MAD ERRORS:
ASSOCIATED IDEAS, ENTHUSIASM, AND PERSONAL IDENTITY IN LOCKE

by

Kathryn Tabb

BA in History, Philosophy, and Social Studies of Science, University of Chicago, 2006

MPhil in History and Philosophy of Science, University of Cambridge, 2008

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Kathryn Tabb

It was defended on

April 10, 2015

and approved by

Margaret Atherton, Distinguished Professor of Philosophy, University of Wisconsin-Milwaukee

Kenneth F. Schaffner, Distinguished University Professor of History and Philosophy of Science, University of Pittsburgh

Kenneth Winkler, Professor of Philosophy, Yale University

Dissertation Advisor: Peter K. Machamer, Professor of History and Philosophy of Science, University of Pittsburgh

Dissertation Advisor: Mark Wilson, Professor of Philosophy, University of Pittsburgh

MAD ERRORS:
ASSOCIATED IDEAS, ENTHUSIASM, AND PERSONAL IDENTITY IN LOCKE

Kathryn Tabb, PhD

University of Pittsburgh, 2015

Associationism — in its most basic formulation, the view that all cognition begins with the compounding of simple sensations into chains of ideas — is frequently held to have been introduced by John Locke in 1700, expanded on by David Hartley and David Hume, and come into its own in the 19th century with psychologists like James Mill and Alexander Bain. The aim of this dissertation is to argue that Locke is not an associationist, and that he has been cast on the wrong side of a fundamental divide over the role of the understanding in the connection of ideas. I show that Locke coins the term "association of ideas" not to launch a new architectonic for psychology based on acquired habit, but to diagnose what he sees as the biggest obstacle to right understanding: madness. Hume's positive embrace of association has often been read back onto Locke, resulting in the easy conflation of the two thinkers under the banner of empiricism. In championing the powers of the active perception over the automaticity of association, however, Locke's psychology stands apart from later empiricist philosophies of mind.

Along with challenging Locke's traditional characterization as an associationist, this project explores the ramifications of Locke's concept of association for his broader commitments. Locke

believes that natural philosophy is possible due to our ability to perceive the truth or falsity of propositions, or, failing this, to make probabilistic judgments about their truth-value. The capacities that allow for these mental acts, reason and judgment (respectively), are gifts from God that allow us to flourish in our environment, despite our mediocre mental endowments. I argue that associated ideas show that these capacities sometimes fail us, compromising Locke's intellectualist picture. Something like false knowledge is possible in Locke's system, insofar as associated ideas generate propositions that are perceived to be true but which are in fact false. I call such propositions "mad errors," and describe their profound ramifications for Locke's ethics of belief and his theory of personal identity.

TABLE OF CONTENTS

PREFACE

I gratefully acknowledge:

John Forrester, Peter Machamer, Bob Richards, Ken Schaffner, Jon Tsou, Mark Wilson, and Bill Wimsatt (for my education),

The National Science Foundation Graduate Research Fellowship Program (Grant #1247842) and The Gates Cambridge Trust (for financial support),

Margaret Atherton, Hasok Chang, Mary Domski, Aaron Garrett, Şerife Tekin, Ken Winkler, Paul White, Charles Wolfe, and the members of STARS (for crucial guidance),

Julia Bursten, Steve Bush, Kathleen Cook, Meghan Dupree, Thomas Cunningham, Jonathan Friedlander, Carla Heelan, Jennifer Jhun, Joe McCaffrey, Michael Miller, Tom Pashby, Elizabeth O'Neill, Lauren Ross, Jeff Sykora, and most of all Margaret Tabb, Tamaara Tabb, William Tabb, and Daniel Williams (for care of my person).

1 INTRODUCTION: ASSOCIATIONS AND ASSOCIATIONISMS

This project begins with the conjecture that early modern philosophy was centrally concerned with mental pathology. The 17th century was an era in which accusations of irrationality and enthusiasm were leveled against religious and political dissenters, and in which nations were engulfed in bloody dissent over what it meant to govern and to worship reasonably. Intellectual error was not a happy prospect but was impossible to deny, and rationality could not be taken for granted. The proper conduct of the understanding was prerequisite for the ambitions that gripped the age, most notably towards civic peace and happiness in the future life. If the mind was healthy, it could perceive the world and its place in it, but the threat of pathology loomed large. Right management of the understanding was a Christian duty.

Despite this growing attention to mental health, and despite the madmen and women who must have been a daily sight for natural philosophers before widespread institutionalization, what we would today call psychosis was not often considered as a philosophical problem. Cartesian dualism extricated reason from the grips of the corruptible, impassioned body and posited that only the latter was touched by insanity. Physician-philosophers working within the Epicurean tradition, such as Thomas Willis, also envisioned a rational soul able to be compromised but ultimately left unaltered by aberrations in the animal and vital functions of the corporeal soul. Unlike the faults that the exercise of the rational faculties could cure, madness required physiological and behavioral interventions, not intellectual ones. Of primary concern to early

modern theorists of mind was how the rational soul overcame obstacles, and the most profound of all, madness, fell outside of this concern because it was caused by the body.

Against this background the originality of John Locke (1632–1704) comes into relief in ways that have not been attended to in the literature. While he has rightly been read as participating in the mental-hygienic project of providing a "postlapsarian cure" for the fallen state of the human mind (Corneanu 2012, p. 1), instead of relegating madness to the body Locke defines it in ideational terms, and insists it is of central concern for the philosopher of mind. Although like many of his contemporaries Locke views false beliefs, prejudices, and intellectual laziness as epistemico-moral failings that can be overcome via the divinely-gifted intellect, his *Essay Concerning Human Understanding* presents a madness as a contrast class, undercutting the optimism of his contemporaries: his psychopathology raises more problems than it solves. Madness, for Locke, is different in kind from other mental weaknesses, and he takes seriously its ability to corrupt every aspect of the human being, including the understanding and even personhood.

Locke's original turn is in defining madness as a certain type of relation between ideas that he calls an *association*. Rather than a failure in the use of the active powers of the understanding, association is a pathology of the powers of passive perception, sensation and reflection that renders certain ideas immune to the effects of reason and judgment. The association of ideas can lead to the intransigent insistence on fallacious ideas that Locke saw as the single biggest threat to civic peace and religious tolerance. When ideas associate and become irreformable, they fundamentally challenge the intellectual humility, reformability, and sociality that characterize Locke's vision of the liberal thinker. When Locke attributes religious enthusiasm or political dogmatism to the association of ideas, he is claiming that it is worse than

irrational; it is a permanently debilitating pathology. Association is foremost a medical problem, albeit one with profound epistemic implications. Locke repeatedly calls association a disease, and goes to pains to assure that his reader not interpret him metaphorically: "I shall be pardon'd for calling it by so harsh a name as Madness," he writes of it, "when it is considered, that opposition to Reason deserves that Name, and is really Madness" (II.xxxiii.4, p. 395).[1] The *Essay*, Locke writes, aims to provide a natural history of human understanding, and he follows Bacon's advice to "attach negatives to our affirmatives" through the consideration of related absences to phenomena of interest (Bacon 2000, p. 112). Madness becomes a way to bring mental health into view: insofar as Locke's general project assesses how rationality can bring human beings closer to God and closer to an ideal political state on earth, association is a central concern.

The initial task of this project is to offer a new account of Locke's theory of association, which was seized upon and transformed by a surge of self-identified associationist psychologists of the 18th and 19th centuries, eclipsed by these later accounts, and accordingly neglected and mischaracterized by historians of philosophy. Scholars have been tempted either to minimize the originality of Locke's view by comparing it with those of his contemporaries, or to exaggerate its influence by conflating it with those of later associationists. There is, indeed, little agreement about the extension of the term. In his entry for the *Dictionary of the History of Ideas* (1968), Robert Young traces the roots of associationism back to Plato, Aristotle, Juan Luis Vives and Thomas Hobbes, but describes it as maturing in the 18th century from "the systematic exploitation of Locke's paradigm for interpreting experience" under the care of Edmund Law,

[1] This citation from the *Essay* and all that follow refer to the book, chapter, and section of the quotation, followed by the page number in (Locke 1974).

John Gay, David Hartley, David Hume, and Erasmus Darwin, before coming to fruition into the 19th century. Howard Crosby Warren's monograph, *A History of the Association Psychology* (1921), begins with Aristotle, and treats René Descartes, Hobbes, Locke, George Berkeley, Hume, and Hartley before moving on to the main 19th-century figures like the Mills, Alexander Bain and G. H. Lewes. David Rapaport's *History of the Concept of the Association of Ideas* (1974) focuses on Francis Bacon, Descartes, Hobbes, Baruch Spinoza, Locke, Gottfried Leibniz, Berkeley, Hume, and Immanuel Kant.

An obvious difference in these genealogies is the inclusion of rationalists in the latter two. Young sees a central tenet of the school to be the associationist origins of all complex ideas, and believes a defining feature of associationist thought to be its empiricism: all complex ideas begin with simple sensations. Warren and Rapaport take a broader view, including those accounts that accept the possibility of innate ideas but emphasizing the importance of connections between ideas to rational thought. Under these readings, associationists are simply those thinkers who recognize the importance of relations between ideas to cognition. Rapaport calls Descartes an associationist because chains of ideas, connected via experience, play a central role in Descartes's theory of memory, in which animal spirits flow from the pineal gland through the same brain pores they have flown before, carving in the engrams that constitute memories.[2] The order in which ideas are experienced impacts the direction recall takes, as when, for example, "if I see two eyes along with a nose, I imagine straight away a forehead & a mouth, & all the other parts of a face, because I am not accustomed to seeing the former without the latter"

[2] See also Kallich (1945); Warren (2013); Gaukroger (2002) and Sutton (1998), all of whom use the term "associationist" to refer to Descartes's theory of ideas connected in memory.

(Descartes 1983, p. 179).[3] In this respect, Descartes was just as much of an associationist as Locke — both were dedicated to the central role of experience in the construction of (some) complex ideas.

Rapaport's account is felicitous insofar as it distinguishes two strains of associationism, each understood broadly enough to accommodate a wide range of early-modern thinkers. The first establishes the connection of ideas as a universal mechanism that acts as the basis of knowledge — Rapaport cites Girolamo Fracastoro, Tommaso Campanella, and Baruch Spinoza as foundational for this tradition (into which Berkeley and Hume would later fall). A separate tradition, which finds its roots in the Baconian idols and includes Descartes and Hobbes, contrasts associated ideas with actively combined ones and blames them for a wide class of intellectual faults. Rapaport claims that these two branches meet in Locke, arguing that while Locke limits association to the pathological connection of ideas, the mechanism he describes in fact underlies all cognition in his system: "Ultimately, this brought the whole psychic field under a unified law, although Locke was unaware that *associatio idearum* and reflection obey the same law" (1974, p. 81). While below I disagree — stridently! — with this reading of Locke, Rapaport's acknowledgement of a parallel associationist tradition in which associated ideas are not universal but rather pathological is valuable.

By grounding the associationist school in empiricist commitments, however, Young gives a more specific and useful delineation of the school, which better tracks those philosophers and psychologists who actually employed the label to refer to their work through the 18th and 19th centuries. According to Young, the membership in this lineage can be defined by allegiance to two principles:

---

[3] My translation.

> "(1) that complex mental phenomena are formed from simple elements derived ultimately from sensations and
>
> (2) that the mechanism by which these are formed depends on similarity and/or repeated juxtaposition of the simple elements in space and time." (Young 1968, p. 111)

Rationalists like Descartes would, of course, roundly reject the first principle. Indeed Descartes bemoans the fact that since "there is nothing whose true nature we perceive by the senses alone, it turns out that most people have nothing but confused perceptions throughout their entire lives" (1985, p. 220). And while Descartes would tolerate Young's second principle with respect to simple perception, he defines learning as a mental phenomenon rather than an embodied one, and unlike later associationists would reject any sort of psychological automatism about higher thought. As Young notes, to extend the term "associationism" to anyone who subscribes to a theory of the connection of ideas in order to displace the Aristotelian psychological picture is to ignore the significance of the second criterion: that propositional thought depends on associative principles, such that thinking and learning are the result of mechanistic processes.

On the other hand, if Young's second principle is insisted on, the associationist school loses another sympathizer — John Locke. Like Descartes, Hobbes, and Pierre Gassendi, Locke rejects the Scholastic notion of intentional species, instead developing a mechanistic model of sensory perception in which the sensoria of human beings deliver up to the understanding appearances, rather than reality. But for these philosophers, including Locke, relations between simple ideas are only part of perception — the passive part. The faculties of mind actively rearrange, compound, and compare ideas in order to generate propositional thought, in stark

contrast to the sort of passive association indicated in Young's second principle. To conflate their positions with later self-described associationists for whom mechanistic habituation is the central mechanism of thought, then, is to get these figures very wrong.

Theories of perception underwent a profound change in the early modern period. The Scholastics' faith that the sensorium delivered up objects and their relations holistically to the rational mind began to seem naive — Galileo had shown how wrong natural philosophers were about the surface of the moon and the nature of motion. Natural philosophers sympathetic with corpuscularianism faced a problem when it came to explaining knowledge: if the essences of objects in the external world could not be known at all, and their properties only indirectly through the effects they had on the sensorium, what justified belief? A common answer was that complex ideas relied on the order and frequency in which simple ideas were delivered up from real things to the understanding — on the connections between ideas that were perceived and stored in memory. Objects and relations that were fractured and fragmented through the process of perception had to be reassembled, and the constant conjunction of simple ideas formed the basis on which complex ideas and propositional thought about real entities were possible, and the profusion of sensations ordered.

The constant conjunction of ideas in experience is a powerful basis on which to understand the mind's capacity to learn, abstract, and navigate its environment. The reinforcement of connections between ideas through further experience can provide an inductive basis for assumptions about their correspondence with reality, and about the causal relations that stand between them. Connected ideas can be the foundation for habits, explaining, for example, why the understanding does not need to repeat, over and over, the inference that the perception of a circle variously shaded may in fact represent a sphere. Certain patterns of experience can

become mastered, and certain chains of ideas can become foundational for other, more complex, inferences. In the early modern period it was widely assumed that the ability to learn about his environment in this way was a gift to man from God, for which his faculties had been ideally designed.

Relying on the powers of mind to construct representations of the outside world makes knowledge of objects and relations quite vulnerable, however. The passive reception of chains of ideas allows for complex ideas to be formed, unscrutinized, in the understanding and habitual behaviors to be shaped automatically, rather than voluntarily. Every natural philosopher who uses chains of ideas as the basis for knowledge of substances owes an explanation for why these chains should be trusted, and for how the corpuscularian picture can be reconciled with an active and rational soul. Figures who have been located in the rationalist tradition, most notably Descartes, were every bit as concerned as empiricists like Robert Boyle with the mental-hygienic practices that assured that connections between ideas were sound. Language became one nexus for the scrutiny of connected ideas —Hobbes focused on how the use of abstract terms could mislead the understanding into conjoining ideas that in fact did not correspond to real relations. Physiology was another. Malebranche, for example, conceived of the difference between necessary and unnecessary connections between ideas as due to the motions of the animal spirits.

Like many of his contemporaries, Locke puts great stake in the ability of the understanding to suspend judgment while its active faculties do their work. Reason can assess a connection between ideas to make sure that they in fact agree with each other by searching for a middle term that can explain the relationship between them. Complex ideas of modes can be refined and perfected. In the case of substances, however, knowledge can only come from an understanding of real essence, which Locke thinks to be impossible for creatures of our mediocre

intellectual endowments. It is the corporeal natures of substances that would be demonstrative, and such knowledge is impossible. Instead objects and relations of substances are subject to probabilistic judgments based on the correspondence of remembered experiences, future experiments, and the testimony of others: the conjunction of ideas allows for belief, even if the nature of that conjunction does not allow for knowledge.

Locke's picture here is in its fundamentals Baconian, and not atypical of his time. Unlike his contemporaries, however, he acknowledges not only the risks of lazy induction, credulousness about the testimony of others, and failures of memory that come from faulty connections between ideas. In addition he acknowledges a further pathology that can result from a reliance on constant conjunction, which, drawing on his long-standing fascination with medicine, he calls *madness*. Madness is caused by the association of ideas, a unique kind of relation that goes beyond constant conjunction by binding ideas so tightly that the active powers of the understanding cannot distinguish them, and treat them as if they were one simple idea. While Locke believes bad habits can be broken and good habits consciously formed, associated ideas are not at all intentional, and have never been attended to by the mind, or scrutinized by the rational faculties. Locke not only acknowledges how the review of connections between ideas produces shades of irrationality, but goes further than his contemporaries by acknowledging real cases in which any such oversight is utterly absent. He believes association to be the worst possible form of mental pathology, "so great a force to set us awry in our Actions, as well Moral as Natural, Passions, Reasonings, and Notions themselves, that, perhaps, there is not any one thing that deserves more to be looked after" (II.xxxiii.9, p. 397).

Insofar as his basic "way of ideas" relies on the active powers of the understanding rather than automatic habituation to judge the streams of ideas delivered up to it by sensation,

Locke's picture is more Cartesian than it is associationist, *sensu* Young. Young writes that "Locke laid the foundation of one aspect of associationism in accounting for the origin of ideas by means of the juxtaposition in experience of simple ideas to form complex ones," but this analysis seems to strip credit from Descartes, Hobbes, Gassendi, and others whose theories about the connection of ideas preceded Locke's, and to mischaracterize him as subscribing to "a mechanistic, though not materialistic, epistemology" (Young 1968, p. 112). Locke's general theory of ideas fits into the connection-of-ideas picture, in which ideas are compounded, contrasted and abstracted through active, rather than mechanistic, powers. But it can only be anachronistically read as associationist in Young's sense.

It is plausible that this anachronistic reading of Locke as an associationist is one of the causes for the prevalence of what Margaret Atherton has called "the standard view," according to which Locke, qua empiricist, is "committed to a highly economical theory of the structure of the mind […] [as] endowed only with simple combinatorial networks" (1998, p. 49). This view is common among non-specialists.[4] While historians of philosophy for the most part acknowledge that Locke views association negatively, they often suggest that he shouldn't; that he embarrassingly fails to realize that association acts as a universal mechanism *within his own system*. Martin Kallich, for example, speaks of Locke's regrettable failure to extend his discussion from unnatural to "natural associations" (1945, p. 145), and Wayne Waxman writes that Locke ignores association's "importance in the explanation of human understanding itself" (2005, p. 371). Young sees fit to broaden Locke's narrow treatment on the grounds that "his

[4] Atherton characterizes this view as "a kind of 'common knowledge' picture of Locke and his relationship to empiricism," and attributes it to Noam Chomsky and Jerrold J. Katz. She suggests, however, that it is implicitly at work in some dominant philosophical approaches to Locke, specifically in the context of his theory of innate ideas (the subject-matter of her article) (p. 48).

discussion of association of ideas was consistent with and consequent upon the whole complex of ideas which led to the establishment of the empiricist tradition in science and philosophy" (1968, p. 112); and John Sutton, in a sympathetic gesture, substitutes "misassociation" on Locke's behalf, implicitly suggesting that Locke had a theory of positive association (1998, p. 198). Rapaport is more overtly critical, rebuking Locke for failing to see the relevance of his own theory: while Locke "brought the whole psychic field under a unified law, [he] was unaware that *associatio idearum* and reflection obey the same law" (1974, p. 81).

It is true that as an account of the mechanisms of cognition the *Essay* is a disappointment, and that Locke's discussion of associated ideas seems to offer only an abortive attempt at deliverance. Ultimately, however, Locke is more concerned with providing a standard of reasonableness for civic-minded Christians than with making a contribution to epistemology or philosophical psychology. In this manner Locke's agenda is quite different from David Hume's, often held up as the figurehead of the associationist school. Locke discusses the association of ideas only twice in his published works, in the fourth edition of the *Essay* and in the manuscript drafted at about the same time and published posthumously as *Of the Conduct of the Understanding* (1706). In contrast, association is ubiquitous in Hume; its principles are invoked to explain such central functions of the understanding as belief, the indirect passions, and sympathy. Hume writes that if anything can entitle him to "so glorious a name as that of an inventor, 'tis the use he makes of the principle of association of ideas" (1938, p. 31). Despite Locke's coinage of the term, this assessment seems warranted: for while Hume's theory of associated ideas is clearly drawn from Locke's, the "use he makes of" it is very different. The difference is due less to conflicting definitions of the term "association of ideas" than to

opposing views on the place that automatic connections between ideas should hold in a theory of human understanding.

In one of the few analyses comparing association in Locke and Hume, John P. Wright notes that "Hume appears to show that what Locke really uncovered was the sordid background of reason itself" (1987, p. 116). Wright means that Hume extends Locke's observation that in certain cases ideas become connected without any oversight from the active understanding, and generalizes it to encompass much of cognition. While for Locke association is the exception that proves the rule of a profoundly active and intentional mind, for Hume it is a broadly applicable explanatory mechanism, formalized in three principles: resemblance, contiguity, and cause and effect.[5] In his system these psychological laws explain healthy connections of ideas in the mind, taking on the roles belonging to the active faculties of the understanding in the Lockean picture.

In the chapters that follow we shall see that, *pace* Hume, Locke finds association far from a satisfying explanation for cognition: he describes it not only as a failure of the understanding but as a medical crisis. Broadly speaking, Locke is at pains to stave off mechanistic interpretations of his epistemology. He emphasizes throughout the *Essay* that human beings have received a divine bounty of mental powers, capable of shaping ideas to match the natural signs provided by God. Although Locke is often characterized as an arch-empiricist, he refers to the contents of the mind, not its structure, as a *tabula rasa*: we are born with faculties of mind, what he terms its "powers," that are essential to its function.[6]

---

[5] Of these only "resemblance" characterizes the relations that connect impressions.

[6] Locke also suggests that we are born with certain behavioral proclivities, most notably a tendency towards sin and self-indulgence that he sees as the result of Adam's fall (Spellman 1988) and a drive towards flourishing which is at the heart of natural law.

The next two chapters make the case that Lockean association should be characterized in terms of its medical valence, its location in passive perception, and its incurability. When taken together, these aspects show that Locke's theory is unique, and collapses neither into the broadly-held views of his day about the connection of ideas nor into the mechanistic Humean picture, in which the active powers of the understanding are eclipsed by the passions. I present Locke's account of associated ideas as mad ideas in Chapter Two, and in Chapter Three show how mad ideas can become incorporated into pathological propositions that I call "mad errors."

On the basis of this new understanding of Locke's theory, I argue that association is relevant for two interconnected areas of scholarly dispute about Locke's views: the ethics of belief, treated in Chapter Three, and personal identity, to which I devote Chapter Four. In both cases, Locke is attempting to solve a problem introduced by his sensationalism, his belief that all ideas come from either outer or inner sensation (reflection). In the first instance, if the certainty of knowledge is not assured through God-given innate ideas, how can it be responsibly established? In the second, if persons are constituted by chains of ideas rather than a mental substance or soul, how is the moral responsibility that undergirds divine judgment possible? The existence of associated ideas does not make it impossible for Locke to answer these questions, I argue, but it does complicate the answers insofar as it reveals human beings to be deeply vulnerable, both as knowers and as actors. Madmen are a testament to this, and one that Locke does not avert his gaze from. His unflinching attention to the fallibility of the human understanding reaps philosophical dividends that have had a lasting effect on psychopathology: I conclude in Chapter Five with a coda about the fate of Locke's own shadow school of associationism, what I call *medical* associationism, in the centuries following him.

2 ACTIVE CONNECTIONS AND PASSIVE ASSOCIATIONS

2.1 A NATURAL HISTORY OF THE HUMAN UNDERSTANDING

There has been a growing insistence by historians that contemporary philosophical divisions — into the epistemological, the metaphysical, and the ethical — are not easily applicable to early modern philosophical projects. In particular, it has been argued that in the 17th century moral and epistemological concerns were intertwined in regimens of self-improvement, which bore a closer relation to the ancient project of *paideia* than the impersonal formalizations of contemporary analytic philosophy. In the early modern period to act the angel was to *think* like an angel, to stretch the limits of the mediocre mental endowments that man became heir to upon exile from the garden, and to use God-given intellectual powers to read the moral law out of the book of nature.

Peter Harrison, for example, has written that "the early modern preoccupation with sin meant that in the realm of epistemology error was often equated with sin, and the human propensity to invest false claims with the character of truth was attributed to Adam's fall" (2007, p. 3). Rather than a secular revolution, Harrison describes the religiosity of the modern period leading to an explosion of primarily *anthropological*, rather than epistemological or metaphysical, philosophy. The driving questions of the period concerning human nature, and humanity's relationship with God, were practical ones. Susan James has suggested that this

emphasis on philosophy as a method of moral-epistemological self-improvement has escaped scholarly notice because it is so different from contemporary epistemology, whose fundamental questions concern the nature of knowledge, rather than the human powers and weaknesses that mediate its acquisition (1997, p. 161).

Sorana Corneanu has described a specific philosophical strain in the early modern period that she calls *medicina-cultura animi*, or the "medicine of the mind," in which "the purification, rectification and reordering of the human mind were […] inscribed among the general aims of experimental natural philosophy" (2012, p. 1). Advocates of this tradition, among whom Corneanu includes Bacon, Boyle, Robert Hooke, Joseph Glanvill, and others, take so seriously the limitations of the intellect that man's ability to access innate ideas, to form true propositions, and to accurately describe the world around him are thrown into question. Within this rubric, epistemological investigation is closely allied with medical practice, insofar as its aim is therapeutic; its aim is, in Corneanu's words, to generate "a package of guidelines to be used as instruments in a curative and cultivating regimen, assumed as the task of a Christian philosopher" (p. 118). Many of the regimens suggested by *cultura animi* philosophers are echoed in medical handbooks that promote "philosophical remedies" for melancholia, such as the focus on positive thoughts or the use of reason to break through irrational preoccupations and passions (Schmidt 2004).

