussion of the “most intricate” problem in the final chapter of the *Descent*. This chapter is also one source of textual support for the biological perfectionist interpretation of Darwin’s moral theory.

While this shift in language is significant, overinterpretation should be avoided. The move toward a law-like characterization of nature “does not mean a clear shift of Darwin’s framework away from his earlier conceptions to Newtonian or Comptean-positivist interpretations.”

Darwin maintained the multiple layers of his concept of nature with varying emphasis throughout his career. I suggest this complex nature concept helps explain the tension in Darwin’s moral theory between moral and biological perfectionism. If Darwin had worked toward a more univocal nature concept, he may have resolved this tension. Left unresolved, this tension persists as an incoherence in Darwin’s response to the “more interesting” and “most intricate” problems.

---

234 Sloan, “‘It Might Be Called Reverence,’” 154.
CHAPTER 7: CONCLUSION

In Chapter 2, I traced the reception of Darwin’s moral theory in an effort to account for the emergence of two sources of tension within the literature. One tension concerns Darwin’s response to the “more interesting” problem of the origin of ethics. On one side of the tension, Robert Richards, following Edward Manier, identifies Darwin as a moral sense theorist indebted to German Romanticism and Naturphilosophie for a teleological, moral concept of nature that enabled the overturning of Utilitarianism. On the other side of the tension, Robert M. Young and others have argued that Darwin’s theory is, rather, the result of the determining influence of his contemporary British context: Paley’s natural theology, British political economy, and Utilitarianism. The second tension concerns Darwin’s response to the “most intricate” problem of human progress. Adrian Desmond and James Moore emphasize the moral perfectionist aspect of Darwin’s moral theory to argue against the opposing thesis of Gasman and Richard Weikart, who emphasize the biological perfectionist aspect of the later chapters of the Descent and its relation to severe forms of Social Darwinism and eugenics.

In the chapters that followed, I argued for three theses. First, I argued that the two sources of tension in the literature reflect tension within Darwin’s own theory of natural selection. Second, I argued that the two sources of tension are connected, because the moral perfectionism described and endorsed by Darwin’s positive moral theory prohibits
the approval of the struggle and suffering required for biological perfection. Third, I argued that this tension in Darwin’s moral theory can be traced to the multiple, non-equivalent layers of Darwin’s nature concept. My argument has implications for Darwin’s relationship to the history of moral philosophy.

In Chapter 3.2, I showed the need to qualify the link between Darwin and Hume that Alessandro Pajewski has recently defended. Although Pajewski is correct insofar as the reading lists testify to Darwin’s consistent and thorough engagement with Hume’s philosophy, he fails to acknowledge that Hume left the status of the moral sense ambiguous. This ambiguity enabled his moral philosophy to be interpreted in multiple ways by his successors in the British moral traditions. Indeed, both Utilitarians and moral sense realists appealed to Hume in defense of their positions.

In Chapter 3.3, I qualified Richards’s characterization of Mackintosh’s influence on Darwin. Following Richards, I acknowledge that in his younger years Darwin drew heavily from Mackintosh to familiarize himself with the history of moral philosophy. Nevertheless, I suggest that Richards has not given much attention to Mackintosh’s ambivalence toward Hume. On the one hand, Mackintosh dismissed Hume’s epistemological concerns in a way that supported Darwin’s natural historical realism. On the other hand, I discussed passages in the Dissertation in which Mackintosh praised Hume’s arguments justifying artificial virtues in terms of their social utility. This element of utility in Mackintosh has implications for Darwin’s relationship to the Utilitarian tradition.

In Chapter 3.4, I drew attention to the strong influence of British and German Romanticism on John Stuart Mill’s innovative departure from the quantitative
Utilitarianism of Jeremy Bentham and James Mill. Richards has not addressed the distinctive features of the younger Mill’s philosophy. He has instead simply merged J. S. Mill together with other Utilitarians Darwin allegedly overturned. Notably, however, Darwin himself failed to acknowledge what was distinctive in the younger Mill’s philosophy. His misunderstanding was challenged by John Morley’s review of the first edition of the _Descent of Man_ – a criticism that Darwin took seriously enough to engage with it further through correspondence, but which Richards has only very briefly treated in his own interpretation.