But they were also the backbone of the natural-philosophical projects that preoccupied members of the Royal Society. Its apologist, Thomas Sprat, describes the goal of the new science as to "supply our thoughts with excellent *Medicines*, against their own *Extravagances,* and will

serve in some sort, for the same ends, which the *Moral* professes to accomplish" (1702, p. 342).[7] Locke himself frames his project of managing the understanding in medical terms:

> "The variety of distempers in men's minds is as great as those in their bodies; some are epidemic, few escape them, and everyone too, if he would look into himself, would find some defect of his particular genius. There is scarce anyone without some idiosyncrasy that he suffers by." (1996, p. 215)

This admixture of the medical and the moral is a signature of the *cultura animi* tradition, and reflects the broader synthesis of ancient and biblical authority that occupied philosophers in the 17th century.

Harrison, Corneanu and others have drawn the connection between the mental-hygienic projects of the modern period and their classical influences. Of especial interest is the Stoic tradition, in which philosophy was viewed as an "art of living," with rationality being the central prudential, as well as epistemic, aim (Sellars 2013). Alongside the revival of Stoic ethics by figures like Justus Lipsius and Guillaume Du Vair, we might see the *cultura animi* tradition as a Christian revival of Stoic ideas about the centrality of rationality in the pursuit of a good life. For early modern physicians of the soul, intellectual healing was also a form of religious transformation. To become more rational was to be better prepared to know the natural law, and was, in itself, an act of thanksgiving for our God-given mental faculties. Thus we see Locke write,

> "He that makes use of the Light and Faculties GOD has given him, and seeks sincerely to discover Truth, by those Helps and Abilities he has, may have this

[7] While Corneanu focuses on the British empiricist milieu, her analysis could almost certainly be extended to the continent: Descartes famously described the preservation of health as the principle aim of his philosophical studies (1983, p. 329).

> satisfaction in doing his Duty as a rational Creature, that though he should miss Truth, he will not miss the Reward of it. For he governs his Assent right, and places it as he should, who in any Case or Matter whatsoever, believes or disbelieves, according as Reason directs him. He that does otherwise, transgresses against his own Light, and misuses those Faculties, which were given him to no other end, but to search and follow the clearer Evidence, and greater Probability." (IV.xvii.24, p. 688)

Insofar as the management of mental maladies was a moral project, it was also an eschatological one, and Locke and his contemporaries were devoted to establishing an ethics of belief that could support the Christian account of divine judgment and punishment. Christianized formulations of ancient distinctions between the active and passive powers of mind emerged — like that of Aëtius, for example, who describes how, upon the "controlling part of [the] soul like paper well prepared for writing on," some conceptions form without conscious effort, "while others come about by our instruction and attention" (Hankinson 2011, p. 63). In the early modern period, the active faculties were those that we must manage if we are to find God through reason's light. Intellectual practices had a direct effect on divine deserts. What sort of ideas are generated actively, and what sort passively; what psychological processes should be considered exemplary active or passive functions, and how their employment relates to sin and deliverance; these were the central stakes of early modern debates about the mind.

Along with this general picture, the moderns inherited from ancient sources a means for determining whether knowledge was formulated rightly, through the proper application of the understanding: the suspension of assent. Corneanu describes how the Ciceronian notion of philosophy as "a kind of science and steadfast opinion of one yielding his assent upon good

grounds only" became a touchstone in the early modern period, in which diverse figures, most notably Bacon, allied the regulation of assent with the very possibility of justified and righteous belief (2012, p. 65). Bacon distinguishes between *anticipations* of the workings of nature, which the philosopher arrives at haphazardly, and *interpretations*, which result from engaging in the proper method. He emphasizes that his aim is in no way to champion the *permanent* suspension of belief, writing, "That which I meditate and propound is not *Acatalepsia* but *Eucatalepsia*; not denial of the capacity to understand, but provision for understanding truly; for I do not take away the authority from the senses, but supply them with helps; I do not slight the understanding, but govern it." The understanding needs must be regulated, to cure what Bacon frequently refers to as "diseases" or "distempers" of the mind that results from, in Joseph Glanvill's felicitous phrase, "precipitate judgments."

Like many of his Royal Society colleagues, Locke embraces Bacon's doctrine of suspension. While the first edition of his *Essay Concerning Human Understanding* (1689) provides a fundamentally hedonistic account in which the will is completely determined in its pursuit of the good, in the second Locke makes allowances for suspension. Beliefs about whether an action will cause pleasure or pain, under this picture, are not automatic but result from the work of the understanding to assess its ideas before acting. Action thus results from an intellectual engagement with the passions that drive men towards and away from different desiderata. Locke believes the improvement of our assessments about how best to relieve uneasiness are made possible through this capacity of suspension:

> "For the mind having in most cases, as is evident in Experience, a power to suspend the execution and satisfaction of its desires, and so all, one after another, is at liberty to consider the objects of them; examine them on all sides, and weigh

them with others. In this lies the liberty Man has; and from the not using of it right comes all that variety of mistakes, errors, and faults which we run into, in the conduct of our lives, and our endeavors after happiness." (II.xxi.47, p. 263)

The importance of suspension to Locke's psychology is clear, but its exact role — as constitutive of free action, or as merely exemplary of it — has been debated. Within the set of acts that commence with suspension and culminate in willed action, Locke writes, *lies the liberty Man has*. This has led some scholars to suggest that without the power of suspension, man is no better than an automaton (Yaffe 2000; LoLordo 2012). At the conclusion of this chapter I will argue against this view, and follow Julie Walsh (2015) in claiming that it is rather the *acts* of the understanding that *follow* the suspension of assent that are essential to human freedom: suspension alone is insufficient. The power to suspend can be seen as a passive one, which allows for the revision of ideas through the acts of the understanding that in fact constitute liberty.

Locke distinguishes the active and passive powers that constitute thinking in Book II Chapter 21 of the *Essay*, where he contrasts the passive "Power to receive *Ideas*, or Thoughts, from the operation of any external substance" with the active powers "to bring into view *Ideas* out of sight, at one's own choice, and to compare which of them one thinks fit" (II.xxi.72, p. 286).[8] The faculties which act on ideas are subsidiaries of the understanding, and include *discernment*, the recognition of sameness and difference between ideas; *composition*, the combination of ideas; *abstraction*, the generalization of an idea beyond its original context; and

[8] Elsewhere Locke notes, however, that "thinking" is popularly taken to only refer to active powers of mind: "Thinking, in the propriety of the *English* Tongue, signifies that sort of operation of the Mind about its *Ideas*, wherein the Mind is active; where it with some degree of voluntary attention, considers any thing" (II.ix.1, p. 143).

memory, which preserves ideas together which have occurred together in experience, storing them so that they can be recalled by *retention*. It is these faculties that, for Locke, explain the generation of complex ideas, and their taxonomy constitutes a "true *History of the first beginnings of Humane Knowledge"* (II.xii.15, p. 162). They are the activities of the understanding that the suspension of assent allows for, and which constitute liberty.

An immediate repercussion of taking seriously Locke's division between active and passive powers of mind is that any naive characterization of him as an arch-empiricist — a characterization that can pave the way to equating his psychology with that of later associationists — must be rejected. It is clear that Locke's infamous *tabula rasa* does not indicate that the mind itself is a void, only occupied through experience. Rather we are born with not only different propensities, but also different mental powers — indeed Locke will disappoint those liberal admirers who read him closely when he admits that we are born with varying intellectual endowments. God has given us the faculties to allow us to know Him, through our investigations of the natural world, our interpretations of scripture, and our intuitive knowledge of Him. We start with a clean slate and must fill it through our labor, but the tools by which to populate it with ideas are already in hand. In the following section I discuss how the active powers of the understanding begin this work through the generation of complex ideas.

2.2 LOCKE'S THEORY OF HEALTHY IDEAS

Locke's contemporaries, notably John Sergeant, Henry Lee, and Gottfried Leibniz, criticized him for replacing deductive reasoning from general maxims with the mechanical determination of the agreement or disagreement of ideas (Yolton 1956, pp. 73–79). But Locke

saw himself as preempting such accusations in the *Essay*. "What room then," he asks in a rhetorical flourish, "is there for the Exercise of any other Faculty, but outward Sense and inward Perception? What need is there of Reason? Very much; both for the enlargement of our Knowledge and regulating our Assent" (IV.xvii.2, p. 668). And indeed, the overarching thesis of this work is that Locke does not replace the mental faculties with the passive connection of ideas, as later associationists would do — his picture is more complicated.[9]

As noted above, Locke describes the understanding as having both passive powers (to be affected from without) and active powers (to cause change from within) that generate ideas. Both active and passive powers contribute to perception, belief, and knowledge. Locke puts great stock in the immediacy of intuitive knowledge, claiming that simple ideas as well as certain propositions are perceived without any intervention by the active faculties. In such cases knowledge is not volitional, but immediate and automatic. But the bulk of Book IV of the *Essay* is concerned with demonstrative and probabilistic knowledge, which only come as the result of an exertion of the mental faculties. I will refer to this work of the active powers as *intellectual labor.* The first stage of intellectual labor is the generation of complex ideas from simple ones, creating the stuff of demonstrative and probabilistic knowledge. The second is the assemblage of these complex ideas into propositions whose truth or falsity we can assess. The generation of complex ideas will be discussed in this chapter, and the creation of propositions in the next.

---

[9] Locke takes care to make clear that, while he employs the traditional term "faculty," he trusts his reader to know he means it in a new way: "These Powers of Mind, *viz*. Of *Perceiving*, and of *Preferring*, are usually call'd by another Name: And the ordinary way of Speaking is, That the *Understanding* and *Will* are two *Faculties* of the mind; a word proper enough, if it be used as all Words should be, so as not to breed any confusion in Mens Thoughts, by being supposed (as I suspect is has been) to stand for some real Beings in the Soul, that performed those Actions of Understanding and Volition" (II.xxi.5, p. 236).

Against the Scholastic tradition, Locke argues that the immediate objects of perception are not types determined by essences, but throngs of simple ideas that are discrete but which can be grouped on the grounds of co-occurrence in experience. Locke believes that simple ideas can be born in the mind from two sources — and only two sources, for there are no other "Windows by which light is let into this *dark Room*" (II.xi.17, p. 164). The first way, through *sensation*, gives humans simple ideas of the outside world. These ideas are generated in the sensorium in response to stimuli from the environment, and are immediately perceived by the mind. When the mind similarly perceives its own operations — that is, the manipulation of ideas by "powers intrinsical and proper to it self" (II.i.24, p. 118) — it generates a second class of ideas about these operations. This process is called *reflection,* and Locke refers to it as "internal Sense" (II.i.4, p. 105). Both classes of simple ideas, those generated from sensation and from reflection, are delivered, as it were, automatically to the understanding. Locke writes,

> "In this Part the *Understanding* is meerly *passive;* and whether or no, it will have these Beginnings, and as it were materials of Knowledge, is not in its own Power. For the Objects of our Senses, do, many of them, obtrude their particular *Ideas* upon our minds, whether we will or no: And the Operations of our minds, will not let us be without, at least some obscure Notions of them. No Man, can be wholly ignorant of what he does, when he thinks. These *simple Ideas*, when offered to the mind, *the Understanding can* no more refuse to have, nor alter, when they are imprinted, nor blot them out, and make new ones in it self, than a mirror can refuse, alter, or obliterate the Images or *Ideas*, which, the Objects set before it, do therein produce." (II.i.25, p. 118)

All ideas resulting from sensation and reflection are introduced to the active perception as simple, unalterable, true components of knowledge, for "God in his Wisdom" has "set them as Marks of Distinction in Things, whereby we may be able to discern one Thing from another" (II.xxxii.14, p. 388). While we cannot know underlying causes of simple ideas, we still have knowledge that they are real — what Locke calls *sensitive knowledge* (IV.iii.5, p. 539). He emphasizes that simple sensations may not resemble their causes — as a corpuscularian, committed to the primary-secondary quality distinction as formulated by his colleague Boyle, Locke recognizes that "the operation of insensible particles on our Senses" may produce a sensation (say, of the color violet) that relates to its source in a manner far more complicated than we could ever conceive. "It being no more impossible," he writes, "that God should annex such *Ideas* to such Motions, with which they have no similitude; than that he should annex the *Idea* of Pain to the motion of a piece of Steel dividing our Flesh, with which that *Idea* hath no resemblance" (II.viiii.13, pp. 136–137).

Unlike Descartes, then, Locke does not define simple ideas in terms of their amenability being intuitively understood — in a sense simple ideas remain opaque to us, since we will never know, for example, the underlying structure that gives rise to the perception of violet. Nonetheless for Locke simple ideas are *real* because they are constant natural signs we can rely on, part of God's plan, and the result of His goodness (Rickless, forthcoming). And unlike Hume, Locke need not argue for the causal relation between impressions of things outside us and our simple ideas. It is not constant conjunction but the origin of our passive powers of perception, carefully crafted for us by our creator, that assure our ideas correspond. The receipt of simple ideas is, therefore, a gift of knowledge from God to us. In its receipt of this sensitive knowledge of the external world the understanding is totally passive. Locke compares the mind, helpless

against the imprint of simple ideas, to a mirror; he also compares it to the organs of the body, which cannot avoid sensing.

Thus like later associationists Locke believes that simple ideas are passively perceived, insofar as they are caused from without and impressed upon the understanding. They cannot be divided into component parts, and result directly from the engagement of the sensorium with external bodies, or, in the case of reflection, with the mind's own ideas. But they are, nonetheless, *real*: "*simple* Ideas *are not fictions* of our Fancies, but the natural and regular productions of Things without us, really operating upon us" (IV.iv.4, p. 564). Even a secondary quality that exists purely in the sensorium, such as whiteness or bitterness, is consistently produced by the same type of external body and thus "has all the real conformity it can, or ought to have, with Things without us" (IV.iv.4, p. 563). Locke gives the examples of the ideas of coldness and hardness, which contribute to the complex idea of a piece of ice, and the scent and whiteness of a lily — qualities which, he says, "being each in itself uncompounded, contains in it nothing but *one uniform Appearance*, or Conception in the mind, and is not distinguishable into different ideas" (II.ii.i, p. 119). Because the understanding is only capable of combining and reordering ideas already present in the mind, it is unable to generate simple ideas, nor destroy those it already has.

In the fourth edition, Locke closes Book II of the *Essay* by offering a typology of the kinds of connections possible between simple ideas.[10] Some ideas, he writes, "have a natural

[10] Locke demurs from clarifying the ontological status of his concept of "idea." He emphasizes that he is using the term in the most general way to describe "whatever is meant by Phantasm, Notion, Species" — lumping together "phantasm," which plays a prominent part in the materialist philosophy of mind of Hobbes, and "species," a cynosure of the Scholastic vocabulary (I.i.8, p. 47). Despite the criticisms by Stillingfleet, Sergeant, Norris and others of his new usage Locke would, at times petulantly, refuse to clarify whether he saw ideas as substances

Correspondence and Connexion one with another: It is the Office and Excellency of our Reason to trace these, and hold them together in that Union and Correspondence which is founded in their peculiar Beings" (II.xxxiii.5, p. 395). These are the connections that bind our complex ideas, which, Locke emphasizes, we do not discover, but rather create to link the simple ideas delivered up to the understanding by experience. As examples of complex ideas Locke gives "*Beauty, Gratitude, a Man, an Army, the Universe*" (II.xii.1, p. 164). While there are as many potential complex ideas as there are combinations of simple ideas, all are substances ("a man"), modes (ideas that depend on substances for their manifestation, such as "gratitude"), or relations ("older," "husband"). We generate complex ideas to resemble, with more or less success, co-occurring properties of objects perceived either in the external world (for substances) or in the mind (for abstract ideas and mixed modes). The mind manipulates simple ideas into complexes through a variety of active powers described above — of recalling, discerning, compounding, abstracting, and comparing. Locke is explicit that simple ideas must be combined through the intellectual labor of the understanding; complex ideas are not delivered up through passive perception like simple ideas are:

> "But as the Mind is wholly passive in the reception of all its simple *Ideas*, so it exerts several acts of its own, whereby out of its simple *Ideas,* as the Materials and Foundations of the rest, the other are framed. The Acts of the Mind wherein it exerts its Power over its simple *Ideas* are chiefly these three, 1. Combining several simple *Ideas* into one compound one, and thus all Complex *Ideas* are made [….]" (II.xii, p. 163)

---

or modes, rejecting the received ontology of his critics (Yolton 1956). But in the *Essay* and in correspondence Locke makes it clear that ideas are not limited to representations of objects in the external world but encompass all ideas that pass through the understanding.

Locke continues to use the language of action throughout his discussion of complex ideas. He describes how ideas of particular substances are generated when "by Experience and Observation of Men's Senses" certain clusters of simple ideas are "taken notice of to exist together, and are therefore supposed to flow from the particular internal Constitution, or unknown Essence of that substance" (II.xxiii, p. 296). In order to generate a complex idea of a substance, then, the individual must first perceive the co-occurrence of simple ideas in what Locke likes to call "constant experience," and then add to that conglomerate the "confused" notion of substance. Thus Locke concludes that "*our specifick* Ideas *of Substances* are nothing else but *a Collection of a certain number of simple* Ideas, *considered as united in one thing*" (II.xxiii, p. 305).

Nonetheless, Locke is often taken as holding forth a mechanistic picture of the generation of complex ideas, analogous to the perception of simple ones. For example, Vere Chappell claims that in Locke's view we also receive via sensation and reflection "compounds consisting of two or more simple ideas joined together" (2011, p. 37). On this reading, complex ideas are not generated by the mind on account of the co-occurrence of simple ideas, but are in fact *constituted* by that co-occurrence, which immediately results in the presentation of a complex idea to the understanding from the sensorium. Complex ideas, under this reading, are also a product of passive perception, since "the reason these simple ideas are joined together in my mind is simply that the visible qualities to which they severally correspond are really joined together" (ibid.). Thus Chappell concludes that complex ideas of substances are, like simple ones, impositions on the mind, and no less passive, contrasting them with the voluntary creations of the mind such as fantastical ideas of substances, or modes.

Chappell might be basing this inference on Locke's frequent description of the *treatment* of certain complex ideas as simple ones by the understanding. While complex ideas are an active production of the understanding, their complexity need not always be the object of *attention*, an elusive faculty Locke often describes in tropes — "as it were, the Eye of the Soul" (II.x.7, p. 152). Once they have been joined together under a name, ideas are often considered together in a single complex as a result of habituation, rather than volition:

> "The Mind being, as I have declared, furnished with a great number of the simple *Ideas*, conveyed in by the *Senses*, as they are found in exterior things, or by *Reflection* on its own Operations, takes notice also, that a certain number of these simple *Ideas* go constantly together; which being presumed to belong to one thing, and Words being suited to common apprehensions, and made use of for quick dispatch, are called so united in one subject, by one name; which by inadvertency we are apt afterward to talk of and consider as one simple idea which indeed is a complication of many ideas together." (II.xxiii, p. 295)

Locke here describes two steps. First, the understanding combines simple ideas together that co-occur constantly in experience and unifies them under a single name. Second, the grouping together of ideas becomes so habitual that the cluster functions as a simple idea. In the case of substances, we accustom ourselves to the assumption that we are interacting with a discrete sort of object, when what we perceive is the frequent co-occurrence of a set of properties. Despite the apparent automaticity of complex ideas, *pace* Chappell Locke states explicitly,

> "Though the Mind be wholly passive, in respect to its simple *Ideas*: Yet, I think, we may say, it is not so, in respect to its complex *Ideas*: For those being

> Combinations of simple *Ideas*, put together, and united under one general Name; 'tis plain, that the Mind of Man uses some kind of Liberty, in forming those complex *Ideas*." (II.xxx.3, p. 373)

Famously, Locke rejects the possibility that we can know the real essences of objects in the external world. The sort of probabilistic belief about substances that comes from experience, often through the habituation of judgment, is acceptable to him as the next best basis for faith in the existence of things. In this he is not original; the experimental tenets of the Royal Society authorized the acceptance of facts upon "the testimony of nature" in lieu of demonstrative proof.

Locke notes, "the *Ideas we receive by sensation, are often* in grown people *alter'd by the Judgment*, without our taking notice of it" (II.ix.8, p. 145). Nonetheless, Locke's famous discussion of Molyneux's problem makes clear that the habitual and unattended-to compounding of simple ideas into complex ones is the result of, at least *originally*, an active engagement with observation and experience, that is, an act of intellectual labor. Molyneux posed the question to Locke of whether a blind man who recovers his sight must *learn* that the ideas of a shaded circle and of a round mass both correspond to the same three-dimensional object. Locke agrees with Molyneux that ideas that we may assume to be perceived automatically as complexes (such as the idea of a sphere) are instead the product of training and experience. Locke puts Molyneux's question to the reader "as an occasion for him to consider, how much he may be beholding to experience, improvement, and acquired notion, where he thinks, he has not the least use of, or help from them" (II.ix.8, p. 146).

Molyneux's problem also shows that the active powers that generate complex ideas are fallible. If the understanding is not exposed to the right sort of experiences, its ideas will not reflect actual clusters of properties in the world, that is, will not have real ideas. The work of the

understanding is to continually adapt its ideas to experience — thus while the blind man "at first sight" would not be able to differentiate between three-dimensional objects, he can learn that the feel and sight of a corner constantly co-occur, and ultimately can compound them. In his discussion of innate ideas, Locke emphasizes that even ideas that we cannot remember perceiving must at one point have entered our conscious awareness (I.iv.20, p. 98).

Tuveson has written that Locke visualizes the understanding "as a living power, reflecting upon the continuum of experiences, arriving at combinations which present a fairly reliable correspondence to external reality" (1960, p. 19). This adaptability in relation to experience is reiterated throughout Locke's discussions of development, that is, "by what steps we enlarge our *Ideas* from our first Infancy." He describes, for example, how the child first has a notion of "Mamma" that is quite particular to his own mother, and then corrects himself to separate out the individual complex idea of her person from the abstract idea of mothers in general (III.iii.7, p. 411). The adult's grasp on a complex idea of a substance can also be transformed through further observation and experiment — thus the nominal essence of a swan is different for the city-dweller and for the ornithologist, and the former's can improve with study (II.xxiii.14, p. 305). This is the heart of Locke's notion of the *tabula rasa*: the mind begins without ideas, but with the tools necessary to transform the simple ideas it receives into useful complexes that can ultimately ground propositional thought. Locke's picture is of the properly-conducted understanding engaging in a process of constant correction and improvement.

Given Locke's acknowledgement that complex ideas can be formulated without much attention and can become so habitual that they appear to the understanding as simple, it is clear why readers have interpreted him as thinking that connected ideas follow each other immediately and automatically in all cases. Locke is at pains, however, to make clear that most complex ideas

are still the result of active judgment. He saw a profound difference between ideas once perceived to go together that over time become connected and *associations*, which automatically follow each other for reasons not grounded in experience. The crucial difference is that the latter sort are not conjoined by an act of the understanding, and so cannot be fixed by one — they are outside the reach the active powers of mind. Crucially, the active powers of the understanding are what assure we have the *right* ideas, not just the most obvious or easiest ones. For Locke the active powers are our link to God's goodness, and give our understanding the capacity to find truth. While ideas can become connected outside of their auspices, these connections are no better than random linkages, leading to intellectual chaos rather than pious clarity.

2.3 LOCKE'S THEORY OF ASSOCIATED IDEAS

Along with describing intellectual error as the result of anticipating, rather than interpreting, connections between our ideas about the natural world, Bacon also used the language of anticipation to discuss madness. "*Anticipations*," he writes, "are quite strong enough to induce agreement, since even if men were mad in one common way together, they could agree among themselves well enough" (2000, p. 38). What protects men from madness is not, then, consensus or authority, but rather an application of the proper method. But while (as Corneanu has demonstrated) Bacon's emphasis on method as a prophylactic and cure for mental weakness was popular during Locke's day, this particular suggestion about madness was not widely adopted. Among Locke's contemporaries, madness was for the most part held to be a problem of soma, rather than of ideas. Hobbes defines it as an overabundance of passion, as does Descartes, though he also attributes it in places to the dark fumes of the melancholic humors and the

meanderings of the animal spirits. Among the Royal Society set Thomas Willis's neurophysiology was highly influential, and his *Pathologiae Cerebri et Nervosi Generis Specimen* (1667) gave complex accounts of mania, melancholia, and other sorts of psychopathology in terms of the motion of the animal spirits through the brain and nerves.

Locke's theory of the association of ideas is in the first instance notable because it provides an ideational theory of madness that situates mental illness against the *cultura animi* framework of self-improvement. I say "against" because for Locke madness does not have the moral-epistemological character of other sorts of intellectual failings that come about when we fail to improve our ideas during suspension. Our ideas of substances can be ill-crafted in all sorts of ways. They can be confused and obscure when we fail to distinguish our ideas from each other, or use words to refer to ideas that are not distinct (II.xxix). They can be inadequate when they only partially or incompletely represent the world as we experience it (II.xxxi). And when connections that we do not experience in reality are drawn between simple ideas the results are fantastical, such as the idea of a centaur or of a human body made out of glass, which are "made conformable to no Pattern existing, that we know" (II.xxx.5, p. 374).

If the notion of existence is also annexed to fantastical complex ideas, Locke calls them "wrong" insofar as they "disagree to those Patterns to which they are referred" (II.xxxii.26, p. 394). In the case of mixed modes, such as justice, the "truth" of the existence of the ideas is given, since all that is required for their realness is that it is possible for the simple ideas contained in them to co-occur in reflection. As long as the simple ideas that constitute them are consistent with each other, they cannot be "chimerical." In the case of substances, however, the situation is quite different. Since our complex ideas of substances aspire to capture the actual co-occurrence of properties in the world — that is, to be representational — they can fail to

correspond to their "archetype," or intensional object, and thus be "false" insofar as they contain an implicit proposition about existence.

All of these are the sorts of errors the mind can fall into "in its Apprehension and Knowledge of Things" (II.xxix.1 p. 363), and they can be corrected through the understanding's labor. Associated ideas are not amenable to these sorts of correctives. Rather than connected through an act of the understanding, associated ideas are pathologically conjoined by connections that, while in fact unnatural, become, as it were, naturalized — that is, which *appear* natural to the understanding. While the fantastical idea of a human body made out of glass can be resolved into its component parts and thus recognized to not correspond with reality, if the complex idea becomes an associated one, the idea of "body" will always be followed by the idea of "glass," even in the face of contradictory evidence about the nature of these substances — the two become "as if they were but one *Idea*" (II.xxxiii.7, p. 396).

Locke distinguishes associated ideas from healthy ideas only in the fourth edition of the *Essay Concerning Human Understanding* (1700), the last edition published during his lifetime. It was clearly important to Locke to introduce the new term, and he devotes an entire chapter to explaining association. His decision to introduce the new theory was made by April of 1695, when he wrote to his friend William Molyneux, "I think I shall make some other additions […] particularly concerning the Connexion of Ideas, which has not, that I know, been hitherto consider'd and has, I guess, a greater influence upon our minds, than is usually taken notice of" (1824, p. 534). Around that same time, presumably after writing to Molyneux, Locke made the decision to coin a term for this particular type of connection, crossing out the word "Connexion"

of ideas and replacing with "Association" in a draft for a new chapter to the *Essay*, what would become Chapter 33. [11]

Associated ideas differ from healthy ones not in terms of their content but in terms of their relationship to the powers of the understanding. As opposed to ideas compounded by the active powers in response to experience, Locke writes,

> "There is another Connexion of *Ideas* wholly owing to Chance or Custom; *Ideas* that in themselves are not all of kin, come to be so united in some Mens Minds, that 'tis very hard to separate them, they always keep in company, and the one no sooner at any time comes into the Understanding but its Associate appears with it; and if they are more than two which are thus united, the whole gang always inseparable shows themselves together." (XX.xxxiii.5, p. 395)

He gives several examples. In one, a young man learns to dance in a room with a trunk, and an association forms between the steps and the trunk. When he is called upon to dance in a different chamber, he finds himself incapable of it, since the trunk is absent. He simply cannot call up the chain of ideas without first perceiving their associate, the trunk (II.xxxiii.16, p. 399). In another, a man eats far too much honey as a small child and is so traumatized by getting sick from it that he is, for the rest of his life, too disgusted by the substance to eat it, since the very thought of him makes him sick (II.xxxiii.7, p. 396). The idea of "honey" has become associated

---

[11] Bodleian Library, MS. Locke e. 1, p. 32. The word would have carried an ominous meaning in Locke's time — according to the Oxford English Dictionary, among its usages in the 17th century is an accord between individuals to get up to mischief. In fact a 1682 usage given by the OED is from the London Gazette, in which an article described "That Seditious paper, The Association, lately found in the Early of Shaftesbury's closet." Shaftesbury was Locke's benefactor, so we can see the term might have been of special significance to him. As Simpson has noted, Milton uses the term to describe Lucifer's band of angels, as well as calling the subversive Eve Adam's "associate" (1999, p. 138).

with the idea of "sickness." Locke emphasizes that examples of association are manifold, and occur not only in madness but also in "the steady calm course of […] Life" (II.xxxiii.4, p. 395). Nonetheless, Locke takes care to emphasize that association, even when isolated to a single complex idea in the mind of an otherwise rational man, is the same phenomenon that is called madness:

> "I shall be pardon'd for calling [the association of ideas] by so harsh a name as *Madness*, when it is considered, that opposition to Reason deserves that Name, and is really Madness; and there is scarce a Man so free from it, but that if he should always on all occasions argue or do as in some cases he constantly does, would not be thought fitter for *Bedlam,* than Civil Conversation […]. That which will yet more apologize for this harsh Name, and ungrateful Imputation on the greatest part of Mankind is, that enquiring a little by the bye into the Nature of Madness, *B*.2 *C*.11 §13. I found it to spring from the very same Root, and to depend on the very same Cause as we are here speaking of." (II.xxxiii.4 p. 395)

I will discuss the origins of Locke's theory at length in the next section and in the following chapter, but here it is worth noting that Locke's discussion of madness in 2.11 provides a further, illuminating, example of association. Locke describes a madman who is certain his body is made out of glass. In Locke's mature vocabulary, we might say that an association has formed between the idea of his own body, and the idea of the adjective "glass." From then on, the madman cannot think of his body without thinking it is made out of glass. The two ideas follow each other so quickly that they are perceived as one simple idea. Locke notes that, nonetheless, mad people can make "right deductions" about what follows from this "Fancy" — believing themselves to be made of glass, they use "the caution necessary to preserve such brittle

Bodies" (II.xi.13 p. 161). Madness does not affect the capacity to make propositions, or reason about them; Locke differentiates madness from idiocy on precisely these grounds, writing,

> "In fine, the defect in *Naturals*[12] seems to proceed from want of quickness, activity, and motion, in the intellectual Faculties, whereby they are deprived of Reason: Whereas *mad Men*, on the other side, seem to suffer by the other extreme. For they do not appear to me to have lost the Faculty of Reasoning: but having joined together some *Ideas* very wrongly, they mistake them for Truths." (II.xi.13, p. 161)

In other words, madmen are not lacking in the capacity to correct *any* ideas, but have rather acquired ideas that are immune to these powers. As I lay out in Chapter 3, reasoning rightly about propositions that contain associated ideas leads to what I will call "mad errors," with disastrous results. But the point here is simply that because they have not been generated through the powers of judgment, whose activities are sanctioned by God to deliver us to true knowledge and justified belief, associated ideas form an unruly mob rather than an orderly procession, running counter to the Stoic ideal of self-control that was so influential in Locke's period (Garrett 2013). To attempt to disband associated ideas through reasoning with the afflicted individual is to "preach Ease to one on the Rack, and hope to allay, by rational Discourses, the Pain of his Joints tearing asunder" (II.xxxiii.13, p. 398). Locke would have been familiar with the popular portrayal of the Stoic as sufficiently protected by his wisdom, to the point that he can

[12] "Naturals" refers to those Locke also called idiots, approximating those whom we would today call intellectually disabled.

find happiness even on the rack.[13] Locke's use of the image, by contrast, underscores the dominance of the non-volitional, passive powers of perception over the intellectual ones:

> "While this Combination is settled and while it lasts, it is not in the power of Reason to help us, and relieve us from the Effects of it. *Ideas* in our Minds, when they are there, will operate according to their Natures and Circumstances; and here we see why Time cures certain Affections, which Reason, though in the right, and allow'd to do so, has not power over, nor is able to against them prevail with those who are apt to harken to it in other cases." (II.xxxiii.13, p. 398)[14]

Locke recognizes several ways in which ideas become associated. The first is through a perversion of habit, which, as described above, Locke generally views as a human good. Nonetheless habit can become an "empire," calcifying unattended-to chains of ideas into hardened complexes that cannot be modified when contradicted by further experience. The riskiest sort of habitual behavior is rumination, in which the mind reviews and reviews again the same chain of ideas, leading to their association. In the *Essay* Locke describes this process in physiological terms, writing that "Custom settles habits of Thinking in the Understanding" due to repeated "Trains of Motion in the Animal Spirits, which once set a going continue on in the same steps they have been used to, which by often treading are worn into a smooth path, and the Motion in it becomes easy and as it were Natural" (II.xxxiii.6, p. 396). As is typical, he disclaims his use of physiological language and emphasizes that he is only speculating on the mechanism

---

[13] Elsewhere in the *Essay* Locke describes the rack as causing pain that "possesses our whole Mind" (II.xxi.53, p. 268).

[14] It should be noted that Locke also accords ideas a tendency towards natural decay, writing arrestingly, "our Minds represent to us those Tombs, to which we are fast approaching; where though the Brass and Marble remain, yet the Inscriptions are effaced by time, and the Imagery moulders away" (II.x.5, pp. 151–152). This perhaps explains time's ability to alter connections between ideas, even associative ones.

that underlies the generation of association despite the "constant vicissitude" of ideas in experience.

Nonetheless, physiological language is invoked more often in discussions of association than other epistemological errors. Locke attributes enthusiasm, which I show in the following chapter to be a result of association, to "a warmed or over-weening brain," a problem in the "vigor of mind" (1996, p. 218). It is never clear whether Locke believes these cerebral dysfunctions to follow from other causes, or themselves to be sufficient etiological accounts of association. Chance physiological dysfunctions may be all that can be said about those cases in which association occurs unexpectedly "in the steady calm course of […] Life" (II.xxxiii.4, p. 395). Many of the examples Locke gives of association, however, are initiated by a strong feeling: disgust upon being sickened by honey, grief at the loss of a child or terror upon hearing a ghost story.[15] A fleeting grimace by a friend, for example, experienced as hurtful, can lead to the terminal association of a beloved face with anger and fear, and the friend can become an enemy. While Locke does not explain the role of the passions in association he notes their ability to completely suspend the "liberty of thought," that is, the intellectual labor of correcting our ideas that allows for free and rational decision-making (II.xxi.53, p. 268).

As noted at the beginning of this section, the ambivalent role physiology plays in Locke's account of madness is unusual for his time. The following section will explore what led Locke to develop a fundamentally ideational account instead of adopting a neurological one like so many of his contemporaries. I argue that, rather than being motivated by Locke's epistemological

[15] See Charland (2008). Rapaport also notes the importance of the passions in Locke's theory of association (1974, p. 68).

concerns, his turn away from physiology was part of a broader stance about medicine that itself had a significant impact on his philosophical commitments.

2.4 AN IDEATIONAL ACCOUNT OF MADNESS

Charles Wolfe has noted the frequent treatment of Locke as an empiricist set on re-founding the philosophy of mind on a physiological basis — from Kant, who calls Locke's project a "physiology of the understanding," to contemporary scholars who see him as a natural-philosophical midwife easing the birth of neuroscience. I agree with Wolfe that "such readings […] are […] *not a matter of interpretation* but are simply mistaken" (Wolfe, forthcoming). A quick comparison of Locke's *Essay* with the likes of Thomas Willis's *Two Discourses Concerning the Soul of Brutes* (1672) shows that, for his time, Locke spared relatively little ink for neurophysiology. In line with Willis, Locke notes that ideas are caused by "different degrees and modes of Motion in our animal Spirits, variously agitated by external Objects" (II.viii.4, p. 133).[16] But where Willis devotes an enormous volume to etiological accounts of cognition and psychopathology in terms of the animal spirits, Locke instead turns his gaze to "All those sublime Thoughts, which towre above the Clouds, and reach as high as Heaven it self" (II.i.24, p. 118). Here I argue that, while not physiological, Locke's language in the *Essay* is decidedly *medical* in the *cultura animi* sense.

---

[16] While Locke does not expand on this superficial physiological account, it clearly follows that of anatomical authorities like Willis and Descartes, the latter of whom describes the actions of the animal spirits in sensation as analogous to the rope of a bell which is pulled to transmit a signal (1985, p. 333) and who both describe the animal spirits flitting through the nerves to the brain to carve tracks in its soft matter, preserving memory or generating fancy.

In the early modern period theories of mental illness underwent a sea change, part of wider development in medicine in line with the mechanistic turn in natural philosophy (Porter 1983). Out of dissatisfaction with Scholastic explanations came a desire to explain the properties of substances and bodies in terms of their underlying structures, rather than in formal terms, such as the *virtus dormitiva* Molière famously parodied in *Le Malade Imaginaire* (1673). The rise of iatrochemical explanations — that is, those integrating alchemical theories into medicine — challenged Galenist accounts of the humoral body that had been dominant for a millennium and a half. The vitalism of Paracelsus and Jan Baptist van Helmont was invoked to explain organic matter as spirited, driven towards development and pathology by active agents that reflected the macrocosmic processes in the microcosm of the body. Van Helmont's theories of the *archei*, seeds or generative principles that caused not only the reproduction of matter but also of disease, localized illness within the body and challenged the holistic approach of the Galenist, with his emphasis on environment and on temperament.

The correspondence of Locke, as well as founding members of the Royal Society like Robert Boyle and Henry Oldenburg and later Isaac Newton, reveals a fascination with the alchemical processes generating spirit through complex relations of the earthly and celestial (Clericuzio 2009, p. 56).[17] The fundamental methods of chemistry — distillation, fermentation, and fusion — became the experimental method of the physic as well as the bases for theories of the body. At the same time, the quantitative methods developed by advocates of the new mechanism also found a foothold in medicine, sometimes integrated with chemical approaches to generate a complex array of novel approaches to medicine (Wolfe 2011).

---

[17] For a discussion of the importance of alchemy to Locke, see Anstey (2011); Walmsley (1998); Milton (2001).

New treatments of madness grew out of the interchange between physicians such as Willis and the early members of the Royal Society, leading to the integration of mechanistic explanations into treatises on mental pathology and its therapies (Brown 1970). What has been called the iatromechanist approach grew in popularity after the *Principia*, culminating in the work of Archibald Pitcairne and leaving a mark on 18th-century writers like George Cheyne and William Cullen (Brown 1987; Wood 2003, p. 106). Taking their cue from Descartes' hydraulic picture of the nervous system, iatromechanists sought to explain mental disorder in terms of the movement of the nervous fluids. At Oxford in the late 1650s Locke became interested in these new currents in natural philosophy and studied the lecture notes of Willis, whose groundbreaking work of neuropathology, *Pathologicae Cerebri,* would be published in 1667. Willis described the brain riddled with the animal spirits, "Substances highly subtil, and Aetherial Particles of a more Divine Breathing, which our Parent Nature hath hid in this Sublunary World, as it were the Instruments of Life and Soul, of Motion and Sense, of every thing" (1684, p. 3). In line with the Epicurean tradition Willis posited a sensitive soul as well as a rational one, and assigned perception, imagination, memory, and judgment to the former (Wright 1991).[18]

Physiological accounts of the mind remained compatible with Christian doctrine because they posited a rational soul that remained untouched in madness, but which was overrun by the diseased brain that housed the lower faculties. For example Daniel Sennert, whose work Locke studied intently during his early days as a doctor, posited that damage to the imagination made reason unable to perform its duties — working within a Galenic framework, he attributed this damage to the darkening effects of the black humor of melancholy (Walmsley 2004, p. 16).

[18] It has been noted that Locke was influenced not only by Willis's iatrochemical theories of the sensitive soul but also by his moral-philosophical approach to mental disorder (Schmidt 2004).

Similarly in so far as he recognized (though did not much discuss) mental substance, Willis could agree with Descartes that this rational part continued on through mental infirmity and finally death, immortal, immaterial, and unchanging. While Willis replaced the humoral explanation with a spiritous one, and Descartes mechanized it, both argued that errors in perception — a material cause — impacted the mental. Thus, as Akihito Suzuki has put it, "John Locke was simply echoing contemporary medical opinion, when he wrote that madmen had not lost the faculty to reason but 'by the violence of their imaginations, having taken their fancies for realities, they make right deduction from them'"(1995, p. 423).

On the grounds of this exclusion of the rational substance from the early-modern investigations of psychopathology, Suzuki argues that by the late 17th century mental phenomena were "not a proper medical concern." He takes Locke's contemporary Walter Charleton as exemplary of this stance, quoting his remark in 1680 that "thinking, knowing, judging, reasoning, concluding electing, willing […] being remote from the province of anatomists, I leave them to be handled by philosophers inquiring into the nature of the soul" (1992, p. 96). Suzuki concludes that by the end of the early modern period physicians and philosophers were studying fundamentally different objects, and, *pace* Kant, places Locke firmly in the philosophical camp, despite his professional commitments to the other side.

This bifurcation, however, ignores a third way: the *cultura animi* tradition. As described above, Corneanu argues that scholarly attention to the theoretical turn in early modern philosophy has overshadowed the continued interest in *askesis* alongside *theoria*. Following Bacon's lead, natural and moral philosophy was applied by the likes of Locke, Boyle and Glanvill (as well as lesser-known physician-philosophers like Thomas Wright and Edward Reynolds) to the medical project of discovering regimens of the mind that could cure affective

and cognitive diseases.[19] The *cultura animi* perspective emphasized the importance of utility in physic, in which "knowledge that is 'practical' can perform the required transformations in the soul, while 'speculative' knowledge fails in that endeavor and is thus sterile" (Corneanu 2012, p. 58).

Locke's allegiance to this tradition helps explain how, while his motivations are not physiological, they remain *medical*. As noted above, association is, for Locke, a medico-moral disease of the soul, which attacks the perceptive faculties rather than the active powers of the understanding. Locke uses the language of disease to describe not only problems with passive perception but also a plethora of "weaknesses and defects of the understanding," such as mind-wandering, hasty judgment, and prejudices, that keep men from the truth. "Of these there are as many possibly to be found," Locke writes, "if the mind were thoroughly studied, as there are diseases of the body, each whereof clogs and disables the understanding to some degree, and therefore deserves to be looked after and cured" (Locke 1996, p. 187). This medical language was no mere literary flourish by an ignorant humanist. There is some debate about when Locke began to study medicine in earnest, but it is clear from his notebooks that by the late 1650s he was deep into the medical literature popular at Oxford (Milton 2001). By 1670 he was practicing medicine, and would continue to do so throughout his life, mostly as a favor to friends.

The absence of any deep engagement with neurophysiology in the *Essay* calls for explanation in the work of a man devoted at the deepest levels to medicine and to natural philosophy. The answer, I think, is to be found in Locke's particular medical orientation which,

---

[19] The tradition can be traced back further to the Augustinian conflation of illness and moral weakness (Augustine calls God "my most private Physician") (Corneanu 2012, p. 52). Burton offers an interesting early case of the treatment of both the soul and body to redress a moral and mental failure, and the theme is well explored in Shakespeare and other literary figures.

unlike that of his peers, was staunchly anti-physiological. Given this orientation, while Locke was very much concerned with mental illness, he conceived of it in ideational terms as a widespread problem of development, one that had profound repercussions for human freedom and personal responsibility.

Before it manifests as a metaphysical quietism in the *Essay* (Yolton 1956, p. 97), Locke's aversion to neurophysiology is visible in his earliest and deepest methodological commitments. From his early medical writings it is clear that Locke found speculative natural philosophy problematic specifically in circumstances where evidence from observation and experiment could not validate theory (Anstey 2011, p. 223; Anstey and Burrows 2009, p. 15). There is good reason to think that Locke's refusal to "meddle with the physical consideration of the mind" (I.i.1, p. 43) was not because a medical avenue of investigation lay outside of his project, but rather because he thought physiology was the wrong way, *tout court*, to obtain answers about human understanding and its disorders.[20]

The grounds for such a position can be found in two of Locke's medical fragments, "Anatomia" (1668) and "De Arte Medica" (1669). There is vigorous debate in the literature about the authorship of these fragments, which are written in Locke's hand but have strong resonances with the contemporaneous work of the Puritan physician Thomas Sydenham, Locke's friend and close medical colleague.[21] These debates are part of a wider one over the direction of influence between Sydenham and Locke, and to review and weigh in on them would take its own

---

[20] Nonetheless, his quietism on this front left the door open for the fleshing out of his developmental account of mental illness with neurophysiological content in the 18th century by the likes of David Hartley.

[21] While all parties agree the hand is Locke's, this does not resolve the matter because there is evidence that Locke worked as Sydenham's amanuensis and copied some of his essays into his commonplace books. See Anstey and Burrows (2009) for a review of the debate — though it is not a neutral one, since the authors seek to argue definitively against Sydenham's authorship.

chapter. The former's loyalty to empirical medicine, based on observation and experience rather than hypothesis or the practice of anatomy, becomes much more ardent and eloquent after his friendship with Locke, and his enthusiasm for Hippocrates and Bacon only then develop into a personal philosophy of medicine. In his mature works Sydenham embodies the ancient empiricist tradition, powerfully rejecting speculation and claiming that all ideas about pathology should be based on experience, since "we know what our five senses teach […]. The whole philosophy of medicine consists in working out the history of diseases, and applying the remedies which may dispel them; and Experience is the soul guide" (Wolfe 1961, p. 196). At the same time, Locke expresses admiration and gratitude for his friend's innovative "way of the Physic" throughout his works and correspondence; and the timeframe of Locke's developing thoughts on natural philosophical explanation and epistemology allow for Sydenham's medical skepticism to have impacted Locke during the early stages of his philosophical career (Romanell 1984).

Despite diverse attempts to settle decisively the issue of influence and of the authorship of the fragments, all that is certain is that the two men shared a close collaborative relationship, in the clinical setting and also in their written works.[22] In any event, it is clear not only that Locke

---

[22] A possibility under-explored in the extensive debate over the dubia is that different standards of authorship in the early modern period make the question about authorial ownership, as posed today, not appropriate. Thus, while agreeing with the contemporary consensus that Locke is writer of the texts based on their stylistic similarity to his work and dissimilarity to Sydenham's as well as recent evidence from computational analysis, I believe it is impossible to ignore the evidence that Sydenham, not Locke, undertook (and left unfinished) a project of the magnitude suggested in the synopsis given in "De Arte Medica"; and that the authorial modifications in Sydenham's hand to Locke's draft of "Anatomia" suggest that he had input into the content of the draft; and that Sydenham was critical of speculative philosophy from before his friendship with Locke (Walmsley 2013). On the other hand, there is also strong evidence that Locke helped Sydenham write the more philosophical sections of his *Observationes medicae* (Meynell 2006). Locke's penchant for co-authorship is evident in his work with Shaftesbury on *The Fundamental Constitutions of Carolina*, *Letters from a Person of Quality*, and other works. Ultimately I agree with Milton that "the factors that make the task [of assessing authorship] so intractable also to

shared Sydenham's position on speculative hypotheses in medicine, but that this position was part of a larger allegiance to the iatrochemical school of physic (Anstey and Burrows 2009, p. 10). Among the new criticisms of Galenic medicine popular among iatrochemical physicians was the rejection of dissection as a method of medical knowledge-creation (Wolfe 1961).[23] The knowledge necessary for the general practitioner was all available at the bedside.[24] Corpses and animals could satisfy the curiosity of the natural philosopher, but were irrelevant for the object of medical study — the human body in vital motion. The author(s) of "Anatomia" argue along these lines,

> "All that anatomy can doe is only to shew us the grosse & sensible parts of the body, or the vapid & dead juices, all which after the most diligent search will be noe more able to direct a physitian how to cure a disease then how to make a man." (Sydenham 1966, p. 85)

While useful for helping a young physician or surgeon get oriented in the human body, gross anatomy is categorically incapable of revealing the "organicall constitution and that texture whereby [the body] operates." Sydenham's own writing echoes this: "However much, by

---

some extent reduces its importance" (Milton 2001). Thus I will refer to the author(s) of the fragments rather than rejecting the possibility that the conceptual content, though not the wording, was in part Sydenham's.

[23] While Anstey (2011) has criticized Wolfe's account by noting that the rejection of gross anatomy was popular amongst other "chymical" physicians in Locke's day, he does not respond to Wolfe's contrast between Locke and Boyle on the specific question of medical anatomy. I believe that Locke's opposition to anatomy in the fragments is best read as a statement about the role of experimentation *in medicine*, which is less problematic for Anstey's larger project than Wolfe's broader observation that Locke and Boyle are opposed regarding anatomization *in toto*.

[24] This view was *not* shared by the majority of the advocates of the new science outside of medicine, such as many of the members of the Royal Society, and by many important physicians of the era, notably William Harvey. For two (quite different) taxonomies of these diverse orientations within the new science see Anstey (2011) and Wolfe (2010).

seriously inclining our minds, we may discover what Nature does, and by what organs she does it, the way in which she does it will always be unknown to man" (Wolfe 1961, p. 210).[25]

"De Arte Medica" acknowledges that it is part of human nature to seek out causes rather than simply observing "the operation of nature and the event of things" (Dewhurst 1966, p. 80). It offers a creation story for the hypotheses of contemporary medicine in which earlier thinkers "were curious in imagining the secret workmanship of nature" and fashioned these fantasies into a system which ultimately "has at last but confined and narrowed mens thoughts, amused their understanding with fine but useless speculations." He who would build up medicine on axioms in order to make it into a science can pride himself on having "enlarge[d] the art of talkeing," but has done little to improve the practice of physic. Importantly, "De Arte Medica" places medicine firmly under the heading of *techne* rather than *episteme* — in explaining why a physician has no need of chymical knowledge, it analogizes him to a cook, who does not owe "his skill in rosting and boyling to his study of the elements" (Sydenham 1966, p. 81). Throughout his career Locke discusses the importance of not applying the faculty of the understanding to tasks beyond its God-given purview. In "De Arte Medica" can be found an early application of this stance to medicine, in which it is explicitly excluded from scientific endeavor.[26]

---

[25] Wolfe notes that Sydenham's passages on medicine reflect his wider commitments both in natural philosophy and in theology. An ardent Puritan, he rejected attempts to investigate the "abyss of causes" because it lay outside God's plan for us. The mature Locke expresses similar sentiments routinely, as discussed below.