The clarification of the moral psychologies of Hume, J. S. Mill, and Mackintosh in Chapter 3 formed the foundation of my extended comparison of these authors in Chapter 4 on basic elements of moral philosophy. In Chapter 4.2 and 4.3, I turned to Darwin’s Notebooks to argue that he did not in fact intend to overturn Utilitarianism as Richards has claimed. Rather, he persisted in his belief that he could reconcile moral sense realism with Utilitarianism by reducing both to natural selection. This more positive sense of Darwin’s relationship to Utilitarianism is supported by the significance he attributed to Mill’s endorsement of the logical form of his theory and research methodology, as well as Darwin’s own approval of Mill’s _On Liberty_. Yet, in the sections that followed I showed that the reduction of Utilitarianism to natural selection implied a change in the content of moral theory that Mill would have opposed. Mill attributed many aspects of moral development to individual psychology and cultural factors, whereas Darwin explained individual moral development in terms of a moral version of the principle that ontogeny recapitulates phylogeny. This reinterpretation of individual moral
development has implications for the possibility of the humanitarian social reform that J. S. Mill endorsed.

I continued my comparison of Darwin and J. S. Mill by turning to an examination of Darwin’s view of human nature in Chapter 5. I acknowledged in Section 5.1 that Darwin’s positive moral theory does in fact endorse a form of benevolence comparable to that which the younger Mill believed was generated by qualitative Utilitarianism. The difference between the two authors emerges, however, concerning the issue of human progress. In Section 5.3, I traced the history of Darwin’s concept of sexual selection from his response to A. R. Wallace’s claim that humans have escaped natural selection. The controversial concept of sexual selection is fundamental to the biological perfectionism implied in Darwin’s naturalized account of human progress. I explicated the tension in Darwin’s account of morality by contrasting the character of the individual who would endorse biological perfectionism with the character of the individual who would benevolently protect the weak members of society.

Finally, in Chapter 6, I argued that Darwin’s concept of nature is not univocal. Thus, Richards has offered only a partial account of the foundation of Darwin’s moral theory. A more comprehensive account would have to take seriously the aspect of Darwin’s view that treats nature as an “a-teleological system of laws.” Since Richards does not recognize the multiple, non-equivalent layers of Darwin’s nature concept, he does not acknowledge the unresolved tension in Darwin’s response to the “more interesting” and “most intricate” problems for natural selection.

My work suggests a way of thinking about how to situate Darwin within the western moral philosophy tradition that began with Aristotle and Aquinas. The principle
of natural selection in Darwin’s view seems to correspond to what Aquinas called
“natural law.” Like many other philosophers before and after him, Darwin clearly seems
to have been concerned with the relationship between “divine eternal law” and natural
law. Aquinas had argued that natural law reflects eternal law. I would suggest that
Darwin appealed to a kind of providence, especially in the Origin, as a way of
maintaining this traditional harmony between eternal and natural law. Notably, though,
the sections of the texts in which Darwin appealed to a kind of providence are grounded
by the “nature as selector” layer of his complex concept of nature. Darwin, in wrestling
with the meaning and significance of suffering for individuals, is closest to Malthus and
others in the British tradition in appealing to a kind of providence. Darwin didn’t make a
similar appeal to providence, nor did he seem to be as intent on maintaining harmony
between eternal law and natural law, when he discussed natural selection and nature in
terms of the other layers of his nature concept.

In other portions of his work that are more directly supported by the “nature as
vital ground” layer of his nature concept, he suggests that nature itself rather than a kind
of providence maintains the harmony between the good of the individual and the good of
the community. Aquinas had claimed that the revelation of “divine positive law” declared
this harmony between individual and common goods, and that through a combination of
reason and receptivity to this revelation we can live in ways that obey this harmony.
Darwin’s response to the “more interesting” problem of the natural origin of ethics and
moral psychology appealed to the development of moral sense through natural selection
as guaranteeing the harmonious link between individual and common goods. Darwin, in
wrestling with the question of how to give a fully naturalized account of the origin of

116
ethics, is closest to the moral sense realists of the British tradition in moving away from Aquinas’s emphasis on the role of reason and revelation in maintaining this harmony. Darwin, like moral sense realists, claimed that moral distinctions are fundamentally grounded in sensibility, rather than reason. Yet in order to articulate a view of nature robust enough to be the vital ground from which such a sense could evolve through natural selection, Darwin drew upon sources outside his native context, such as Humboldt. I have highlighted in this paper that Humboldt and other German Romantics were common sources for both Darwin and J. S. Mill. Thus, in articulating a view of nature that is at least in part a vital ground for the development of genuine but fully naturalistic ethics and moral psychology, the way in which Darwin moved away from moral rationalism actually brought him into greater proximity to the qualitative Utilitarianism of the younger Mill.