[26] The framework, and even the language, of this point is obviously Baconian. The next sentence continues, "The begining and improvement of the useful arts, and the assistances of human life, have all sprung from industry and observation; true knowledg grew first in the world by experience and rationall operations; and had this method beene continued and all mens thoughts been imploid to adde their own tryalls to the observation of others noe question physick, as well as many other arts, had been in a far better condition then now it is."

To the frustration of some later commentators,[27] "Anatomia" rejects any possibility of science and technology improving such as to reveal, through tools like the microscope, the "minute and insensible" parts by which "nature performs all her operations on the body." But it must be recalled that Locke and Sydenham were committed to the alchemical worldview, within which mechanical parts were simply insufficient to explain organic phenomena without the vital seminal ferments, or active powers, that caused development. "Morbus," a fragment written by Locke in 1666, expresses the importance of seminal principles alongside "the bare mistion[28] of ye parts" for the understanding of organic development and suggests that their study could generate a "more rationall theory of diseases" than either the Galenic or Paracelsian system (Walmsley 2000). Similarly, "There is something therefore in the body and juices," the author(s) of "Anatomia" write, "too curious and fine for us to discern […] *intus mens agitat molem*" — the mind inside that moves the matter (Sydenham 1966, p. 92). The author(s) of the fragment note(s) that whether nature works on the minutest levels through "organicall texture" or a "kinde of ferment" is still an open question.

Could any investigations into the physiology of the human brain, by necessity (in Locke's day) performed on dead tissue, be revelatory of the functioning of the understanding? It seems even as early as the late 1660s, a physiological approach to psychopathology would not have been viable for Locke. The author(s) of the fragments hold(s) that ideas are intimately involved in the functioning of the soma — "Anatomia" notes the physiological changes accompanying "the receipt of suddaine bad news" (Sydenham 1966, p. 91). Yet even "the quickest sighted

---

[27] See, for example, Wolfe (1961). The stance does seem particularly strange given that Robert Hooke and others were already detailing the "minute parts" of organic matter using microscopy in the 1660s.

[28] That is, "mixture."

anatomist, assisted too by the best microscope," could never locate "a fright which causes such diseases as epilepsies, hysterical fits, and fatuity." Any natural history of the understanding must, accordingly, be reflective, based on perceived experience of the understanding "from the inside" rather than on physiology. The mature Locke later writes, "For my design being, as well as I could to copy nature, and to give an account of operations of the mind in thinking, I could look into nobody's understanding but my own, to see how it wrought […]. All, therefore, I can say of my book is, it is a copy of my own mind, in its several ways of operation" (Locke 1889, p. 345). What Locke observed when his understanding turned upon its own operations was not the machinations of the fleeting animal spirits, but rather the perception of ideas.

Locke's early skeptical views, expressed in the popular language of empiricist physicians, were foundational for his mature philosophy more generally. Because God gave us senses fitted to the macroscale on which we live, Locke argues in the *Essay*, our knowledge will always be imperfect and we will never be able to understand the complex processes by which the microscale constitutes the macroscale. This distinction between the underlying hidden corpuscular structures of the world and the qualities of them we perceive — the real and nominal essences of substances, respectively — is at the heart of his metaphysics. To believe we have found real essences of substances is, Locke suggests, impiety. In emphasizing observation and experience over speculation into hidden causes, Locke invokes the Baconian project, and indeed he identifies his own project as that of offering a "*History of the first beginnings of Humane Knowledge;* whence the Mind has its first Objects, and by what steps it makes its Progress to the laying in, and storing up those *Ideas*, out of which is to be framed all the Knowledge it is capable

of; wherein I must appeal to Experience and Observation, whether I am in the right" (II.xii.15 p. 162).[29]

Explain how sensations and perceptions interface with the understanding; "make it intelligible"; "and then," Locke quips in the *Essay*, "the next step will be to understand Creation. For the giving a new determination to the motion of the animal Spirits (which some make use of to explain voluntary motion) […] leaves voluntary motion as unintelligible as it was before." In language that echoes that of the earlier fragments, he insists that to imagine we can understand the operations of our own minds is hubris, "an overvaluing our selves" (IV.x.19, pp. 629–630). In his response to Norris, for example, Locke notes again that all we can really know of ideas is that they are the result of motion. Locke's motivation for the claim is that it is "so hard and almost impossible to keep in our minds the same unvaried idea long together" (Locke 1889, p. 469), suggesting that *something* changing rapidly causes ideas. He goes on in the paragraph to suggest the definition of ideas as "being in motion" is *the best* that can be gleaned — nothing can be known further than that ideas "are perceptions we experiment in ourselves."

While carefully cataloging the diverse processes of the understanding as experienced through reflection, Locke consistently excludes speculation into the aspects of the understanding that cannot be perceived through either sensation or reflection: namely, the inner workings of the living brain, the "mind's Presence-room" (II.iii.1 p. 121). While he often uses the language of contemporary neurology, echoing Descartes and Willis, to describe the "natural Causes and manner of Perception" (II.xiii.4 p. 133), terms like "animal spirits" have little explanatory power,

---

[29] Indeed, in his defense of his *Essay* against the Bishop of Worcester, Locke goes so far as to protest that he is not, in fact, attempting a new architectonic epistemology at all but merely a history: "that which your lordship calls my *new way by ideas* […] if it be new, it is but a new history of an old thing" (1889, p. 343).

merely working to register the constitutive role of the body rather than describing mechanisms of the understanding.[30] Locke professes his own way "dull" and "unphilosophical" in this regard, noting that because he theorizes about the nature of ideas based solely on what can be culled from sensation and reflection, he can conclude little more about their nature but that they "depend on, and are some way or other the effect of motion" (Locke 1889, p. 469).

Nonetheless, Locke's mention of animal spirits (especially in Chapter 33, on association) has excited the attention of scholars, and Sutton is representative in claiming that Locke's recourse to animal spirits in describing association is a "far cry from his official neutrality about the operations of the mind" (2010, p. 247). Wright (1987) maintains that Locke holds a "psycho-physiological account" of association, inspired by his reading of Malebranche, who discusses the connection between neurological and mental phenomena explicitly. Wright points out that in the same letter to Molyneux where he first mentions his new chapter on the connection of ideas, written in 1695, Locke mentions his recent work on Malebranche in his response to Norris. Locke does not discuss Malebranche's views in his letter — in fact, Wright's claim that "Locke notes that he has almost completed his study of Malebranche's philosophy" is a bit misleading, since Locke only notes in response to Molyneux's query that he in fact gave up the study he began years before,[31] and in a letter a month prior states that he has no plans to ever return to it

---

[30] For example: "If then external Objects be not united to our Minds, when they produce *Ideas* in it; and yet we perceive *these original Qualities* in such of them as singly fall under our Senses, 'tis evident, that some motion must be thence continued by our Nerves, or animal Spirits, by some parts of our Bodies, to the Brains or the seat of Sensation, there to *produce in our Minds the particular* Ideas *we have of them"* (II.xiii.12, p. 136). Other examples from the *Essay* can be found in II.x.5 and II.xxvii.13).

[31] "My opinion of P. Malbranche agrees perfectly with yours. What I have writ concerning seeing all things in God, would make a little treatise of it self. But I have not quite gone through it, for fear I should by somebody or other be tempted to print it" (Locke 1979, p. 352).

"being myself fully satisfied about it" (1979, p. 287).[32] However, support for Malebranche's influence on Locke's theory of association can be drawn from the fact that Locke had first read Malebranche's *De la recherche de la vérité* two decades earlier in France (Milton 2001, p. 232), around the time he made his first notes about false ideas causing madness in his notebooks.

As Wright emphasizes, Locke's new vocabulary in his chapter on association bears a strong resemblance to that of Malebranche, who writes about "natural" and "unnatural" connections of ideas. However, Malebranche's natural connections are those that link traces that the animal spirits leave on the brain in response to external stimuli to the ideas of those external stimuli. He defines unnatural connections as the happenstance occurrence of a thought along with a particular stimulus to which it bears no relation: "Our having had certain thoughts at a time when there were some new traces on our brain often suffices to make it impossible for these traces to recur without our having the same thoughts again."[33] Finally, Malebranche recognizes connections due to "the will of men" (1997, p. 103). He emphasizes that natural connections strike the mind more deeply due to its repeated exposure to them, and this is as it should be. On the other hand "All the connections that are not natural can be and should be broken, because different circumstances of time and place are bound to change them so that they can be useful to

---

[32] Evidence from his archive reveals he left off working on his draft of the manuscript that would ultimately be published as "An examination of P. Malebranche's opinion of seeing all things in God" (1706) in 1693 — quite definitively, since the last words of the manuscript are "Thus far 1693." Locke never returned to the manuscript.

[33] In *The Passions of the Soul* Descartes also recognizes perceptions that "differ from the others" insofar as "our will is not used in forming them." These "imaginings" therefore "cannot be numbered among the actions of the soul, for they arise simply from the fact that the spirits, being agitated in various different ways and coming upon the traces of various impressions which have preceded them in the brain, make their way by chance through certain pores rather than others" (1985, p. 336).

the preservation of life" (1997, p. 106).[34] Locke adjusts the vocabulary to refer to connections *between ideas*, rather than connections joining objects to perceptions of qualities. Nonetheless, his theory is in line with Malebranche's insofar as it contrasts the natural and constant perception with the contingent, and the individual perception with the universal.

While the connection with Malebranche is an important one, what Locke takes from that source is ideational, rather than physiological, and the above discussion demonstrates that calling Locke's account "psycho-physiological" is to exaggerate the case — his description of habitual associations as the result of the animal spirits in this chapter merely functions to reinforce what he says throughout the *Essay*, that there is surely some physical correlate for the passage of ideas through the mind in chains. In this respect the mention of animal spirits is no more explanatory than it is elsewhere in the *Essay,* where Locke often uses the motion of the animal spirits interchangeably with more abstract descriptions of the temper of the brain or the "motion" of ideas (e.g. II.x.5, p. 152, II.xxvii.27, p. 347, Locke 1996, p. 12, p. 226). In most cases, Locke believes "giving a new determination to the motion of animal Spirits […] clears not the difficulty one jot" (II.xxi.55, p. 269). It is the nature of perception, described through subjective observation of the understanding by the understanding, that constitutes Locke's method of natural history, and it was this aspect of Malebranche's thought that may have excited his attention.

---

[34] Malebranche's own theory of mental illness interestingly distinguishes "visionaries of the senses," who see things that are not present, from "visionaries of the imagination," who are "excessive in all instances; they raise up what is low; they exaggerate what is small; they bring near what is remote. Nothing appears to them as it is. They wonder at everything, they exclaim about everything without judgment or discernment" (Malebranche 1997, p. 165). Malebranche suggests that the difference between the former and the latter is that "visionaries of the imagination" can pass as sensible, whereas the fantasies of "visionaries of the senses" are easily recognized: "everyone has reason for not believing that one can so easily become a cock or a king" (Malebranche 1997, p. 180).

Given Locke's commitment to medicine as a curative art that is distinct from — indeed in cases hindered by — physiology, his abandonment of neuroanatomy is compatible with his treatment of association as medical. From Bacon onwards, the early-modern fascination with the moral betterment of the self through reason bled into medical practice, foreshadowing the obsession with mental hygiene that overtook England in the 18th century. Hooke writes of the importance of "rectifying the operations of the Sense, the Memory, and Reason," while Descartes, who saw medicine as one of the great aims of natural philosophy (1985, p. 151), emphasized that using reason to control the passions is key to moral as well as physical health.[35] Likewise attending to the two sorts of human ailments, medical and spiritual, became a single task for Locke, managed through the attention to ideas via the powers of the understanding, rather than the mind's material or immaterial substances. This focus on the right conduct of the understanding places Locke within the early modern movement that saw natural science as still committed to the ancient aim of philosophy: the living of a better life (Corneanu 2012; Garrett 2013). This context, in tandem with his suspicion of psychophysiology, led him to generate a highly original account of mental illness.

2.5 FREEDOM AND ASSOCIATION

I have shown that in his chapter on the association of ideas Locke presents a novel theory of madness as a pathology of ideas rather than of substance, recognized through phenomenology rather than physiology and treated through an original integration of philosophy and physic.

[35] See, for example, Descartes' letters to Princess Elizabeth, especially (1985, pp. 249–251).

Locke does not give a mechanistic account of association as a pathology of the sensorium, nor does he contrast it with a nonmaterial treatment of healthy ideas as mental substance. Instead he characterizes madness as a result of the formation of ideas by forces outside the control of the understanding, ideas which enter the understanding through its receptive capacities which are most closely reliant on the sensorium: sensation and reflection. While psychopathology is no less ideational than healthy cognition, associated ideas are not the product of intellectual labor.

Besides clarifying Locke's unique theory of association, this account can shed light on broader themes in his philosophy. This is so because Locke frequently uses madness as a trope, and understanding what he meant by it can clarify other areas of his thought such as his views on the ethics of belief (see Chapter 3) and personal identity (see Chapter 4). More immediately, associated ideas can act as a foil for healthy ones, and reveal to us what, for Locke, constitutes normal psychological functioning. The pathology that results from the automatic and passive connection of ideas makes clear that for Locke, mental health is reliant the active labor of the understanding. Indeed, Locke notes explicitly that the labor the healthy, rational mind performs is closely tied to human freedom:

> "If to break loose from the conduct of Reason, and to want that restraint of Examination and Judgment, which keeps us from chusing or doing the worse, be Liberty, true Liberty, mad Men and Fools are the only Freemen: But yet, I think, no Body would chuse to be mad for the sake of such Liberty, but he that is mad already." (II.xxi.50, p. 265)

Locke has already told us that all the liberty man has lies in his ability to suspend his assent and correct his ideas. Here he notes the irony of his account, in which the constraint reason places on the connections between our ideas is constitutive of human freedom. This is the

heart of Locke's concept of freedom, which is explicitly compatible with determinism — Locke writes that the determination of our will by our active faculties "'tis not an Abridgment, 'tis the end and use of our *Liberty*; and the farther we are removed from such a determination, the nearer we are to Misery and Slavery" (II.xxi.48, p. 264). Indeed, "every Man is put under a necessity by his constitution, as an intelligent Being, to be determined in *willing* by his own Thought" (ibid.). Even God, Locke maintains, cannot but choose what is good, not despite but *on account of* his absolute freedom.

While passages like this suggest that Locke equates freedom with the ability to arrive at the good through intellectual labor, there is some ambiguity. It is clear that somewhere in the process of suspending the will and correcting his ideas, the rational man becomes free, but it is not obvious where his freedom is located. Is it in the act of suspension, or the act of correction? Locke writes, for example,

> "If the neglect or abuse of the Liberty [man] had, to examine what would really and truly make for his Happiness, misleads him, the miscarriages that follow on it, must be imputed to his own election. He had a Power to suspend his determination: It was given him, that he might examine, and take care of his own Happiness, and look that he were not deceived." (II.xxi.56, p. 271)

Some have read this passage to imply that the power of suspension can be equated with liberty, but recently Walsh (2015) has offered a careful textual analysis of the language Locke uses to describe suspension (as a "source" or an "inlet") in order to suggest that suspension is merely the means to freedom, not constituent of it.

The text remains obdurate, but my discussion can offer some further support for Walsh's interpretation. Locke notes that the suspension of the will until deliberation has taken place only

belongs to "intellectual Beings," and follows naturally from "the inclination, and tendency of their nature to happiness" — "an obligation, and motive to them" (II.xxi.52, p. 267). As noted above, Locke believes the intellectual faculties of the madman to be perfectly intact, in contrast to the idiot.[36] Nonetheless, Locke is explicit that madmen are not free in any meaningful sense — and as we will see in Chapter 4, this is supported elsewhere by his assertion that they cannot be judged as moral agents. While we have no reason to think that madmen cannot suspend their assent, Locke is clear that they cannot do work on their ideas. Therefore freedom can be defined in terms of that activity that the madman is incapable of: not the power to suspend judgment, but the means to make use of this suspension through the scrutiny and correction of ideas. Locke's account of association endorses an interpretation of his theory of freedom as constituted by intellectual labor.

If this is right, the association of ideas has substantial ramifications for Locke's broader theory of human nature. In the current chapter I have shown how associated ideas frustrate some of the powers in which freedom lies, namely the understanding's sundry powers that allow for the evaluation of connections between ideas. In the next chapter I show the devastating effects of association on the understanding's ability to manipulate chains of ideas, or propositions. Because associated ideas compromise the ability of the understanding to correct its own ideas, they also affect its capacity to reason probabilistically and justify its beliefs. As I show, associated ideas both directly (through complex ideas) and indirectly (through propositions) wreak havoc on moral agency, forcing Locke to rethink his ethics of belief in light of the phenomenon of madness.

---

[36] Idiots "perceive but dully," and their paltry store of complex ideas means that they "will have little matter to think on" (II.xi.12, p. 160). For more on Locke's theory of mental retardation see Goodey (1994).

3 MAD ERROR

3.1 ASSOCIATED IDEAS AND THE ETHICS OF BELIEF

Locke came of age during the English Civil War, and was profoundly disturbed by the instability and violence that surrounded him. Born into a Puritan family, his religious and political loyalties were nebulous across his lifetime. His attitude is well captured in a letter to his father in July of 1659, in which he condemns both Royalists and Parliamentarians equally for ignoring the path God has chosen for humanity, that of serenity and fellowship.[37] Douglas Casson describes the young Locke as a political and religious skeptic, writing seemingly heartfelt paeans within a few years praising first Oliver Cromwell and then Charles II as harbingers of peace (2011, p. 76). His strongest loyalty seems to have been to the abolition of the violent chaos he was born into, and he tried on a variety of different political and religious mantels in pursuit of this aim — authoritarianism as well as liberalism, Anglicanism as well as antitrinitarianism — before settling into his mature views. A crucial element of this normative program is the

[37] "I hope I am to be pardond on both sides if I am not quick sighted enough to see either that glorious fabrick of liberty and happinesse, or those goblins of warre and bloud which either side would perswade us they behold over our heads ready to drop downe on us, that which I looke to is the hand that governs all things, that manages our Chaos and will bring out of it what will be best for us and what we ought to acquiesce in, I have long since learnd not to rely on men" (Locke 1976, p. 83).

justification of belief, the active intellectual labor that goes into determining what we have reason to believe.

Casson catalogues Locke's parallel frustrations across the diverse fields to which he contributed, noting some fascinating moments of interplay. Locke was closely involved with the Royal mint, and uses the metaphor of recoinage to describe the revisions to the philosophical discourse for which he so fiercely advocated (ibid., p. 9). He also accused religious enthusiasts, whom he saw as abusing their own mental faculties as well as the credulity of others to spread religious discord, of counterfeit "coining" (ibid., p. 6). All of Locke's enemies — philosophical and religious zealots, political extremists, counterfeiters and charlatans — were in some sense guilty of the same crime: the abandonment of the authoritative powers sanctioned by God, and the debasement of the public sphere with the currencies of the private, the unlawful, the unaccredited, and the occult.

I defended the view in the previous chapter that the active powers of the understanding are, for Locke, the organs of human liberty. There I described how our *complex ideas* must be subjected to the scrutiny of the active powers, but Locke also believes that *propositions* must be labored upon to justify our belief. Our potential to come to right judgments is the direct result of our capacities to view ideas sequentially and to manipulate them, capacities given by God but is not in His image, since "we cannot say God reasons at all; for he has at once a view of all things" (Locke 1824, p. 251).[38] The application of these capacities to our propositional beliefs is a moral

[38] Locke often compares the human intellect to the divine one and, even more frequently, the angelic one. The trope of the angel allows Locke to demonstrate not only the shortcomings of humanity's natural state but also how individual men and women can rise a little higher up the chain of being. While we will never have "as clear Ideas of the radical Constitution of Substances" as the angels (III.xi.23, p. 520) and never be able "at one Glance to see the Connexion and Agreement of very many Ideas," we come closest when, "by single and slow

duty: Locke writes that "the last resort a man has recourse to in the conduct of himself is the understanding," for while the will determines action, the will is itself determined by "some precedent of knowledge, or appearance of knowledge, in the understanding" (1996, p. 171).

If man neglects to apply his intellect to the pursuit of his *summum bonum* he fails to pay "the Obedience due to his Maker, who would have him use those discerning Faculties he has given him, to keep him out of Mistake and Errour" (IV.xviii.24, p. 687). Views about theological, natural philosophical, and political matters that cannot be attributed to reasoning and judgment are condemnable, not due to their content *per se* but due to the way they were obtained — this is why, for Locke, the assumption that a proposition qualifies as innate knowledge just because we do not recall its origin is a dangerous one. Even if unscrutinized propositions are, by luck, true, they are no less sinful. To neglect the pursuit of proper inquiry and judge too hastily is to "fight against God, who is the God of truth, and do the work of the devil, who is the father and propagator of lies; and our zeal, though ever so warm, will not excuse us" (Locke 1996, p. 185).

Locke's stance on moral culpability and the active faculties of reason and judgment are of a piece with his claims about the active powers that manipulate simple ideas, and he recognizes an analogous sort of pathology that compromises our agency over our mental propositions. Along with *mad ideas*, complex ideas that are combined through associations rather than healthy connections, Locke recognizes what I will call *mad errors*, propositions that, due to the presence of mad ideas, are equally intransigent against the powers of the understanding. Locke describes mad ideas — that the body is made out of glass, or that one is king of England — that are easy to recognize. The sorts of faulty propositions that motivate his discussion of mad error, however,

---

Steps, and long poring in the dark," we assure that our ideas correspond to actual relations in the world and that our terms are used with care and precision (IV.iii.6, p. 543).

are more dangerous for being less obvious. It is these more subtle manifestations of madness that have all the feel of certain knowledge that Locke was most concerned with. In fact, it does not seem too strong to say his interest in association is motivated by his horror at how religious authorities proclaim divinely-delivered principles in order to sway their congregations and block access to the moral law. Locke needed a way to explain how so many minds could be moved against the current of reason (Tully 2009, p. 23).

Along with "Of the Association of Ideas" (Chapter 33), Chapter 19 of Book IV, "Of Enthusiasm," was the only chapter Locke added to the fourth edition of the *Essay* in 1700. While Locke recognizes divine inspiration as a source of knowledge, here he argues that a revelation is only as good as the evidence for its divine origin. Without such evidence, it is no more than madness. Enthusiasm, "which laying by Reason would set up Revelation without it," is the inevitable product of taking as foundational "the ungrounded Fancies of a Man's own Brain" (IV.xix.3, p. 698) on account of an excess of "Melancholy […] mixed with Devotion" (IV.xix.7, p. 699). Of enthusiasts Locke writes,

> "Reason is lost upon them, and they are above it: they see the Light infused into their Understandings, and cannot be mistaken; 'tis clear and visible there; like the Light of bright Sunshine, shows itself, and needs no other Proof, but its own Evidence: they feel the Hand of GOD moving them within, and the impulses of the Spirit, and cannot be mistaken in what they feel." (IV.xix.8, p. 700)

John Passmore (1978) has used the case of the enthusiast to argue that Locke's account of the ethics of belief puts him in a bind, insofar as his epistemology cannot account for cases like these in which men are wrong even when they have all the necessary evidence in front of them to see their error. Religious enthusiasm seems to be a case in which Locke's account of the automatic

action of the will in response to reasons breaks down; despite being exposed to the truth, religious zealots ignore what they perceive and continue to believe their falsehoods. Passmore writes that in his discussion of enthusiasm Locke faces his "greatest difficulties in defending his intellectualist account of belief," and concludes that Locke's treatment of enthusiasm marks a shift away from rationalism and towards a more psychological vision of belief as driven by the passions: "That dictum which in earlier versions of the *Essay* was little more than a passing remark — 'what suits our wishes is forwardly believed' — comes to occupy the centre of the stage" (1978, p. 205).

Ayers has rejected Passmore's "rather startling interpretation" of Locke's theory of belief as desire, arguing that while the passions may impede the suspension of assent, since they "are not responsible for the original conceits" we need not "suppose that Locke experienced any sweeping conversion to irrationalism" (1993, p. 112). Enthusiasts suffer from misconceptions rather than unreasonable passions; Ayers claims that Locke attributes their beliefs "not to the conventional passions or desires […] but to the imagination" (1993, p. 111). Ayers has been criticized for not defending his claim that enthusiasts are motivated by intellectual rather than emotive forces, and indeed he provides little textual support in his brief treatment of the topic (Williston 2002, p. 72). But that Ayers is right is clear from the two-stage treatment of enthusiasm that Locke offers up, describing first the presence of an intransigent idea and only in the second instance the false and *peccable* belief that the idea has been generated through revelation. Below I argue that it is this latter judgment that Locke condemns as unreasonably impressionistic, motivated by religious fervor rather than an indifferent love of truth. The first stage is simply an example of mad error, against which the active powers of the understanding are helpless, and for which the enthusiast cannot be blamed.

Both Passmore and Ayers for the most part ignore Chapter 19's associate, Chapter 33 — though Ayers dismisses the both of them in one breath as contributing little that is new to the *Essay* (1993, p. 112).[39] But Locke's discussion of association can shed important light on his theory of enthusiasm. Enthusiasm and madness are both problematic for Locke's account of human understanding insofar as the usual powers of correcting falsehoods, assigning probabilistic judgments to truth claims, and generating new knowledge are shown to be defeasible. I conclude the current chapter by showing how, in characterizing enthusiasm as in part a medical failure, Locke preserves his intellectualist account of the human understanding while clarifying the circumstances under which it is insufficient to explain human action. As Tully (2002, p. 22) notes, Locke is concerned with "the factors that do govern assent, as opposed to those that ought to," and accordingly Locke is the first to admit that in some cases human understanding fails miserably. He is interested, in other words, in the full gamut of human understanding, from rationality to irrationality, sanity to insanity, agency to patienthood. Insofar as he explains failures of the understanding as *medical* rather than *moral*, however, he is able to maintain his stance that epistemological and Christian betterment progress hand in hand as the result of self-discipline.

The next section describes how our propositional judgments are formed out of our complex ideas, and how they impact the affective relations we have with possible actions, what Locke calls the "uneasinesses" which directly determine the will. I will focus especially on Locke's theory of habitual propositional knowledge. Section 3.3 shows how associated ideas compromise the truth-value of propositional thought, like a rotten ingredient that poisons a dish even when

---

[39] "If there was anything new in the fourth edition explanations of error in both 'Of Enthusiasm' and 'Of the Association of Ideas,' it lay in the more conspicuous role of the physiology of the imagination. Even that had been anticipated in a journal entry of 1682" (Ayers 1993, p. 112).

the recipe is followed to perfection. In Section 3.4 I show how propositions that contain associated ideas, which I am calling mad errors, differ from other sorts of intellectual failures. Finally in my last section I return to what has been called Locke's ethics of belief, where my discussion will reveal that Locke was no naïve subscriber to Clifford's Principle, that "It is wrong always, everywhere, and for anyone, to believe anything upon insufficient evidence" (Clifford 1974, p. 77). I show how Locke's theory of mad errors can motivate his mature account of religious toleration, which excuses false belief while condemning the actions of enthusiasts.

3.2 THE LIGHT OF REASON

In the *Essay,* and even more directly in the *Conduct*, Locke concerns himself with finding the best regimens for strengthening the faculties that generate moral knowledge. About these faculties Locke is fairly pessimistic: for the most part we labor in a "twilight of probability" and when it comes to the essences of things our groping after knowledge "leaves us in the dark" (II.xxiii.23, p. 308). But our saving grace is the light of reason, which Locke celebrates as revelatory—it is through this light that we know the moral law. The metaphor of light is ubiquitous in Book IV of the *Essay*, where Locke discusses knowledge, and also in the *Conduct*, where Locke describes "the clear light" of the mind that leads "into truth and knowledge" (Locke 1996, p. 184). But the metaphor is most richly drawn at the very beginning of the *Essay*, where Locke compares the man who gropes for absolute certainty to

> "An idle and untoward Servant, who would not attend his Business by Candle-light, to plead that he had not broad Sun-shine. The Candle, that is set up in us, shines bright enough for our purposes. The Discoveries we can make with this,

> ought to satisfy us: And we shall then use our Understandings right, when we entertain all Objects in that Way and Proportion, that they are suited to our Faculties; and upon those Grounds, they are capable of being propos'd to us; and not peremptorily, or intemperately require Demonstration, and demand Certainty, where Probability is only to be had, and which is sufficient to govern all our Concernments." (I.i.5, p. 46)

How does reason assist with this prudential sort of discovery? Reason, Locke tells us, consists in those faculties that assist in the generation of knowledge, that is, in the "Perception of the Agreement, or Disagreement, of our own *Ideas*" (IV.xvii.2, p. 668). All ideas come through experience, but knowledge is primarily about the *relations* between ideas, and our awareness of those is a product of reason (an active power) and reflection (a passive one). Locke refers to the perception of the agreement or disagreement between ideas afforded through these methods as "seeing" (e.g., II.xvii.2, p. 669). Reason illuminates the relations between ideas that have been brought together in propositions, and can generate new propositions to provide deductions for those propositions that are initially opaque to us.[40]

When Locke describes ideas as "agreeing," he refers to ideas that, when joined in an affirmative proposition, generate a truth: for example, "gold dissolves in aqua regia." A proposition in which these ideas are in opposition ("gold does not dissolve in aqua regia") would be a false one, and reason would allow us to perceive it as such. On the other hand, ideas that

[40] Locke makes it clear that the term "proposition" can refer both to the joining of ideas (mental propositions) and the joining of names (verbal propositions). Mental propositions can exist quite independently of the latter, although it is (for obvious reasons) impossible to demonstrate this in argument (IV.v.2, p. 574).

"disagree" generate true negative propositions: "a triangle does not have four sides." Thus Locke writes:

> "For *Truth, or Falshood*, being *never without some Affirmation, or Negation,* Express, or Tacit, it is not to be found, but where signs are joined or separated, according to the agreement, or disagreement, of the Things they stand for. The signs we chiefly use, are either *Ideas*, or Words, wherewith we make either mental, or verbal Propositions. *Truth* lies in so joining, or separating these Representatives, as the Things they stand for, do, in themselves, agree, or disagree: and *Falshood* in the contrary." (II.xxxii.19, p. 391)

Knowing whether a proposition is true or false, then, results from our experience with the ideas in question, and our familiarity with their properties and relations.

Locke recognizes four ways in which ideas can agree or disagree: in regards to identity / diversity, relation, co-existence (necessary connection), and real existence. The first two allow for propositions of the form "a circle is not a square," or "an elephant is bigger than a mouse." Other propositions use "is" or "is not" to make claims about the constant co-occurrence of ideas in experience, such as "gold melts when put in fire" or "man is not an omniscient all-powerful being." The truth of some propositions can be perceived immediately, and assent occurs without the active powers of the understanding contributing at all. An example of such a proposition is that "*Three* are more than *Two*" (IV.ii.1, p. 530). Such *intuitive knowledge*, "like the bright Sun-shine, forces itself immediately to be perceived" (ibid.). In these cases, reason is not required to elucidate the relationship between ideas because the agreement between them needs no mediation to be discovered.

In other cases, knowledge of whether a proposition is a "mental truth" — that is, whether its "Ideas are so put together, or separated in the Mind, as they, or the Things they stand for do agree, or not" — is not passively delivered up by perception. In these cases, the active powers of the understanding must do their work, bringing other ideas to bear. These intermediate ideas, or "Proofs," function as middle terms, bridging two ideas in a demonstration that produces knowledge; reason lays "them in a clear and fit Order, to make their Connexion and Force be plainly and easily perceived" (IV.xvii.3, p. 669). Locke gives an example of this in an exchange with Stillingfleet, where he writes that "by the clear ideas of thinking in me [a simple idea accessed via reflection], I find the agreement of the clear idea of existence, and the obscure idea of a substance in me, because I perceive the necessary idea of thinking, and the relative idea of a support; which support, without having any clear and distinct idea of what it is, beyond this relative one of support, I call a substance."[41] This is *demonstrative knowledge* (IV.ii.ii., p. 532).[42]

When proofs are not available, the understanding instead relies on judgment, the power to make probabilistic assessments of whether our ideas are arranged so as to reflect "as in Reality Things are" (IV.xiv.4, p. 653). When judging, the understanding does not perceive the agreement of ideas but merely perceives reasons to believe that the probability of agreement is high (such as the frequent co-occurrence of two ideas in experience). Affirmation can be given on the basis of such probabilistic judgment, though it then falls short of assent. Unlike intuitive knowledge,

---

[41] As quoted in Yolton (1956, p. 255), who notes that this passage shows that Locke believes certainty of demonstration depended on the clear perception of connections between ideas, rather than the clarity of the ideas themselves. This is because truth or falsity, for Locke, is a property of propositions, not ideas.

[42] Locke gives another example in the *Essay*: "Men can determine themselves," which relies on the previous knowledge that "Men shall be punished" and a chain of intermediate ideas: "God the punisher,—just Punishment,—the Punished guilty—could have done otherwise—Freedom—self-determination" (IV.xvii.4, p. 673).

which can be produced through passive perception, judgment is always an act of the understanding, since it relies on the use of intermediate ideas to assess the probability of agreement or disagreement. And unlike either rational or intuitive knowledge, judgments are often false — we may be confident in a proposition because it seems to have a high probability when it is in fact wrong, and we may also make errors in assessing probabilities themselves due to the fallibility of sensation and reflection and the weakness of our intellectual endowments or of our will.

In establishing the truth or falsehood of a proposition, the understanding draws on three sources to fill in the connections between ideas or to assess the odds that such a connection can be drawn. The first is the person's own experiences. Knowledge is, Locke writes, "founded on particulars," and thus relies first and foremost on those connections between ideas that are preserved in the memory (IV.xii.3, p. 640). Another way to discover the relations between ideas is to increase the scope of this experiential knowledge-base by performing further "rational and regular experiments" (IV.xii.10, p. 645). And finally, the testimony of trustworthy authorities can also considered.[43] These resources are brought to bear by the active powers of the understanding during the suspension of assent. Reason, by assessing the agreement or disagreement of ideas, can generate or quell unease in the mind by elucidating connections between ideas and thus changing the grounds on which the will is determined to action. Because actions follow directly from the will, the exercise or neglect of reason not only determines whether we believe truly but also whether we act rightly.

---

[43] Locke cashes trustworthiness out in terms of the number of witnesses, their integrity, their consistency, and the uniformity of their opinion (IV.xvi.4, p. 656).

Locke pays particular attention to "maladies of assent" (Corneanu 2012, p. 90). Particularly dangerous is allowing ideas to join up in chains without attentive perception of their agreement and disagreement. This is what Locke considers a bad habit. But Locke also goes further than most in celebrating the cultivation of positive habits as transformative, both for the individual and society at large (Smith 2006, p. 832).[44] While Locke emphasizes that all propositional thought must, at the first, be attended to, he recognizes that often people benefit when the movement from one idea to the next becomes habitual — analogous to the habits of perception that allow us to see a shaded circle as a sphere, he recognizes habits of judgment as essential to our functioning as limited beings. Thus Locke writes,

> "A man is said to know any Proposition, which having been once laid before his Thoughts, he evidently perceived the Agreement, or Disagreement of the *Ideas* whereof it consists; and so lodg'd it in his Memory, that whenever that Proposition comes again to be reflected on, he, without doubt or hesitation, embraces the right side, assents to, and is certain of the Truth of it. This, I think, one may call *habitual knowledge*: And thus a Man may be said to know all those Truths which are lodg'd in his Memory, by a foregoing clear and full perception, whereof the Mind is assured past doubt, as often as it has occasion to reflect upon them." (IV.i.8, p. 528)

Habits are beneficial when they make automatic chains of ideas that have been established to be trustworthy by judgment. In Locke's developmental picture, personal habits impact the slow aggregation of true propositions that lead the individual to have a body of

---

[44] Descartes's theory of habit, for example, is similar to Locke's, but with a more negative valence — see his Letter to Arnauld, 29 July 1648 (Descartes 1983, p. 221).

demonstrative knowledge and justified beliefs: "change of Customs and Opinion bringing with it new Combinations of *Ideas*" (II.xxii.7, p. 291).[45] Likewise bad habits of mind occur when propositions formed too hastily are repeatedly assented to without being scrutinized, escaping the attention. An example of a bad intellectual habit is a prejudice, in which a false proposition is assented to without being properly scrutinized and then believed habitually, such that the act of judging it to be true happens so quickly that it escapes attention, and it appears in the mind with all the authority of an intuition. Habits in some cases are better seen as chains of ideas representing actions, rather than propositions — a process, not a product. Locke in places describes habit as the "power or ability in Man, of doing any thing, when it has been acquired by frequent doing the same thing" (II.xxii.10, p. 293). The musician plays one note following the other automatically because he has willed his fingers to move in the same pattern over the frets many times, generating, as it were, a real state of affairs which he can perceive and remember. The chain of ideas constituting the strain can become *naturalized,* treated as though intuitive, and playing the first note can bring the others tumbling along to complete the tune. So while Locke talks about thoughts as habits, this may be shorthand for referring to the habitual *perception* of chains of ideas, whether or not they form propositional knowledge or beliefs.

Since all perception is conscious, it is clear that habits must, too, be conscious — though they may happen on the very edge of the ambit of awareness. Ideas that habitually follow each other, like the steps in certain repeated actions, must at first have been conjoined with attention,

---

[45] At the same time, Locke is particularly concerned about the development of bad intellectual habits in children, since in their case the contingent and capricious nature of perception is not sufficiently countered by the active powers of the understanding. Locke writes in *Some Thoughts Concerning Education* of one unfortunate juvenile behavior that "the only cure for this, as for any other miscarriage, is by use to introduce the contrary habit" (Locke 1996, p. 107). He emphasizes the importance of, through habit and learning, improving children's natural dispositions towards belief-formation.

and Locke acknowledges that it would be absurd for the grounds of assent to be recalled every time such judgments are used as evidence in the making up of new propositions. "It suffices," he writes,

> "that [men] have once with care and fairness, sifted the Matter as far as they could; and that they have searched into all the Particulars, that they could imagine to give any light to the Question; and with the best of their Skill, cast up the account upon the whole Evidence: and thus having once found on which side the Probability appeared to them, after as full and exact an enquiry as they can make, they lay up the Conclusion in their Memories, as Truth they have discovered. and for the future, they remain satisfied with the Testimony of their Memories." (IV.xvi.2, p. 658)

Bad habits too have once been the subjects of attention, though not of *sufficient* attention. Propositions that become prejudices are thus "potential objects of contemplation" (Lähteenmäki 2008, p. 96), annexed to consciousness, even if they resist reason's attentions.

Accordingly bad habits can be viewed normatively, as the result of failures of the individual to rigorously appraise their thought: first when the original proposition was formed, and every time since as the chain of ideas slowly became habitual. Locke notes that the more habitual a physical or intellectual act is, the harder it will be to scrutinize, because it will pass more and more quickly before the understanding. Thus Locke can pun on his usual metaphor of reason as sunlight, writing that to allow bad habits to flourish is to "put our selves in the dark, or in the power of the Prince of Darkness" (IV.xix.13, p. 703). The moral failing is not the *acting* on a false idea, for Locke argues that once the powers of active perception offer up a judgment to the will, the will must act on it. It is also not a mistake of perception. Rather the sin is the failure

to apply the active powers to appraisal during the suspension of the will: "*Errour* is not a Fault of our Knowledge, but a Mistake of our Judgment giving Assent to that, which is not true" (IV.xx.16, p. 706).

In section 3.4, I argue that despite Locke's tendency to discuss mad errors within the context of the "empire of Habit," his interpreters must not conflate habitual ideas and mad error, despite their close relation. But first, the following section examines the notion of the mad error in more detail. Locke envisioned the *Essay* as a Baconian natural history of the understanding, and accordingly enumerated not only instances of its powers, but also of its failures. The centrality of the active faculties of reason and judgment in the management of propositions made their failures important targets of investigation. The reliability of intuition was also a crucial linchpin in Locke's account, and its faultiness in the case of the mentally ill was undeniable for a medical practitioner like Locke. In the previous chapter we saw that if ideas become associated during passive perception such that they are presented to the active understanding as one simple idea, they frustrate the active powers and cannot be revised. Such is the case of madmen who believe, for example, that they are the king of England, or that their flesh is made out of glass. When these ideas enter into propositions, they result in false propositions that appear, in the light of reason, to be true: "all must pay my royal personage due obsequence;" "nothing must touch my brittle skin." The reasoning in such cases is sound — the ideas really do agree — but the ideas themselves are "wrong" in the sense explained above. Thus enthusiasts believe they are receiving knowledge from God but merely chase *ignes fatui*, "For all the Light they speak of is but a strong, though ungrounded persuasion of their own Minds that it is a Truth" (IV.xixi.10, p. 702).

3.3 MAD IDEAS AND MAD ERRORS

Locke chose to add his chapter on association to Book II of the *Essay*, which is devoted to the treatment of simple and complex ideas.[46] But there is some ambiguity as to whether he saw madness as a phenomenon affecting ideas or propositions. In some cases, he seems to refer to madness as resulting from the erroneous joining of ideas into propositions:

> "In fine, the defect in *Naturals* seems to proceed from want of quickness, activity, and motion, in the intellectual Faculties, whereby they are deprived of Reason: Whereas *mad Men*, on the other side, seem to suffer by the other extreme. For they do not appear to me to have lost the Faculty of Reasoning: but having joined together some *Ideas* very wrongly, they mistake them for Truths." (II.xi.13, 161)

In this quote, madness seems to reduce to the sort of error Locke describes in his discussion of wrong assent or error in Chapter 20 of Book IV, in which he describes how judgment is ineffective against "*Propositions that are not in themselves certain and evident, but doubtful and false, taken up for Principles*" (IV.xx.7, p. 711). This reading seems to have support from Locke's earliest surviving reflections on madness, found in his medical notebooks kept around the time he was first drafting the *Essay* two decades before he wrote about association. Locke tells us that these early thoughts on insanity influenced his mature view of association: "This consideration of the thing itself [that is, madness] at a time when I thought not in the least on the Subject which I am now treating of [that is, association] suggested it to me" (II.xxxiii.4, p. 395). In July of 1676, while traveling in France, Locke wrote in shorthand in his journal, "Query

---

[46] I thank Margaret Atherton for making clear to me the import of Locke putting the chapter where he did.

whether mania be not putting together wrong ideas and so making wrong propositions from them, notwithstanding the reasoning be right? But *idiocy* is a fault in the faculty of reasoning" (Dewhurst 1963, p. 70).[47]

But there is some ambiguity in Locke's assertion that "mad Men put wrong *Ideas* together, and so make wrong Propositions, but argue and reason rightly from them" (II.xi.13, p. 161). The passage could be read as claiming that madmen choose the "wrong" ideas to assemble, thus creating wrong propositions, as the above quote seems to suggest. But it could just as well be interpreted as saying that the ideas madmen assemble into propositions are intrinsically wrong ideas — begging the question of what "wrong" means here. In the former case, madness would be due to the employment of fantastical or chimerical ideas in existential propositions.[48] For example, the idea "God-with-shape" is a fantastical idea, and a proposition about it could result in absurdity, but not necessarily a falsehood — it could take the form of a conditional, such as "*if* God had a shape, *then* He would take up physical space." What would make it a falsehood would be the addition of a claim about existence in the proposition, apart from the associated idea: "God-with-shape exists, and (therefore) takes up physical space."

The problem with this sort of reading is that it is contradictory. Locke states that madmen have not lost the faculty of reason, and that their reasoning is right. He tells us, after all, that "a distracted Man fancying himself the King" may think with perfect sanity with regard to what

---

[47] Wright (1987) reports Henry Schankula's crucial observation that Dewhurst mistranslates Locke's "fatuitas" here as madness, which in Locke's day was interchangeable with mania, rendering the sentence incoherent. A more accurate translation would be foolishness or idiocy, making it consistent with quotation that follows from 1677. While as noted idiocy is most closely analogous to the contemporary construct of mental retardation, mania and madness capture the symptoms of modern-day psychosis (hallucination along with cognitive and affective defects).

[48] That is, the class of complex ideas of substances which Locke defines as having "no Foundation in Nature, nor having any Conformity with that reality of Being, to which they are tacitly referr'd, as their Archetypes" (II.xxx.1, p. 372).

should follow from his incumbency, such as the "suitable Attendance, Respect, and Obedience" of subjects (II.xi.13, p. 161).[49] And Locke has told us that the combining of ideas into affirmative or negative propositions is the very work of reason. The second reading avoids this contradiction by viewing the ideas that make up propositions as wrong, with the propositions thus being only indirectly false. But this reading is problematic as well, since Locke makes clear that the terms "true" and "false" cannot meaningfully be applied to ideas of modes and substances, insofar as our ideas are "nothing but bare Appearances or Perceptions in our Minds" (II.xxxii.1, p. 384). Complex ideas are not able to introduce truth-value into a proposition.

A few weeks after Locke wrote the above in his journal, he returned to the topic of madness:

> "Query whether madness be not the wrong application of mad ideas to things that exist, but are neither having of wrong ideas nor wrong reasoning, and then so that it seems to exist wholly in *propositions, not simple ideas nor syllogismes*, for example [those] made in phantasy [such as] him to be either king or candle." (Dewhurst 1963, p. 71)[50]

---

[49] Some of Locke's examples of mad delusions, such as having a glass body or being a king, may be familiar to the reader from Descartes' *Meditations* (1984, p. 13). While such examples are prevalent in early modern accounts of madness (as well as contemporary ones), it is still worth considering a possible influence. Unfortunately space does not permit a comparison of Locke's theory of madness with Descartes's more physiological account, or of the very different roles madness plays in their larger projects (but see Frankfurt [2008]).

[50] The bracketed additions are Dewhurst's. In the original manuscript (MS. Locke f. 1, pp. 358–359) the selection is written partially in Locke's shorthand (he used a variant of the stenography invented by William Cartwright and widely attributed to Jeremiah Rich regularly in his journals). I am quite certain that Dewhurst has made a translation error, and in the main text have replaced his "proposition into simple ideas or syllogisms" with my own translation (between the asterisks).

In the earlier passage quoted from his journals Locke attributes madness to wrong ideas, and he does so again in the *Essay*. In this passage, however, he contrasts wrong ideas with "mad ideas," which are applied to things "that exist." Unlike the earlier quotation and Locke's mature formulation of his theory of madness in the *Essay*, this passage makes it clear that when claiming that madness is a pathology of propositions, Locke is not referring to the sort of construction of propositions that constitutes syllogistic reasoning, but a narrower class of existential propositions. Mad ideas become problematic due to their relation to things in reality, and the sort of ideas that Locke is worried about in this passage are those that refer to things in the outside world — that implicitly posit existence.

There is a place for this sort of idea within Locke's mature taxonomy. As mentioned above the *Essay* the chapter on associated ideas follows directly upon the one treating on true and false ideas, and Locke's conflation of mad ideas with propositions in the above passage would suggest that they are a subclass of the latter. As noted, Locke emphasizes that truth values can properly be attributed to propositions, rather than simple or complex ideas, insofar as they constitute facts about relations between ideas. However, he notes that the terms "true" and "false" can be useful for distinguishing a unique type of "tacit" proposition, which often contributes unnoticed to a complex idea:

> "Though Truth and Falshood belong, in Propriety of Speech, only to Propositions; yet *Ideas* are oftentimes termed *true or false* (as what Words are there, that are not used with great Latitude, and with some deviation from their strict and proper Significations?) Though, it think, that when *Ideas* themselves are termed true or false, there is still some secret or tacit Proposition, which is the Foundation of that Denomination: as we shall see, if we examine the particular Occasions, wherein

> they come to be called true or false. In all which, we shall find some kind of Affirmation, or Negation, which is the Reason of that Denomination." (II.xxxii.1, p. 384)

Thus one may refer to a manticore as a mythological creature, and in so doing simply employ a fantastical idea. A fantastical idea can be colloquially called false, however, when it is employed in a proposition that gives it existential heft. If one uses "manticore" to refer to a *real* creature that one might see on an exotic voyage, one's idea of the manticore smuggles in the property of existing, and can be assessed for its truth-value. Given the infelicity of calling complex ideas false, Locke writes that ideas such as these "may very fitly *be called* […] *wrong* Ideas, according as they […] disagree to those Patterns to which they are referred" (II.xxxii.26, p. 394). Locke's ambivalence about attributing madness to wrong ideas in the quotations above seems to be resolved by this formal definition, though it would seem in his journals he used "wrong" colloquially, rather than in the technical sense developed later.

Associated ideas are thus complex ideas that are wrongly presumed to exist, and which can cause existential claims of which they are a part to be false. Thus while Locke's chapter on association ideas is a contribution to his discussion of complex ideas, not of propositional knowledge, as the conclusion of Book II it also functions as the segue to the following books on language, knowledge and belief. For Locke makes clear that association has terrible implications for human rationality. He writes,

> "Intellectual Habits and Defects this way contracted are not less frequent and powerful, though less observed. Let the *Ideas* of Being and Matter be strongly joined by either Education or much Thought, whilst these are still combined in the Mind, what Notions, what Reasonings, will there be about separate Spirits? Let

custom from the very Childhood have join'd Figure and Shape to the *Idea* of God, and what Absurdities will the Mind be liable to about the Deity?"

In these examples a mad idea — the combination of God with figure and shape — becomes a building block for a proposition — a "Notion" or "Reasoning" — that is false. The ideas in a proposition, say, "God-with-shape takes up physical space," may agree with each other, and it may be reasonable to affirm the truth of the proposition. But perception of agreement between the ideas "God-with-shape" and "physical space" does *not,* in this case, constitute a truth, because falsehood is implicit in the mad idea of "God-with-shape."

This sort of proposition which the reason judges to be certain but which is in fact false due to the presence of a mad idea is what I mean by *mad error*. Mad errors have the same perceptual *feel* as intuitive knowledge — the light of reason shows them to be true. And indeed our reason *correctly* tells us the ideas agree, since surely if the fantastical notion of "God-with-shape" did in fact exist, it would follow that He would take up physical space. In the healthy mind, the mistaken connection between the idea of "God" and "shape" would have been ferreted out, but in the case of the associated idea, the fantastical nature of the idea is unidentifiable, its existence is tacitly assumed, and it contributes to an intransigent belief that has all the feel of certain knowledge.