In his later career as well as in his discussion of the “most intricate” problem, Darwin placed more emphasis on nature as an “a-teleological system of laws.” Here his relationship to the western moral tradition is most complex and perhaps distressing in its ambiguity. Darwin presciently suggested in the final chapter of the Descent that there could be as-yet-undiscovered laws of heredity. If discovered, these laws would help to clarify the character and range of possibilities for the process of human progress. Darwin explicitly linked the discovery of such (then-) hypothetical laws to the ability to formulate just legislation concerning treatment of the poor and ill members of society. Darwin admitted throughout his writings that the suffering of individuals was unfortunate and yet trusted that the pain or even death of individuals could ultimately contribute to the common good, understood as what he called the general good. Thus, the discovery of
laws of heredity might even offer a conscientious person like Darwin a way of
reconciling legislation that permits or even promotes individual human suffering with
concern for the common good. In this area of Darwin’s work, these laws of heredity
hypothetically replaced the descriptive principle of natural selection as Darwin’s version
of Aquinas’s “natural law.” With no appeal in these portions of the texts to something
like providence as a kind of “divine eternal law,” Darwin suggested that natural law is the
highest form of law. Thus, here Darwin presents, very hesitantly and without devoting
much attention, a view significantly distinct from that which Aquinas had defended many
centuries prior. Darwin suggested the possibility that with more knowledge of nature, we
might no longer need to appeal to a kind of providence as the prototype of natural law.
Nor would we need to appeal to a kind providence to maintain the harmony between
individual and common good. Instead, natural law, including the laws of heredity, would
become the prototype for human positive law and would be the foundation of rational
consideration of the morality of individual human suffering. Darwin seems to have
approximated the distinctive logical rigor of the Benthamite qualitative Utilitarians on
this point, for it appears that reason, which discovers laws of nature, would then be
involved in making the complex moral distinctions necessary for revising legislation to
more effectively maximize the common good. In contrast, Aquinas believed that
appealing to revealed positive law was the only justifiable basis for revising or resisting
human positive law. Darwin was likewise furthest from his contemporary qualitative
utilitarian, J. S. Mill, on this point. Any concept of human progress within the younger
Mill’s philosophy would be grounded in his Romantic conception of happiness. This
conception of happiness seems to have been largely unconcerned with human progress in
the biological sense, because the good of the individual was always the starting point of Mill’s consideration of what was best for the community. The best community is that one which enables the flourishing of all individuals.

In trying to maintain a sense of both biological and moral perfectionism, Darwin tried to reconcile moral sense realism and qualitative Utilitarianism within the broader framework of natural history. Darwin never completed this ambitious project and has thus bequeathed to us fundamental questions concerning moral origins and human progress. Richards has admirably illuminated the aspects of Darwin’s moral theory that are grounded in the view of nature as a vital ground from which has developed a morality that providentially harmonizes individual and common good.

My work has shown that Darwin’s moral theory includes other aspects that equally definitive of Darwin’s response to the “more interesting” and “most intricate” problems for natural selection. Richards has yet to examine the complex relationship between Darwin, J. S. Mill, and Romanticism. Consequently, Richards has not appreciated the extent to which Darwin’s main project of reconciling moral sense realism with Utilitarianism, grounded in the view of nature as selector, was in tension with the very Romantic sources he drew upon in his youth. More problematic is Richards’s failure to acknowledge the a-teleological law layer of Darwin’s nature concept from which Darwin draws in his response to the “most intricate” problem of human progress. If Richards acknowledged these multiple layers of Darwin’s nature concept, I believe he would move away from his conclusion that Darwin is something like the British culmination of German teleological approaches to the relationship between nature and
freedom. Instead, I think he would agree that Darwin is more like a node where several approaches to this relationship intersect, but also collide.


Bowler, Peter J. “Do We Need a Non-Darwinian Industry?” *Notes and Records of the Royal Society*, 63, no. 4 (2009): 393-398.


______. “Darwin’s Notebooks on Transmutation of Species, Part II: Second Notebook [C] (February to July 1838).” Edited by Gavin de Beer. _Bulletin of the British Museum (Natural History): Historical Series_ 2, no. 3 (May 1960): 75-118.