In the fourth edition of the *Essay* Locke describes mad errors as a universal affliction, beginning his discussion of associated ideas with the observation that everyone is familiar with the presence of something "really Extravagant" in the reasoning of even their most sensible acquaintances. When it comes to certain topics, men otherwise rational would be "thought fitter for *Bedlam*, than Civil Conversation" (II.xxxiii.1, p. 395). A central question must be how, if mad errors are so widespread, these pathological propositions can be differentiated from intuitive

knowledge. Locke considers the possibility of mad error to be a threat to his definition of knowledge, and considers a possible line of criticism:

> "If it be true, that all Knowledge lies only in the perception of agreement or disagreement of our own *Ideas*, the Visions of an Enthusiast, and the Reasonings of a sober Man, will be equally certain. 'Tis no matter how Things are: so a Man observe but the agreement of his own Imaginations, and talk conformably, it is all Truth, all Certainty. Such Castles in the Air, will be as strong Holds of Truth, as the Demonstrations of *Euclid.* That an Harpy is not a Centaur, is by this way as certain knowledge, and as much a Truth, as that a Square is not a Circle. But *of what use is all this* fine Knowledge *of Men's own imaginations,* to a Man that enquires after the reality of Things?" (IV.iv.1 p. 563)

Locke offers two sorts of arguments to assure his reader that there is a privileged distinction in his psychology between knowledge and mad error. The first is the authority of sensitive knowledge, which is guaranteed by the fact that the mind is incapable of manufacturing simple ideas, guaranteeing that such perceptions are "ordained and adapted" to "the Wisdom and Will of our Maker" (IV.iv.4, p. 564). Simple ideas, by definition, cannot be mad. But this distinction will not satisfy the reader who is concerned about distinguishing mad errors from propositional knowledge *in practice*: to the afflicted, associated ideas can *seem* like simple ideas, so judgments about whether knowledge claims are justified cannot be made from the inside.[51]

---

[51] Though the obliviousness of the afflicted is qualified in Chapter 33, where Locke discusses the case of a man who, indulging and surfeiting on honey, is unable to disassociate the taste of it from the feeling of severe sickness. Locke's description of the makes clear that the man knows the source of his association, but he does not go so far as to say that he knows his antipathy to be irrational.

Indeed, the assurance Locke hopes to provide with his notion of sensitive knowledge will only reinforce the wild certainty of the madman regarding his own intuitions.

Locke notes that some complex ideas only refer to archetypes of the mind's own making, such that propositions involving them need not, in fact, conform to real things in order to be true. This point is not relevant to mad errors, which as described above always refer to purportedly real things, rather than the mind's own archetypes. Regarding substances, Locke emphasizes the importance of comparing relations between ideas judged to exist by the mind with new ideas, such that probabilistic judgments can be refined. For madmen, however, even ideas about substances seem certain, and as such the adjustment of their judgments on the basis of new probabilities is impossible.

It would seem, then, that the arguments Locke makes for the reality of our knowledge cannot rule out the possibility that our perceptions of truth are illusory, since they rely on our capacity to differentiate the *perception* of agreement between ideas from the mere *judgment* of agreement — a capacity that is lost in madness. This accords with Locke's claim in Chapter 33, that no process of introspection can cure the association of ideas. The specter of mental illness has profoundly bleak implications: how can anyone ever know that his or her perception of agreement between ideas is justified at all?

Locke, of course, did not intend anything like this radical skepticism, and I shall argue in what follows that he has an *in principle* defense of the reality of knowledge: it can be distinguished by the acts of intellectual labor that constitute it. The constitutional history of mad errors is far different from healthy knowledge or justified belief. This does not counter the problem of *identifying* madness, however. On that point I suggest that, by relegating the

association of ideas to the *medical* sphere by identifying it as madness, Locke relegates the threat sufficiently for his own interests.

First, however, I must address a possible concern, which is that Locke, contrary to the narrow reading I have given of association, in fact accepts the existence of unmediated thoughts in the form of *habit*. Under this reading, castles in the air — that is, chains of ideas that do not correspond to reality — include all sorts of unattended-to ideas that follow each other immediately, outside of the oversight of the understanding. Many of these unattended to chains of ideas are important for the day-to-day functioning of men and women, who, unlike angels, cannot be fully aware of every idea that passes through their minds. Before proceeding with my argument for the role of mad errors in Locke's larger project, in the following sections I present evidence for why various other types of error should not be conflated with those caused by mad ideas.

3.4 WHAT MAD ERROR IS NOT

In previous sections I described how Locke's account of madness amounts to a concession that in some cases we may perceive a proposition to be true that is, in fact, false. Yet when an associated idea is part of a proposition, the error is not, in fact, one of reason — we can perceive correctly that the subject and predicate agree, and still be wrong about the truth-value of the proposition. This is because the agreement is based on a falsehood implicit within the associated idea itself, which reason cannot recognize or correct. If this type of mad error is common in human cognition, Locke's intellectualist picture becomes worthless as a guarantor of the reality

of our knowledge. On my reading it thus becomes crucial to differentiate mad errors from other sorts of intellectual errors that are correctable.

Locke's commentators often extend his account of associated ideas beyond the bounds he gave it. Most frequently, mad errors are equated with habit. Wolterstorff, for example, writes that Locke's discussion of the association of ideas aims to highlight the confusion of "ideas produced by custom with the agreement and disagreement that they possess intrinsically" (Wolterstorff 1996, p. 161). Chapter 33 is also seen as gathering together Locke's rangy reflections on our epistemic failings: Wright writes that Locke's earlier discussions in the *Essay* of "Wrong Assent, or Error […] seems to presuppose the theory of association of ideas" (Wright 1987, p. 110). Here I will consider these interpretations in turn, and show how Locke's account of mad error is more precise than each allows for.

Turning first to the comparison of associations with habits, the first thing to remember is that habits are, for Locke, chains of ideas that "*produce actions in us, which often escape our observation*" (II.ix.10, p. 147). These acts can be either physical or mental, that is, perceptual; in the *Conduct* Locke writes of habits, "as it is in the body, so it is in the mind" (Locke 1996, p. 187). To say an action is habitual is to say that it can be undertaken outside of the ambit of attention. Locke describes two different sorts of habits, with very different etiologies. On the one hand, chains of ideas that result from *acts* of the understanding can become, we may say, *actively* habituated; this is how we form beneficial habits. Such habituated chains of ideas perform two important functions in Locke's psychology. As we have seen human freedom, for Locke, is circumscribed by an individual's ability to suspend their assent until the active powers of perception have assessed the relationship between ideas. This power, he writes in *Some Thoughts Concerning Education,* is the result of diligent labor of habituation:

> "It seems plain to me that the principle of all virtue and excellency lies in a power of denying ourselves the satisfaction of our own desires where reason does not authorize them. This power is to be got and improved by custom, made easy and familiar by an *early* practice." (Locke 1996, p. 29)

Improved habits of perception allow disciplined judgment to overcome mind-wandering (Locke 1996, p. 206) and make philosophy possible (ibid., p. 187).

Beneficial habits also allow men and women to navigate their environment without being constantly preoccupied by the exhausting enterprises of reason and judgment. Recall that Locke allows for *habitual knowledge*, in which demonstrations of the agreement or disagreement between ideas can be stored up such that the thinker need only recall the fact of the proved demonstration, rather than perceiving again every step of the proof (IV.i.8). Probabilistic judgments can also be actively habituated, such that the act of judgment can happen without the attention of the active powers of perception, almost automatically. Locke says *beliefs* can be habitual, though this must be shorthand for what he really intends: that the process of judging, that is, the stepwise performance of the active understanding, happens with little attention.

Locke also talks of habit or custom in a negative light, when propositions form not as the result of an intentional mental act but through the lazy acquiescence of the understanding — what we might call *passive habituation*. Most frequently he talks about the force of dubious authority figures in swaying the judgment of those in whom the active faculties are not sufficiently honed, like children. It is here that the case for mad errors *qua* bad habits becomes compelling, but while the relationship is a close one, they are not synonymous. Locke writes of association,

> "This sort of Unreasonableness is usually imputed to Education and Prejudice, and for the most part truly enough, though that reaches not the bottom of the Disease, nor shows distinctly enough whence it rises, or wherein it lies. Education is often rightly assigned for the Cause, and Prejudice is a good general name for the thing it self: But yet, I think, he ought to look a little farther who would trace this sort of madness to the root it springs from, and so explain it, as to shew whence this flaw has its Original in very sober and rational Minds, and wherein it consists." (II.xxxiii.4, p. 395)

While bad habits can cause mad errors, Locke emphasizes that they are not reducible to them. Most obviously, Locke notes that the reader must "look a little farther" because, as we have seen, "this sort of unreasonableness" is due to a specific sort of problem of ideas. Bad habits can form without containing associated ideas, and in this case they will be rectifiable through the efforts of the understanding, and the instantiation of inverse good habits. But while passive habituation leads to prejudices that are curable, it is also a risk factor for the association of ideas: habituation can play a role in forming associated ideas through their frequent co-occurrence in passive perception. And when associated ideas form, bad habits can lead to mad errors. For "When two things in themselves disjoin'd, appear to the sight constantly united; if the Eye sees these things rivetted which are loose," Locke asks rhetorically, "where will you begin to rectify the mistakes that follow?" (II.xxxii.18, p. 401). The only intervention for association is prophylactic: children must be protected from fantastical ideas, false principles, and other forms of indoctrination.

The close connection Locke draws between habits and mad errors may suggest to his reader that since habits are not all bad, associated ideas cannot be either. And there is some

evidence that Locke sees some "sane" chains of ideas as phenomenologically identical to mad ones once they become habituated. In his discussion of habits of perception, Locke suggests that good habits can too lead to the collapse of ideas into a complex in which the relations appear certain and necessary, and which cannot be altered. His famous discussion of Molyneux's problem is one such example, and he describes a similar case in his discussion of association in the *Conduct*.[52] Arguing for the incurability of associated ideas he writes,

> "This is for caution against this evil, before it be thoroughly riveted by custom in the understanding; but he that would cure it, when habit has established it, must nicely observe the very quick and almost imperceptible motions of the mind in its habitual actions. What I have said in another place about the change of the ideas of sense into those of judgment may be proof of this. Let anyone not skilled in painting be told when he sees bottles and tobacco pipes, and other things so painted, as they are in some places shown, that he does not see protuberances, and you will not convince him but by the touch: he will not believe that, by an instantaneous legerdemain of his own thoughts, one idea is substituted for another." (Locke 1996, p. 219)

This is the closest Locke comes to using the terms of association to refer to beneficial habits that become automatic. What he is really interested in is not mechanical connections between ideas per se (*pace* Hume, for example); but rather the failures of the active understanding that cause pathology. Thus the association of ideas is fundamentally negative,

---

[52] The brief discussion of associated ideas in the *Conduct* was written at the same time as Chapter 33 of the *Essay*, around April 1695. Locke decided to exclude it when he added the new material to the *Essay*'s fourth edition, but his amanuensis recopied it in the late 1690s when Locke began to revise the notes that would be published (posthumously) as the *Conduct*.

providing a contrast class to the active powers of the understanding that are his true fascination. If Locke's theory allows for "good" associations they remain irrelevant to his larger project, which sounds an alarum about the dangers of passive perception. Given the practical aims of his project, it is not surprising that he does not discuss immediate connections between ideas that result from active habituation, except to suggest, as he does in the *Conduct*, that perhaps those capacities we take to be innate are the result of this sort of process (Locke 1996, p. 173). His treatment of association is intended to pick out a particular threat to the understanding caused by the adoption of wrong ideas that has unique implications for personal identity, religious toleration and epistemic responsibility; the naturalization of right ideas interests him less.

Along with differentiating it from bad habits, the incorrigible nature of mad error also differentiates it from other types of wrong assent. Locke recognizes four varieties of error from the first edition of *Essay* on: the want of proofs, the want of ability to use them, the want of the will to use them, and wrong measures of probability. The first and third types of error Locke details are best characterized as epistemic failures, whose remedy he is quick to detail: "GOD has furnished Men with faculties sufficient to direct them in the Way they should take, if they will but seriously employ them that Way" (IV.xx.3, p. 708). The second and fourth are more complicated. The want of ability Locke attributes to "defects in the Organs of the Body, particularly adapted to Thinking; or in the dulness or intractableness of those Faculties, for want of use; or, as some think, in the natural differences of Men's Souls themselves" (IV.xx.5, p. 709). At its extreme, the want of ability results in idiocy, but Locke recognizes some gradient.[53]

[53] Disappointingly Locke never attends to the question of how epistemic responsibility relates to mental capacity, but in conflating those of poor mental capacity with beasts, he suggests they should not be treated as persons.

Locke's discussion of the fourth type of error, "wrong measures of probability," is the second place where earlier additions of the *Essay* can be viewed as anticipating Locke's later discussion of association. Wrong measures of probability occur when individuals believe in a proposition not in accordance with *reasons*, but rather due to some external factor that effects their assessment of the relations between ideas, such as the submission of their judgment to false propositions, received hypothesis, passions or inclinations, or the authority of others. The first of these influences Locke describes in terms now familiar, of intransigent mad error: "What is inconsistent with our *Principles*," Locke writes, "is so far from passing for probable with us, that it will not be allowed possible" (Locke 1996, p. 711). Accordingly Locke urges that we take great care in attending to what we, and our children, adopt as principles. Taking up the question of children, he notes, in language anticipating his discussion of association in the *Conduct*, that certain principles

> "fastened by degrees, are at last (equally, whether true or false) riveted there by long Custom and Education beyond all possibility of being pull'd out again. For Men, when they are grown up, reflecting upon their Opinions, and finding those of this sort to be as ancient in their Minds as their very Memories, not having observed their early Insinuation, nor by what means they got them, are apt to reverence them as sacred things." (IV.xx.9, p. 712)

Correcting such entrenched principles is beyond the power of reason.

Note that here Locke is not describing association as a case of wrong assent — "riveted" ideas may contribute to true *or* false propositions — but as a key risk factor for it. This passage emphasizes the difference between associated ideas and the wrong measures of probability that may follow from them, which I am calling mad errors. Unlike the other sorts of errors Locke

describes in Book IV, mad errors cannot be revised because they are not reducible to a wrong assessment of probability. They are not, in other words, errors of judgment, but are cases in which reason and judgment are compromised by the pathology of associated ideas. Rather than just being cases of poor reasoning, the afflicted's perceptions are in "open defiance of their senses" (IV.xx.11, p. 714).

"Having joined together some *Ideas* very wrongly," Locke writes of madmen, "they mistake them for Truths" (II.xi.13, p. 161). What can it mean, in Locke's psychology, to mistake something for a truth? As we have seen, Locke locates the risk of epistemic error in belief, rather than in knowledge. While people can err in failing to pursue knowledge, once they have knowledge it is infallible. To know a proposition to be true is to *perceive* the fact of its truth, that is, the agreement of its subject and its predicate, and Locke seems to universally maintain that we cannot perceive agreement where there is none. Thus Wolterstorff, for example, writes,

> "What would Reason's being in error consist of? Presumably it would consist in Reason, with respect to some falsehood, producing in one an experience which is phenomenologically no different from the 'just knowing' experience — this in turn producing a belief. Locke, so far as I can see, held firmly to the conviction that that never happens; as indeed he rejected the possibility that we have in us some other faculty which, with respect to falsehood, perfectly mimics the 'just knowing' experience." (Wolterstorff 1996, p. 93)

Wolterstorff maintains that in cases where people are wrong, their access to the truth is blocked on account of a stronger belief which either monopolizes their awareness or inhibits the act of perceiving itself (1996, p. 96). And yet Locke describes the effects of association precisely in terms of misleading phenomenology — a conflict between perception and reality. For the

madman, Locke writes, wrongly-connected ideas cause things to "constantly appear" other than they are, and "naturally so" — this is the central character of association that makes it "a disease of the mind as hard to be cured as any" (Locke 1996, p. 218). For two ideas to be associated is precisely for one to follow the other with the immediacy that is, in the healthy mind, reserved for relations between ideas that are necessary due to the facts of the matter. "Such unnatural connections become by custom as natural to the mind," Locke writes, "as sun and light. Fire and warmth go together, and so seem to carry with them as natural an evidence as self-evident truths themselves" (ibid.). Associated ideas appear, like simple ones, to be knowable through sensitive knowledge, and when they are parts of propositions they are as truth-functional as healthily connected ideas.

Contrary, then, to Wolterstorff's claim that "On Locke's view, it never genuinely appears to us that we are 'perceiving' some fact when we are not; what happens rather is that sometimes we believe we are when we are not," mad error is not a second-order belief about a perception, but a true perception of agreement between ideas that results in something that is not, actually, a fact. It is not that madmen think they are certain when they are not, it is that they are certain when they should not be. Madmen do not take some proposition to be true on the grounds of a faulty assessment of probability, but rather because of a perception of a relationship between ideas. This is not a false perception, but it produces a false proposition. Thus Locke writes that the afflicted "firmly embrace falsehood for truth; not only because they never thought otherwise, but also because, thus blinded as they have been from the beginning, they never could think otherwise" (Locke 1996, p. 218).

The active powers of the understanding that are compromised by association are not reason or judgment, but those that regulate the construction of complex ideas: our ability to compound,

abstract, compare, and otherwise analyze the connections between our simple ideas. Cemented together in passive perception, associated ideas become "almost one" such that the afflicted "confounds them in his mind;" he "scarce distinguishes them" (II.xxxiii.11, p. 398). This problem at the level of connections of ideas (rather than relations between ideas in propositions) is nowhere present in Locke's account of false beliefs, which he defines as the acceptance of a proposition on the basis of faulty evidence. Unlike the major failures in understanding Locke details in the *Conduct*, namely the want of sagacity, deference to the opinion of others, the overcoming of reason by passion, or the development of undesirable habits, association is not the failure of the active powers of the understanding over which the agent has control, but rather of the passive power of reflection, the inner perception that characterizes intuitive knowledge. This capricious process disturbs the basic materials of thought, and renders the thinker unable to execute his sole epistemological, and moral, responsibility — the assessment of agreement between ideas. Because association confounds even the most self-disciplined intellect and preempts the ability of reason to correct the proclivities of the will, it differs from the other intellectual failings Locke describes in the *Essay* by being more like a somatic shortcoming than a moral one — a failure of the mechanisms of mind, rather than the person at its helm.

3.5 MADNESS AND ENTHUSIASM: LOCKE'S SOLUTION

Madness in the early modern period was a term in flux, but for the most part it was used to describe an extreme degree of melancholic disease characterized by raving and violent behavior. While the boundary between melancholia and madness thus was much more defuse than the contemporary distinction between psychoses and affective disorder, melancholy severe enough

to be called madness was at times distinguished on three grounds that are pertinent to the current discussion. First, towards the end of the 17th century, madness began to be seen as diagnosed through the presence of delusions and hallucinations, which had previously been associated with milder types of melancholy (MacDonald 1981, p. 16). Secondly, mania was increasingly described in more medical, and less moral, terms than melancholia.[54] Finally, while melancholics were often chastised for intellectual and moral weakness, mania was considered beyond the reach of reason. The madman's false utterances and violent and bizarre behaviors were seen as beyond his or her control. Glanvill, for example, wrote of lunatics, "grant to them that they may be serious, believe themselves infinitely, and feel all those Warmths which they pretend" (Sena 1973, p. 300). These indications of madness fall into line with Locke's descriptions of mad error as beyond the reach of the will and ultimately unrectifiable by the powers of the understanding — and phenomenologically indistinct from knowledge.

In drawing the connection between madness and the sort of false "knowledge" that interested him — particularly, the wrong principles of the religious dissident — Locke was not original. The equation of dissident beliefs (especially radical stripes of Puritanism) with psychopathology, particularly mania, was a widespread strategy invoked by Latitudinarians and Broad-Church Anglicans in their attempts to delegitimize and undermine the religious zealots who staked their claims at the fringes of Protestant reform (MacDonald 1981). The aim of the accusation of madness was not to challenge the theology of dissenters, nor to charge them with fraud, but rather to remove them from the set of actors shaping the religious development of the nation after the upheavals of the 17th century by denying that their claims had meaning (Sena

[54] Willis, for example, described melancholy as treatable through rational self-government while madness demanded physical restraint and somatic therapies (Schmidt 2004, p. 584).

1973, p. 298). As in our own day, the mad and the manic of Locke's time often hallucinated revelation, divine sight, or the presence of angels and devils. It was a powerful tool of delegitimization to put Quakers, Barkers, and other sects intent on proselytizing on the basis of their religious visions in the same camp as the truly mad.

In a 1656 letter to his father Locke wrote mockingly that Quakers would do well to go bareheaded (as he observed one doing in protest of courtly obsequy) on the grounds of "the head to hott being dangerous for mad folk."[55] That same year Henry More published his *Enthusiasmus Triumphatus,* which Locke purchased in 1662 as part of a collection of More's philosophical writings (Harrison and Laslett 1965). More argued that enthusiasts (especially Quakers) were not willfully wicked, and certainly not controlled by the devil, but rather suffered a somatic disease. The aim of the volume was purportedly not to give reason for the persecution of enthusiasts as lunatics, but rather to promote tolerance of, and compassion for, their affliction. It is apparent, however, that More had a more cynical and politic aim: the disenchantment of heterodoxy. "For where the naturall causes of things are laid open," he writes, "there that stupid reverence and admiration which surprises the ignorant, will assuredly cease" (More 1656, p. 2). Like Locke, More emphasizes that madness can be completely topical — he gives many examples of such delusions, including ones that appear in Locke's writings, such as the man who believes himself made out of glass. In one instance he describes a man who, after falling into a pool full of minnows, is so sure he has swallowed some that he spends the next decade of his life seeking a cure for the frogs he can hear croaking in his belly. Despite having (in an attempt to learn about his condition) turned to the study of physic, this man, reasonable in all other things, cannot move past his delusion (More 1656, p. 12).

---

[55]Locke to John Locke, Sr., 25 October 1656. MS. Locke c. 24, ff. 169–70.

Locke describes the misbegotten certitude of the enthusiast in terms of the perception of agreement between wrong ideas:

> "Reason is lost upon them, and they are above it: they see the Light infused into their Understandings, and cannot be mistaken; 'tis clear and visible there; like the Light of bright Sunshine, shows itself, and needs no other Proof, but its own Evidence: they feel the Hand of GOD moving them within, and the impulses of the Spirit, and cannot be mistaken in what they feel"

Here Locke is dismissive of the equation of visual and ideational perception that he elsewhere relies on so heavily, accusing enthusiasts of circularity insofar as, "their Perswasions are right, only because they are strong in them. For, when what they say is strip'd of the metaphor of seeing and feeling, this is all it amounts to" (IV.xix.9, p. 700). Viewing the propositions that enthusiasts champion as mad errors makes it clear why they fall short of knowledge, although they are phenomenologically identical with it. It also makes clear why Locke insists that enthusiastic delusions cannot be corrected through reason:

> "*Reason* must be our last Judge and Guide in every Thing. I do not mean, that we must consult Reason, and examine whether a Proposition revealed from God can be made out by natural Principles, and if it cannot, then we must reject it: But consult it we must, and b it examine, rather it be *Revelation* from God or no: And if *reason* finds it to be revealed from God, *Reason* then declares for it, as much as for any other Truth, and makes it one of her Dictates. Every Conceit that throughly warms our Fancies must pass for Inspiration, if there be nothing but the Strength of our Perswasions, whereby to judge of our Perswasions: If *Reason* must not examine their Truth by something extrinsically to the Perswasions

themselves; Inspirations and Delusions, Truth and Falsehood will have the same Measure, and will not be possible to be distinguished." (IV.xix.14, p. 704)

Locke recognizes that sometimes persuasion can ape knowledge, insofar as reason really does "perceive" the agreement between ideas. What he insists on is thus not that the mad error of the enthusiast itself be scrutinized, but rather the additional inference that the proposition in question is divinely inspired be justified. While the associated ideas at the heart of enthusiasm cannot themselves be intervened upon, Locke emphasizes that it is within the power of the afflicted — and of those moved to believe in their claims — to assess the additional step that constitutes revelation. "He therefore that will not give himself up to all the Extravagancies of Delusion and Error," he writes, "must bring this Guide of his *Light within* to the Tryal" (ibid.). Associations abide, but deference to them as the word of God can be intervened upon through reason.

More's and others' treatment of dissidents as not simply wrong but mad meant that enthusiastic claims were not appropriate topics for philosophical or theological debate, and it implied that the enthusiast's proper place was not at the head of an impressionable congregation but within the grim isolation of the madhouse. This strategy generated a dilemma: equating enthusiasm with irrationality risked making all religious claims vulnerable to rationalist critique. In Locke's additions to the fourth edition of the Essay we find a deft resolution to this problem. False principles incorporating associated ideas cannot be identified from within, since they are perceived in a manner that is experientially identical to intuition. Locke recognizes the implications for religious disputes, noting in his first *Letter Concerning Toleration* that when two churches are in disagreement, each may look equally mad to the other, even if one is based on reasoned principles and the other on false ones, for "every church is orthodox to itself; to others,

erroneous or heretical" (1824, p. 19). Nonetheless, Locke gives an account of mad error that makes it objectively different from justified belief, insofar as it cannot be managed by the active understanding.

Thus when the theological context is brought into view — and we must remember that settling matters of religious difference purportedly motivated the writing of the *Essay* itself — it is clear why Locke felt the need to draw a bright line between association and other forms of faulty connections between ideas. To protect his account from the related charges of skepticism and relativism, Locke needed to show that while enthusiasts really do have mad ideas, their sort of delusion is not an extreme of everyday epistemological error, nor a matter of relative judgment, but a distinct sort of mad error. Thus, the addition of Chapter 33 helped shelter Locke's ethics of belief from the charge of irrationalism and to promote his political agenda. This may explain why, after his return to the question of toleration towards the end of his life, Locke renewed his interest in his earlier discussion of madness rather than simply expanding on his discussion of wrong assent in Book IV.

On the other hand, the conflation of enthusiasm with madness in particular brought with it the risk that religious dissidents, like madmen, would not be held responsible for their actions, and not seen as agents capable of sinning against the Church or against God. But while Locke does not hold enthusiasts responsible for their false principles, he does accuse them of committing an epistemico-moral failure by not questioning whether their delusions have a divine origin. The only religious beliefs that can be insisted on, Locke maintains, are those that are either clearly established through natural reason or certain to have been delivered through revelation (i.e., of Biblical origin). As Locke notes, a true prophet should have no trouble using his active faculties to meet this bar, since "God when he makes the Prophet does not unmake the

Man" (IV.xix.14, p. 704). One can imagine cases in which the individual's assessment of the origin of his principle was itself an association — a madman imagining himself Moses, complete with a hallucinated burning bush. But Locke seems confident that most enthusiasts, despite the mad core of their "internal light," are able to meaningfully ask themselves whether "This seeing is it the perception of the Truth of the Proposition, or of this, that it is a Revelation from God" (IV.xix.10, pp. 700–701). While association may lead to a false affirmation of the former, even the madman has his faculties of reason sufficiently intact to assess the latter.

By viewing association as a medical problem, distinct in kind from wrong assent, we can temper Passmore's claim that, "In short, the existence of the enthusiast constantly undermines Locke's hopeful view of man." Enthusiasm does indeed show the danger associations can play in the civic life of men, something Locke had long previously acknowledged about madmen. But by distinguishing mad errors from the grounds of their justification Locke is able both recognize that some religious views really are a form of madness, and hold those who try to enforce those views on others culpable for their actions. Furthermore, the notion that what enthusiasts are in fact responsible for is not the *content* of their principles but rather their treatment of them as revelation buttresses Locke's arguments for toleration in the *Letters*, in which his concern is really about the imposition of ecclesiastical laws by those claiming to know God's will. Locke's description of enthusiasm in the terms of mad error thus supports his wider contention that religious heterodoxy is not in itself punishable by either the state or religious authorities. Nonetheless, without evidence for the divine provenance of purported revelation, he maintains that there is little reason to tolerate what may well be the "castles in the air" of a zealot. Accordingly, anyone who forces his religious views on others should be penalized by the civil magistrate. Locke's theory of madness thus allows him to promote religious toleration without

authorizing nonconformist views, and condemn the authority of false prophets without undermining his program of toleration.

Locke's views on enthusiasm in turn raise the question of whether there are parallel measures that can be taken against associated ideas in general — that is, whether the chains of ideas of the typical madmen also contain additional, justificatory premises that can be scrutinized even if the associations themselves cannot. To address this in the following chapter I investigate to what extent Locke has a positive theory of the sober and rational understanding, that can serve as a basis for explaining how associated ideas impact his ethics of belief. I argue that Locke's definition of personhood proves relevant here, insofar as it shows what is necessary for an individual to be a moral agent. Moral agency is, for Locke, cashed out in terms of the ability to do intellectual labor basis of past experience, rendering the capacity to annex experience to the consciousness the central criteria for mental health. While Locke defines personhood in terms of rationality, personal identity can be disturbed by madness, and thus can be devastating to the (mad)man even as it leaves the person intact.

Within this framework we can view association not only as a risk to rationality, but also to the ability of the individual to be a free agent, to operate in a republic, to approach religion from a place of reason and judgment, and to take on moral responsibility. The madman is not a *subject* who can be said to "transgress against his own light" (IV.xvii.24, p. 688) but rather is an *object* of a failure of passive perception. The very possibility of this perversion of consciousness makes medical psychology central to Locke's concerns. To be free, he believes, is to be self-governing — and to be self-governing is to be guided by the active powers of the understanding, rather than the animal or sensitive powers. Perhaps Locke's profoundest contribution to psychology was to make the development of the active powers of perception the new nexus for intervention on the

individual, insofar as he "in effect transferred the clear identity from the ego to the separate ideas, the simple impressions" (Tuveson 1960, p. 29).

4 ASSOCIATION AND THE SELF

4.1 PERSONAL IDENTITY AND THE ASSOCIATION OF IDEAS

Locke wrote to Molyneux in 1692, inviting criticisms of the *Essay* and asking his friend what additions he would like to see in the second edition. Molyneux complied generously with Locke's request, sending everything from line edits to the repeated but unheeded suggestion for additional support of Locke's claim that a demonstrative moral theory was possible.[56] Among the significant additions Locke made in response to his ongoing conversations with Molyneux was the Chapter 27 of Book II of the new edition, entitled "Of Identity and Diversity." A consideration of Locke's associationism suggests a new answer to the vexed question of whether Locke's influential theory of personal identity is naïve, misguided, or just misunderstood.

Debates over personal identity in the early modern period focused on the nature of the substance that receives divine judgment. Ambiguity about what, exactly, the Bible describes the resurrection *of* during final judgment gave rise to theological conundrums: Could the immaterial soul be complete without the body, or did it need the body to carry responsibility for its earthly actions? What happens to the conscious self between death and Judgment? What makes the sinner in life *the same* as the sinner judged after death? (Bynum 1990). Given the centrality of eschatology in 17th-century philosophy, it was necessary that Locke offer answers to these

[56] Molyneux begged off, perhaps regrettably, from Locke's plea for suggestions about which repetitious passages to cut.

questions compatible with his "way of ideas." Locke's solution was to champion a new object for divine judgment distinct from the body and the spirit: the self.

Locke's starting point is to raise the contentious specter of mortalism, the view that the soul can perish temporarily with the body, an extension of his insistence (*pace* Descartes) that the mind is not always thinking. In his *Reasonableness of Christianity* (1695) — published anonymously due to its provocative claims — Locke makes little mention of the soul but describes how consciousness is annihilated in death and only raised up again by Christ on the day of judgment. Throughout his discussion of personal identity, Locke emphasizes that immaterial substance is insufficient to explain the persistence of selfhood: "'Tis evident the *personal Identity* would equally be determined by the consciousness, whether that consciousness were annexed to some individual immaterial Substance or no" (II.xxvii.23, p. 344).[57]

Locke also dismisses the importance of sameness of body during resurrection. In his response to Stillingfleet he writes,

> "The resurrection of the dead, I acknowledge to be an article of the Christian faith; but that the resurrection of the same body, in your lordship's sense of the same body, is an article of the Christian faith, is what, I confess, I do not yet know" (1824, vol. 2, p. 357).

Indeed, he points out that readers of Corinthians "will see that [St. Paul] plainly distinguishes between the dead that shall be raised, and the bodies of the dead" (ibid., p. 372) — and argues at length that this mode is consistently employed throughout the New Testament. That Locke was self-consciously attempting to alter the terms of the eschatological debates during the years he

---

[57] While Locke emphasizes and reemphasizes the fallacies that result from viewing the soul as the condition for personal identity, as this quote suggests he is agreeable to the suggestion that consciousness is annexed to an immaterial substance (II.xxviii.24, p. 345).

was thinking about the association of ideas can be seen by his own careful shift in vocabulary from the fourth edition of the *Essay*'s discussion of resurrection from "the body" to "the dead" (Thomson 2008, p. 56).

Instead of the body or the spirit, Locke believes what will face Christ is the *person*, identified via that set of ideas corresponding to actions produced by the active powers of the understanding that are united by the same consciousness. Locke's account of personal identity is summarized in Section 2 below. In it Locke emphasizes that "person" is a "forensick" term; his concern is with establishing what actions an agent is responsible for. Every Christian is an individual consciousness, and the experiences appropriated by that consciousness form a unified self that can stand before divine judgment and be righteously judged. Personhood is constituted, Locke believes, by those ideas that are annexed to the self in the proper way, as "the memories of a past consciousness" (II.xxvii.23, pp. 344–345). These memories delineate the self, and thus restrict the actions for which it can be held accountable. Despite frequent misinterpretations, the consciousness that signifies that ideas *are* memories is the identity condition for personal identity, rather than memories themselves. This radical interpretation of personal identity as "the sameness of a rational being" (II.xxvii.9, p. 335), well divorced from the more intuitive definitions of his day based around either the body or the soul, leads Locke to bite several bullets.

Of particular relevance here is Locke's insistence that if the agent is not conscious of having committed a past action, they are not responsible for it, even if they were observed to do it by others. In other words, even if the human organism committed an action, the perceiving self inside of it need not be guilty. Even more troubling to some critics, if I am conscious of having committed an action, I *am* culpable for it, even if, in the view of all outsiders, it was committed

by someone else. The worry is that divine judgment is in some way arbitrary if it is based on what individuals *remember*, as opposed to what they *did*, since memory can be faulty.[58] Locke's account suggests that not all bad actions are sinful, but only those that, contingently, can be recalled. Besides theological objections — the view seems to make God's judgments somewhat haphazard, reliant not on His law on his creations' mental states— the account seems to go against our intuitions. Bernard Williams, for example, urges us to imagine our body being tortured. Even if we are assured that in advance of the torture we would be stripped of all our memories, we expect to still be *ourselves* in a way that is meaningful (Williams 1970).

In what follows I argue that Locke's account of association can make these bizarre ramifications of his theory more comprehensible by shedding light on what Locke means when he says we just *are* the consciousness of those ideas we have appropriated. Appropriation, I maintain, is the equivalent of those *acts* of the understanding that are association's contrast class. Besides contributing to our reading of Locke on personal identity, then, the discussion will serve to underscore the centrality of the active understanding to Locke's theory of human nature.

First I clarify Locke's motivations for his account in Section 4.2, which situates Locke's chapter in the debates over mortalism of his day, and over the puzzles for moral responsibility that result from weaknesses of the flesh. Building on arguments made in Chapter 2, I argue that Locke's rejection of the physiological as the appropriate realm of investigation and intervention by the physician leads him to translate the moral-medical discourse around personal identity into ideational terms. Rather than the fleeting animal spirits that course through the brain, the

---

[58] If Locke's account depends on memory being veridical, it falls into another trap, that of the circularity objection most famously proffered by Reid. Veridicality of memory must rely on the person being the same in some metaphysical sense as an earlier person, and this threatens to undermine Locke's psychological account.

appropriate target for intervention is the ideas that, Locke suggests in places, they constitute. One aim of Locke's theory of personal identity is to show that the substance of which consciousness is a mode is irrelevant to its management by the active faculties of the understanding. Important for this case is the question of the nature of madness, and in Section 4.3 I turn to the demonstrative role of mad error in Locke's case for personal identity. Locke uses madness as an exemplar of failures of consciousness. I argue that Locke sees associated ideas as failing to be *self*-conscious, due to their opposition to intellectual labor. He believes continuity of personality to be suspended wherever madness occurs. In Section 4,4 I show how my reading can bolster an appropriationist reading of Locke's account of personal identity, and reflect on its implications for to Locke's broader commitments about human nature and its absence — what has been called the "soft underbelly of reason" (Gaukroger 2002).

4.2 LOCKE'S ACCOUNT OF PERSONAL IDENTITY

Locke's new chapter on identity and diversity considers two sorts of identity relations.[59] The first is synchronic identity, that is, the unique relationship a thing has with itself. Synchronic identity relations are established on the basis of location in place and time, for, Locke concludes, "whatever exists any where at any time, excludes all of the same kind, and is there it self alone." The relation of identity "refers always to something that existed such a time in such a place, which 'twas certain, at that same instant, was the same with itself and no other" (II.xxvii.1, p. 328). The sort of thing that exists in space and time is determined by its real essence or, due to

[59] The difference between Locke's theory of identity in this chapter and that suggested in the first edition are substantial — see Ayers (1993, p. 206).

the mediocrity of human understanding, its nominal essence. Locke's account of diachronic identity is simply an extension of this relation across time — the idea of the thing itself specifies what is essential to it, and sameness of this essential property indicates identity. In investigating what constitutes an identity relation for different sorts of things, Locke emphasizes that the essential properties that must remain constant across time and place vary depending on the sort of thing in question: for example, "in these two cases of a Mass of Matter, and a living Body, *Identity* is not applied to the same thing" (II.xxvii.3, p. 330).

In other words, once the particular criterion for identity for a given type of thing is established, the evaluation of whether identity obtains between an object at time *t* and an object at time *v* is no more difficult than seeing whether the criterion used to establish identity for that class of things remains *the same* across times *t* and *v*. The question is whether the synchronic relation of the essence of the thing with itself at time *t* is the same relation as the diachronic relation of the essential property of a thing at time *t* and the essential property of a thing at time *v*. Since synchronic identity is established via location in place and time, the identity relation means that *all* of the thing's properties will be identical. In the case of diachronic identity, the essential property is identical but other properties need not be. For example, the same horse when a filly and when an old nag will have the same *life*, the identity criterion, for Locke, of being a living creature, but the matter that makes up its body will be different at these different stages.

Locke believes that the delineation of groups of properties into kinds is underdetermined by the ontology of the world— as Guyer has put it, in Locke's view the complexity of the external world leaves us with the "burden of choice" when it comes to classification (2011, p. 118). Accordingly the sorts of properties identity is constituted by are necessarily based on the

understanding's own categories. While the identity of a mass of atoms can be established through the identity of matter (no addition or subtraction of atoms) if this criterion were applied to a living thing, like a horse, identity relations would not track the understanding's own ideas about nominal essences. Ultimately, all identity relations are for Locke is a way to establish the appropriateness of applying names to things: "for such as is the *Idea* belonging to that Name, such must be the *Identity*: Which if it had been a little more carefully attended to, would possible have prevented a great deal of that Confusion, which often occurs about this Matter, with no small seeming difficulties" (II.xxvii.6, p. 332).

Locke emphasizes this point especially in the case of human beings, writing,

> "'Tis not therefore Unity of Substance that comprehends all sorts of *Identity*, or will determine it in every Case: But to conceive, and judge of it aright, we must consider what *Idea* the Word it is applied to stands for: It being one thing to be the same *Substance*, another the same *Man*, and a third the same *Person*, if *Person, Man, and Substance,* are three Names standing for three different *Ideas.*" (II.xxvii.7, p. 332)

Man, Locke says, is merely a type of animal form, and being the same man only requires having the same life over time. Missing here is the traditional criterion of being a *rational* animal: even a madman has human identity, in Locke's view. He even goes so far as to say that if a parrot could discourse more eloquently than most human beings, rather than calling the bird a man one would exclaim that it were a uniquely brilliant, even rational, parrot.[60] Alternatively, a human who could only parrot back words without sense would still be called a man.

---

[60] The incredible narrative that Locke inserts into his chapter while making this case is worth mention. He retells a story told to him by, he claims, an extremely credible source, who

A person, on the other hand, is not defined in terms of its substance or species but rather in terms of its capacity for certain forms of intellectual activity; it is

> "a thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places; which it does only by that consciousness, which is inseparable from thinking, and as it seems to me essential to it." (II.xxvii.9, p. 335)

Recall that for Locke reason is the ability to compare ideas using (when necessary) intermediate chains of ideas in order to establish relations between them, while reflection is an inner sense that can observe mental phenomena. The ability to consider oneself to be the same across time and space relies on both the former and the latter capacities. Reflection allows for consciousness insofar as "Consciousness is the perception of what passes in a Man's own mind" (II.i.19, p. 115). And reason allows for the comparison of one's current perception of oneself as oneself — one's ipseity, or "the perception I have of my actions 'from the inside'" (Balibar 2013, p. 100) — with previous such perceptions. It is this sort of comparison that allows each of us to assess, *from the inside*, whether past actions were perceived by the same consciousness as currently constitutes our self.

Any being that thinks, according to Locke, has consciousness, since it is "impossible for any one to perceive, without perceiving, that he does perceive" (II.xxvii.9, p. 335). Consciousness is, then, itself a form of perception, but it does not, despite the previous quote, seem for Locke to be a higher-order perception of one's perceptions, like reflection. It is not an

---

encountered a parrot capable of all manner of rational discourse. While the bemused Locke seems to present the story with credulity, it had no obvious impact on his view on animal rationality, which at other points in the *Essay* is revealed to be skeptical. Nonetheless this passage suggests he believes personhood to be possible in particular animals, despite their lack of a human identity.

idea about our ideas, but rather a perception of *ownership* that accompanies our present sensations and reflections:

> "When we see, hear, smell, taste, feel, meditate, or will any thing, we know that we do so. Thus it is always as to our present Sensations and Perceptions: And by this every one is to himself, that which he calls *self*: It not being considered in this case, whether the same *self* be continued in the same, or divers substances. For since consciousness always accompanies thinking, and 'tis that, that makes every one to be, what he calls *self*; and thereby distinguishes himself from all other thinking things, in this alone consists *personal Identity*, *I.e.* the sameness of rational Being: and as far as this consciousness can be extended backwards to any past Action or Thought, so far reaches the Identity of that *Person.*" (II.xxvii.9, p. 335)

Consciousness is also, Locke says here, all that is needed to constitute the identity of a person — much in the way that the same hooves and mane are not needed to make the filly the same horse as the nag, the same powers of reason and reflection are not needed for to make the person the same over time (Weinberg 2011, p. 400). All that is needed is the same consciousness. And in what does this sameness consist? Rather than a form of higher-order thought, consciousness is property of our ideas that tags them as ours — a feeling of "mineness." It follows that, as Margaret Atherton has put it, "the effect of having experiences or registering experiences via consciousness will be to build up a distinctive outlook" (2008, p. 298). In discussing remembrance, Locke describes how "the Mind many times recovers the memory of a past consciousness" (II.xxvii.23, pp. 344–345). What makes this "past consciousness" a *memory* is that the self's *own* consciousness is perceived as annexed to those ideas about the past which

are the object of the reminiscence. Personal identity is only established through the recognition of past ideas as *my memories* insofar as memories are ideas to which my current consciousness is annexed.

So it is the continuity of the same consciousness, my consciousness, rather than the chain of memories, that is the identity condition for persons.[61] It follows that when Locke says personal identity consists in "the sameness of rational Being," this signifies the *felt* sameness of being — assessed from the inside. Since no one can know another's thoughts (he can only guess them through the words which the other chooses to represent his ideas), personal identity can only be established by the individual *for him or herself*. Consciousness is established phenomenologically through the dual faculties of reason and reflection: "If we take wholly away all Consciousness of our Actions and Sensations, especially of Pleasure and Pain, and the concernment that accompanies it, it will be hard to know wherein to place personal Identity" (II.i.12, p. 110).

Since judgments about the identity of persons are always self-reflexive, and Locke acknowledges how deeply this limits the utility of the concept for establishing moral responsibility in ways that will be discussed below. For now let us note that personal identity differs from animal identity because while an outside observer might establish whether the same life or organization continued uninterrupted in a man, identity in the case of persons can only be established through reflection and is therefore by definition only calibrated subjectively. When establishing that a woman is the same woman I knew as a child, I must establish that her

---

[61] This is important because for Locke, a string of particulars cannot constitute identity — all things whose "Existence is in succession, such as are the Actions of finite Beings, *v.g. Motion* and *Thought*, both which consist in a continued train of Succession" are necessarily diverse (II.xxvii.2, p. 329).

organization or life has proceeded uninterrupted since the time of our initial acquaintance in grade school. When I establish whether a memory is, in fact, mine, I do *not* attempt to establish a genealogy of any feature of the event in question but rather assess whether the memory of the past action is *currently* accompanied by an idea of consciousness — a feeling of ownership, *mineness*. Locke emphasizes that a person establishes whether it "can consider it self as it self, the same thinking thing in different times and places […] *only* by that consciousness, which is inseparable from thinking" (II.xxvii.9, p. 335, italics mine).

Locke notes that in the ordinary usage man and person are interchangeable, but emphasizes that his aim in differentiating them is to establish a forensic term that isolates the moral responsibility of the agent, even if it seems to conflict with what we know of the man. He offers a provocative thought experiment:

> "Had I the same consciousness, that saw the Ark and *Noah's* Flood, as that I saw an overflowing of the *Thames* last Winter, or as that I write now, I could no more doubt that I, that write this now, that saw the *Thames* overflow'd last Winter, and that view'd the Flood at the general Deluge, was the same *self*, place that *self* in whatever Substance you please, than that I that write this am the same *my self* now whilst I write (whether I consist of all of the same Substance, material or immaterial, or no) that I was Yesterday. For as to this point of being the same *self*, it matters not whether this present *self* be made up of the same or other substances, being as much concern'd, and as justly accountable for any Action was done a thousand Years since, appropriated to me now by this self-consciousness, as I am, for what I did the last moment." (II.xxvii.16, p. 340)

Here Locke insists that if he "had the same consciousness" of seeing the deluge described in Genesis as he did of seeing the overflowing of the Thames the previous year — even if that memory necessitated being thousands of years older than was possible — both would contribute to his personality, and he would be equally as responsible for any actions he remembered taking in Biblical times as those he recalled from the previous year. A few pages later Locke makes clear he is also committed to the inverse, also counter-intuitive implication of his theory:

> "For whatsoever any Substance has thought or done, which I cannot recollect, and by my consciousness make my own Thought and Action, it will no more belong to me, whether a part of me thought or did it, then if it had been thought or done by any other immaterial being any where existing." (II.xxvii.24, p. 345)

Here Locke offers up another thought experiment, the case of "two distinct incommunicable consciousnesses acting the same Body, the one constantly by Day, the other by Night." This is a Jekyll-and-Hyde situation, in which consciousness is split between two different characters, with different memories. Locke asks rhetorically whether "the Day and Night-man" would not be as distinct persons as, say, "Socrates and Plato," despite the fact that the body they inhabit actually performed both the actions undertaken during the day, and those undertaken during the night. He concludes they would be two separate people (II.xxvii.23, p. 344).

Cases like these are a source of ire for critics such as Flew (1951), who follow Reid (1785) in arguing that Locke's criterion has no way to distinguish between what a person *actually* did and what they think they did. Locke would put it as a difference between what the *man* did and what the *person* did; what the bodily or mental substance does may differ from what the person does, because the person is constituted by sameness of consciousness. His critics are

not satisfied by this because it makes Locke's theory of divine judgment somewhat arbitrary — not all actions will be judged, because many are of forgotten over the course of life; and what actions each person will be judged for on the day of judgment cannot be predicted by what they were observed to do. In other words, there is no principled relation between those ideas that are annexed to consciousness and those we are morally responsible for that would seem able justify Locke's psychological interpretation of personhood. In the following section I argue that the theory of associated ideas presented in the previous two chapters suggests such a justificatory relation.

4.3 ASSOCIATION AND APPROPRIATION

As mentioned above Locke emphasizes that personal identity is a forensic term, "appropriating Actions and their Merit" (II.xxvii.26, p. 346). Critics have objected that, given the subjectivity of the calibration of personal identity by consciousness that Locke allows for, his identity condition for persons is ill-suited for forensics (Weinberg 2011). What use is a metric for establishing culpability in a court of law that can only be testified to by the accused himself? Locke acknowledges his theory's shortcomings for earthly justice; his motivations are eschatological.[62] The term "person" is a term of art for moral science, a language used between

[62] "For though punishment be annexed to personality, and personality to consciousness, and the Drunkard perhaps be not conscious of what he did; yet Humane Judicatures justly punish him; because the Fact is proved against him, but want of consciousness cannot be proved for him" (II.xxvii.22, p. 344). Molyneux raises a concern with this aside of Locke's. Surely, Molyneux argues, drunkards should be found guilty because their sin of intoxication caused their crime, and they were conscious when they began to drink. Locke acknowledges the point, but dismisses it as outside of his purview, writing, "But drunkenness has something peculiar in it when it destroys consciousness; and so the instances you bring justifie not the punishing of a drunken fact, that

the individual and God. Locke's aim is to allow people to assess their prospects for the afterlife, and remind them that God will only punish those sinners who *own up to* their crimes:

> "The Sentence shall be justified by the consciousness all Persons shall have, that they *themselves* in what Bodies soever they appear, or whatever Substances soever that consciousness adheres to, are the *same*, that committed those Actions, and deserve that Punishment from them." (II.xxvii.26, p. 347)

Accordingly, his definition of the person need not assure continuity across the lifetime, but only assure fair treatment at the instant of final judgment. A question remains, however, about how well it succeeds at *this* goal — why should God's assessment of our sinfulness be based on what *we* think? What makes consciousness the appropriate arbiter of responsibility?

The self is conscious of all the ideas it perceives, but only some ideas enter into the repository of our memory. These are ideas of which we have what Locke calls *past* consciousness — when we perceive them we know that we have perceived them before. The most obvious reason we could posit for these ideas being in our memories is that they are veridical — that is, that something that is us, some sort of substantial unity, really did the acts we remember doing. Locke is not particularly interested in this sort of metaphysics of memory. Rather, he his concerned with method: *how* were the ideas produced, when they were produced? The extent to which we remember ideas seems to be a product of the *attention* with which we originally perceive them:

---

was totally and irrecoverably forgotten, which the reason that I give being sufficient to do, it well enough removed the objection, without entring into the true foundation of the thing, and shewing how far it was reasonable for humane justice to punish a crime of a drunkard, which he could be suppos'd not conscious of, which would have uselessly engag'd me in a very large discourse, and an impertinent digression" (Locke 1979 Vol. 4, p. 785). He fails, in my opinion, to adequately respond to Molyneux's point that ignorance of past actions can be feigned just as well on the grounds of madness as on intoxication (Locke 1979, p. 58).

> "That there are Ideas, some or other, always present in the mind of a waking Man, every one's Experience convinces him; though the mind employs it self about them with several degrees of Attention. Sometimes the mind fixes it self with so much earnestness on the Contemplation of some Objects, that it turns their Ideas on all sides; remarks their Relations and Circumstances; and views every part so nicely, and with such intention, that it shuts out all other Thoughts, and takes no notice of the ordinary Impressions made then on the Senses, which at another Season would produce very sensible Perceptions: At other times, it barely observes the train of Ideas, that succeed in the Understanding, without directing, and pursuing any of them: And at other times, it lets them pass almost quite unregarded, as faint shadows, that make no Impression" (II.xx.3, p. 228)

Ideas that are attended to are remembered: "*Attention* and *Repetition,*" Locke tells us, "help much to the fixing any *Ideas* in *the Memory*" (II.x.3, p. 150). The connection between attention and active perception is also touched on throughout the *Essay*.[63] Locke describes attention is a key component of intellectual acts, writing, "Thinking, in the propriety of the English Tongue, signifies that sort of operation of the Mind about its Ideas, wherein the Mind is active; where it with some degree of voluntary attention, considers any thing" (II.ix.1, p. 143). In the cases where the formation of complex ideas is closely attended to, the connections between them tend to be natural; in the cases where propositions are formed with attention, they tend to be true. To attend to an idea is to see if it is in need of adjustment or correction by the active

---

[63] "Other Truths require a train of Ideas placed in order, a due comparing of them, and deductions made with attention, before they can be discovered" (I.iv.22, p. 99); "yet he will have but a confused Idea of all the Parts they are made up of, till he applies himself with attention, to consider them each in particular" (II.i.7, p. 107)

powers to bring it in line with what is real and true. And correctness, for Locke, is next to Godliness — to have right ideas is to have ideas of the natural law. Locke's hatred of mind-wandering, inattention, and prejudice are because epistemic acts are *moral* acts — ideas formed with insufficient attention are condemnable.

Ideas that result either from our attentive labor or from our lazy neglect are equally our handiwork, the result of voluntary acts (or acts of omission) of the understanding. Past consciousness is, Locke tells us, "but a present representation of a past action" as ours (II.xxvii.13, p. 337). When that past action is represented as ours it is because it was a voluntary action, the result of intellectual activity. When we represent it, we represent the ideas in a proprietary fashion. Actions that do not follow from mental activity cannot be volitional, and are not registered in our consciousness as done by us. Given the close connection between those ideas of which we feel ownership — that is, that are annexed to consciousness — and our moral culpability, it makes sense that it is our self-made registry of our intellectual virtues and vices that we are judged by. As discussed in Chapter 2, human freedom is, for Locke, the result of the active powers of the understanding. Any acts not directed by those powers are not free, and therefore not subject to judgment.

Support for this interpretation comes from the fact that for Locke, mad errors disrupt personal identity:

> "But if it be possible for the same Man to have distinct incommunicable consciousness at different times, it is past doubt the same Man would at different times make different Persons; which, we see, is the Sense of Mankind in the solemnest Declaration of their Oppinions, Humane Laws not punishing the *Mad Man* for the *Sober Man*'s Actions, nor the *Sober Man* for what the *Mad Man* did,

thereby making them two persons; which is somewhat explained by our way of speaking in *English*, when we say such an one *is not himself*, or is *besides himself*." (II.xxvii.20, pp. 342–343)

Why, given the associative theory of madness presented in the previous chapter, would Locke believe madmen to "at different times make different Persons"? Recall that association is a unique mental malady that differs from other weaknesses of the understanding in being a pathology of ideas, rather than reason or volition. When the madman applies his inner sense of reflection to his ideas, he is unable to evaluate the connections between them. Complex ideas appear to be simple ones because his active powers of understanding, while itself functioning normally, cannot disentangle the fantastical constructions of the diseased faculties of sense. If he had correctly assessed — that is, been conscious of — the simple ideas that got compounded, he could undo the association and replace his unnatural ideas with natural ones. But because the active powers of the understanding were excluded from the formation of the complex he sees a simple idea, rather than a compound one, and is helpless.

Because associated ideas are phenomenologically indistinguishable from necessarily connected ideas, they are treated by the understanding as "true" in the colloquial sense that Locke recognizes: they are truth-functional when used in propositions. Accordingly, the madman has beliefs and (what he takes to be) propositional knowledge that were not produced through the active understanding, that is, were not voluntary. Insofar as he passively perceives without actively perceiving, the madman is not conscious in the sense Locke uses it in his definition of personhood, akin to *self*-consciousness, which is "inseparable from thinking, and as it seems to me essential to it: It being impossible for anyone to perceive, without perceiving, that he does perceive" (II.xxvii.9, p. 335). Thus mad ideas and the mad errors they contribute to are not

candidates for past consciousness of the type described above, in which the perceiver can recognize that he has perceived the chain of ideas before, because the mad idea was never originally perceived in an active, generative fashion.

Locke draws a close relation between the past consciousness of ideas and the ownership of actions, using "past selves" and "past actions" somewhat interchangeably to describe what it is that is united across time by sameness of consciousness to constitute the person. Given Locke's account of voluntary action as that which is determined by our desires and beliefs, it makes sense that he slips between the ideas causing the action and the action itself. We might say that an action becomes ours by being the result of our free volition, that is, the result of the active application of our powers of understanding to our ideas. It is this sense in which personal identity is a forensic term: "person" is used by the self to refer to its acts of "appropriating actions and their merit," that is, for its claiming as its own actions which result from acts of will (II.xxvii.26, p. 346).

Scholars sympathetic to this sort of interpretation of personal identity as the subjective identification of certain actions as ours have latched onto this language of appropriation to describe their view (Behan 1979; Winkler 1991; LoLordo 2012). According to appropriationist readings, the person just *is* the agent responsible for those actions that a single consciousness *appropriates*, or claims as its own; "Locke is interested in what we *take* the self to include, or in the self that comes about as a result of such takings" (Winkler 1991, p. 153). Those supporting this view rely on quotes like the following:

> "This personality extends it *self* beyond present Existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to it self past Actions, just upon the same ground, and for the same reason, that it

> does the present […] And therefore whatever past Actions it cannot reconcile or *appropriate* to that present *self* by consciousness, it can be no more concerned in, than if they had never been done." (Locke, *Essay,* II.xxvii.26)[64]

For the most part scholars sympathetic to the appropriationist reading have not written much about the nature of the appropriative act, despite that elsewhere in Locke's writings, "to appropriate" — from the Latin *appropriāre*, from *ad-proprius*, to render one own, to deliver to the self — plays a crucial role. More than simply claiming something as one's own, Locke uses the term to indicate the *rightful* transfer of an object from the public sphere to oneself, or maybe better, the transformation of an object from something outside of oneself to a part of oneself (Olivecrona 1974, 222). It is at the core of his theory of the origins of property in the *Second Treatise of Government* (1689); here the person *appropriates* objects through labor, through acting upon what is commonly held to make it one's own. By cutting down a tree with the strength of his body, the farmer appropriates the lumber for himself, by mixing his labor with it:

> "Thus the Grass my Horse has bit; the Turfs my Servant has cut; and the Ore I have digg'd in any place where I have a right to them in common with others, become my *Property*, without the assignation or consent of anybody. The labour that was mine, removing them out of that common state they were in, hath *fixed* my *Property* in them." (Locke 1960, p. 289)

Michael Ayers is suspicious of any meaningful connection between Locke's discussion of appropriation in the context of personal identity and his later employment of it in the political

---

[64] Locke uses the term once more in Chapter 27, describing the man who remembers Noah's flood as "being as much concern'd, and as justly accountable for any Action was done a thousand Years since, appropriated to me now by this self-consciousness, as I am, for what I did the last moment" (II.xxvii.16, p. 341).

setting. The concept of ownership in each setting is, he argues, quite different. While "external property is united to us by a merely legal or moral relation […] [o]ur actions themselves […] are 'appropriated' to us by an entirely natural and given principle of unity, namely consciousness, rather than by some acquisitive act of acknowledgment or 'owning' on our part" (Ayers 1993, p. 268). Etienne Balibar takes much more seriously Locke's conception of the person as an individual "proprietor of himself," whose intellectual labor is analogous to manual labor in being primarily an expression of human freedom (2013, p. 72). Following from the close 17th-century connection between consciousness and conscience (in French the same word, *conscience*, serves both purposes, and the English variation only emerged around Locke's time), Balibar argues that Locke constructs "a parallelism of responsibility and property, of self-consciousness and 'property in oneself'" (ibid., p. 101).[65] What is appropriated by consciousness to the self is actions, but of course, for Locke this means *ideas* of actions. Healthy complex ideas, Locke tells us, are connected through the workmanship of man, and he defines thinking as the *activity* (rather than the *essence*) of the soul. We act on the basis of our ideas, which drive our beliefs and desires about the world — that is, we act on the basis of reasons. What allows us to experience past actions as our own is that they were the result of active thought, that is, were free actions, determined by our volitions.

---

[65] This move can be traced to Pufendorf and Grotius, both influential on Locke's theory of property. Olivecrona writes, "According to Grotius, means of subsistence could be appropriated in the state of nature without any preceding compact. What a man had collected became his own. It was included in the *suum*. Therefore it constituted an injury to rob him of it. The concept of injury had its usual significance in this connection. To take some fruits from him who had collected them was an injury in the same sense as giving him a blow or damaging his reputation. It was an attack on his personality. The underlying idea was that the personality of the collector had been extended so as to encompass the fruits" (Olivecrona 1974, p. 223). Actions were analogously appropriated to the person; for Grotius and Pufendorf, Olivecrona writes, "my actions are 'my own' because they are directed by the ego" (p. 224).

In contrast, we can see why association — the *passive* construction of mental products — produces ideas that are not the property of the moral agent. It is true associated ideas enter the understanding and becomes the stuff of thought. In fact, we have seen that madmen can reason correctly about their mad ideas, like the man who thinks his body is made out of glass and is careful with it. We may say that the madman is conscious of the idea of his body being made out of glass, but not *self*-conscious of it, not having generated it. He did not make the idea, did not store a creative act in the repository of his memory, and therefore is unable to correct it through his labor. When he wishes to return to the original action in which the simple ideas were connected in order to appraise the connection — an ability that undergirds Locke's moral epistemology — he cannot do it, because he never committed the act.

Thus we may see why, though the madman has ideas he is conscious of, when he is sane he does not stand in a relation of ownership towards those ideas. They are not of his person. To punish the sane man for what he did when mad is "to punish one Twin for what his Brother-Twin did, whereof he knew nothing, because their outsides were so like, that they could not be distinguished" (II.xxvii.19, p. 342). The madman is not subject to the natural law — unable to discover it through reason, and spared from being punished when he breaks it on account of his condition. In this respect the madman serves as a foil for the unified person, and helps us see that for Locke personhood is a product of the labor of active faculties of mind, which either act in accordance with the moral law or allow for the "shameful neglect of what is in their power" (II.xxi.70, p. 281). To be a moral agent capable of receiving divine deserts, the individual must, when suspension allows for it, act on her ideas. Her successes and failures are all she takes with her on Judgment Day, and her consciousness keeps the logbook that she lays at the Gates for

divine review. This interpretation makes sense of Locke's obsession with the management of the understanding, since to be rational is to gain access to the next life:

> "But when infinite Happiness is put in one Scale, against infinite Misery in the other; if the worst, that comes to the pious Man, if he mistakes, be the best that the wicked can attain to, if he be in the right, Who can without madness run the venture? […] Must it not be a most manifest wrong Judgment, that does not presently see, to which side, in this case, the preference is to be given?" (II.xxi.69, p. 282)

4.4 THE PERSON AS A MEDICAL OBJECT

Locke's description of the madman above as having two "distinct incommunicable consciousness" at different times, causing him to lose, for forensic purposes, his personal identity, is analogous to his day-man/night-man example. However unlike that case — and the many other whimsical ones Locke provides, such as the prince and the cobbler, Socrates sleeping and Socrates awake, and so on — madness is *not* a thought experiment. It is rather, Locke claims, a universal affliction: "there is scarce a Man so free from it, but that if should always on all occasions argue or do as in some cases he constantly does, would not be thought fitter for *Bedlam*, than Civil Conversation" (II.xxxiii.4, p. 395).

The example of the madman serves to show that Locke's distinction between the human and the person is not purely academic. Rather it gets at the complex relationship between the powers of the understanding, human freedom, and divine restitution. In cases of complete mental wellbeing, in which reason and reflection allow for perfect memory, the human being and the

person can collapse into each other. But while in the moment the madman can exhibit reasonable thought, he is not capable of the developmental product of consciousness that is the self: he simply cannot perform the right kind of intellectual labor on his ideas to make them part of himself. Thus he is subject to precisely the sort of paramnesia Locke's critics worried about, in which what he thinks are veridical beliefs are really just fantasies. Only when mental disease has led to the destruction of the smooth advancement of consciousness through the appropriation of ideas does it become clear that what supports moral responsibility is not the immaterial soul nor its unicity with the body, but the correct management of the understanding.

Reading Locke's account of personal identity this way is to emphasize the normative nature of his epistemology, which takes medicine as the regulative ideal and extends its reach to the diagnosis and treatment of those mental failings that become moral. Once association has set in, its consequences are grave — madness has repercussions not just for this life, but for the next one. But failures of judgment are even more problematic, insofar as they are sins against God. Locke argues that once a human being knows what the good is — that is, what can ease his discomfort — he will do it. However, Locke argues,

> "Though his will be always determined by that, which is judg'd good by his Understanding, yet it excuses him not: Because, by a too hasty choice of his own making, he has imposed on himself wrong measures of good and evil; which however false and fallacious, have the same influence on all his future conduct, as if they were true and right." (II.xxi.56, p. 271)

Sutton refers to "Locke's sad realism about the cognitive effects of the fall," (Sutton 1998, p. 249); in line with his contemporaries mentioned above in the *cultura animi* tradition, Locke clearly sees the human condition as a result of original sin. But he also emphasizes that

God has gifted humans with the understanding best suited to their circumstances, and further that through the use of this divine gift we can even alter our circumstances to become more angelic (I.i.5, p. 45, Locke 1996, p. 215). "If I may guess at Things unknown," Locke muses in an unusually speculative moment,

> "I am apt to think, that Angels have now, and the Spirits of just Men made perfect, shall have, in a future State, of Thousands of Things, which now, either wholly escape our Apprehensions, or which, our short-sighted Reason having got some faint Glimpse of, we, in the Dark, grope after" (IV.xvii.14, p. 683).

In this respect the *Essay* is really a study of human foibles, and the *Conduct* a wistful instruction manual for comporting oneself like a higher being.

There are, however, limits to the improvement of the self, and these limits are precisely what defines human beings, and offer the key to their nominal essence that functions as Locke's identity condition. After mentioning the outstanding memory of Pascal, who exhibited perfect recall, Locke notes,

> "For this of Mr. *Pascal* was still with the narrowness, that humane Minds are confin'd to here, of having great variety of *Ideas* only by succession, not all at once: Whereas the several degrees of Angels may probably have larger views, and some of them be endowed with capacities able to retain together, and constantly set before them, as in one Picture, all their past knowledge at once." (II.xi.9, p. 154)

At the foundation of Locke's definition of persons as rational beings, then, lies the chain of connected ideas. In the nature of these connections — natural or unnatural — lies the fate of the human being.

Personhood brings with it the responsibility to attempt to improve these connections; Locke's Protestant commitments show here and he is in step with his time, in which it became common to equate moral goodness with the maintenance of "'character' in the laudatory sense" (Leites 1988, p. 120), especially the maintenance and strengthening of the understanding. Original in Locke's account is the recognition that sometimes we are not even granted the freedom to try to improve our lot. To neglect the understanding is to offend against God — "*we can hinder both Knowledge and Assent, by stopping our Enquiry,* and not imploying our Faculties in the search for any Truth," Locke writes. "If it were not so, Ignorance, Error, or Infidelity could not in any Case be a Fault" (IV.xx.16, p. 717). But it is also to be granted the potential to receive divine forgiveness, that is, grace — the sinner always has that capacity. The exclusion of the madman from rational agency on the basis of his disrupted personhood shows that this personal freedom cannot be assured. In the distinction between association and other forms of mental impairment can be found the criterion for moral responsibility.

My discussion of association as a powerful contrast class to healthy relations between ideas serves those scholarly accounts that differentiate Locke from archetypical associationists like Hume insofar as his ideational treatment of the self is *not* a mechanistic one: as Tuveson, for example, writes,

> "Experience for Locke is not a mere automatic connection of impressions, as if an adding machine were being set up. He always has the sense of a living being, with inclinations of its own, responding in a myriad of ways to a world which affects it in as many ways." (1960, p. 40)

What associated ideas stand in contrast with is precisely the adaptive nature of human understanding, which, through the powers of active perception that constitute its freedom, grows

and conforms in relation to its environment. Locke's commitment to the developmental nature of the understanding is most obvious in his normative tracts on human understanding, *The Conduct of the Understanding* and *Some Thoughts Concerning Education.*

Childless himself, Locke approached the topic of the upbringing of children in a concerted way in the 1680s in response to the pleas of his friend, Edward Clarke, who was eager for counsel about the education of his sons. Locke ultimately published what would become a wildly popular compendium of child-rearing advice based on the letters he sent to Clarke, first in 1693 and then in two additional printings during his lifetime in 1695 and 1699. Both *Some Thoughts* and the *Conduct* engage head-on the problem raised by Locke's developmental account of rationality: in light of the contingency of all chains of ideas, how do we delineate the healthy from the pathological ideas, and protect our children from the latter? It is within this developmental context that Locke first writes about association, though, as described in the previous chapter, he attributes his discussion to his reflections on madness and the connection of ideas decades earlier. No doubt Locke's interest in pedagogy served his political agenda. As Casson has written, "Instead of assuming the presence of preconstituted, rational individuals, Locke seeks to shape his readers into the type of people who will be able to sustain stable and just institutions" (2011, p. 261).

As we have seen the threat of madness was first recorded by Locke in his 1677 journal, in which he writes, "[O]ne may see of what moment it is to take care that the first impressions we setle upon our minds be comformable to the truth and the nature of things, or else our meditations and discourse there upon will be noe thing but perfect raving" (Dewhurst 1963, p. 89). After revising *Some Thoughts Concerning Education*, which is littered with references to the power of habit and the importance of extraordinary care in the education of young people, Locke

formalized this early anxiety about the imminent threat of madness under the heading of "associated ideas" in his new chapter for the *Essay*. In his mature formulation, the boundary of rationality is policed by the understanding, which must parse ideas as they enter the mind through perception and make sure they are not unnaturally conjoined. In childhood, the understanding must be trained to perform this duty, but if it does not gain the appropriate vigilance, the faculties of reason and judgment become ineffectual mental appendages.

The importance of his concomitant reflections on child-raising and the conduct of the understanding for the present topic is clear from Locke's insistence, in both *Some Thoughts* and the *Essay*, that in the education of children "there is not any one thing that deserves more to be looked after" than the association of ideas (II.xxxiii.9, p. 397). In the *Conduct* Locke gives a normative rather than historical treatment of association in which he emphasizes that one can only hope for a successful defense against bad habits and chance associations through inoculation — a careful attention to the ideas that get unnaturally connected during education. Children, he writes, must be carefully observed so that "they never suffer any ideas to be joined in their understandings in any other or stronger combination than what their own nature and correspondence give them." In other words, children may not be able to themselves recognize when a relation between two ideas has become habitual, and thus are particularly at risk of forming associations. Because of children's natural submission to authority, the careless words of adults — fables about ghosts, sprites, or devils lurking in the dark, for example — can leave lifelong associations. Learning to interrogate their own ideas not only saves children from the mania of association but also teaches them to do good science, to engage the world around them through induction and experiment.

In his discussion of learning Locke mentions the importance of careful exercise and the slow building of strength of the understanding, to avoid putting minds "to a stress beyond their strength" which can undermine the performance of the rational faculties and destroy their vigor. Locke describes a battery of personal failings that can get in the way of the right exercise of reason, including laziness and fear — he applies the proverb, "Use legs and have legs" (Locke 1996, p. 215) to suggest that the failure to be reasonable is due to a failure to develop the mental faculties, the way one would a muscle. He urges self-discipline, reminding his readers that failing to activate their reason is, literally, to give up their freedom:

> "Men know the value of their corporeal liberty, and therefore suffer not willingly fetters and chains to be put on them. To have the mind captivated is, for the time, certainly the greater evil of the two, and deserves our utmost care and endeavors to preserve the freedom of our better part. And in this case our pains will not be lost; striving and struggling will prevail, if we constantly, in all occasions, make use of it."

Occupation for the mind, Locke continues in fine Protestant fettle, is the best possible medicine, and "it may not be amiss […] to make use of so profitable a remedy that is always on hand" (ibid., p. 227). As seen above, active perception takes effort, and while associations are immune to its curative effects the assessment of the connections between ideas the mind *can* recognize as in disagreement will hone the understanding.

Much of the *Essay*, and nearly the whole of the *Conduct*, is dedicated to advising readers on how best to govern their understanding in order to maintain their rationality without being swayed by passions or confused by language or bad habits. "I am only speaking," Locke writes, "of what they should do who would deal fairly with their own minds, and make right use of their

faculties in the pursuit of truth; we fail them a great deal more than they fail us" (ibid., pp. 211–212). Locke condemns men who would blame their sins on nature, since, he argues, moral failings are due to a mismanagement of assent, "the great principle and foundation of all virtue and worth" (ibid., p. 25). On the other hand, just men who utilize their reason to praise God and honor their covenant with Him can expect, in the next life, to be rewarded with a higher faculty of understanding that will render the clumsy tool of reasoning unnecessary: the complete certainty of the angels (IV.xvii.14, p. 683).

What Locke achieves in such passages is a new conception of the self as the sum of those ideas annexed to the same consciousness. The chains of ideas that are codified by habit into personal principles, factual knowledge, and probabilistic belief come to define the person. The influence of this view on the way personal development was conceived following Locke is suggested by shifts in theories of child-rearing away from the breaking of the child's innate depravity through discipline towards a more sympathetic education in the right habits of mind (Spellman 1988, p. 212). Association is a prime example of Locke's novel focus on the ideas themselves over the faculties that produce and moderate them.

Patrick Romanell points out that Locke substitutes the Greek "Semiotike" for "logic" in his taxonomy of knowledge, and notes that in the early modern period "semiotics" had a medical connotation, referring, within Galenic medicine, to symptomatology. "For the science of knowing diseases in particular," Romanell argues, "became in Locke the physician-philosopher the science of knowing in general" (1984, 148). In providing a natural history of the understanding, Locke pursues a nosology of the cases in which human understanding succeeds and fails, including its most common pathologies. Locke's treatment of various mental pathologies through the *Conduct*, and his emphasis on a range of dangers that face children

during development, recognize the severity of the risks posed by association as unique within this array. A legacy of Lockean association, less celebrated but no less worthy of scholarly attention than the anti-Lockean associationism found in Hume, is the integration of pathology into the developmental picture of the self.

5 CODA: CASTLES IN THE AIR

"I Doubt not but my Reader," Locke writes in Book IV of the *Essay*, "may be apt to think, that I have been all this while only building a Castle in the Air; and be ready to say to me, to what purpose all this stir? Knowledge, say you, is only the perception of the agreement or disagreement of our own *Ideas*: but who knows what those *Ideas* may be? Is there any thing so extravagant, as the Imagination of Men's Brains?" (IV.vi.1, p. 562). Locke imagines his reader objecting that his account cannot differentiate between "the Reasonings of a sober Man," such as the "Demonstrations of *Euclid*," and the "Visions of an Enthusiast" which he refers to, repeating the odd phrase he has just applied to his own philosophy, as "Castles in the Air" (ibid., p. 563). It was a phrase of growing popularity in his day, used to describe an idle daydream or fancy, and by Burton to refer to the ravings of madmen.[66] For Locke, the second application here clearly indicates a mad error: a chain of ideas containing irreparable associated ideas that stop it from being annexed to the consciousness.

Almost a century and a half later a young Charles Darwin (1809–1882), newly returned from his voyage around the world on the H.M.S. *Beagle*, began to think about insanity. In the privacy of a notebook (labeled M for its "metaphysical" contents) Darwin jots down a series of notes about "castles in the air."[67] One of his sisters reports never having such fancies — another

---

[66] Malebranche, when describing madness, uses a more archaic French form of the phrase, "un château en Espagne": madmen "wonder at everything, they exclaim about everything without judgment or discernment […] they build castles in Spain" (Malebranche 1997, p. 165).

[67] It is clear from Darwin's correspondence that he was familiar with the phrase in the nontechnical sense of fancy or mind-wandering. In a letter to his erstwhile professor John Henslow detailing his excitement over his upcoming voyage a young Darwin writes, "What changes I have had: till one to day I was building castles in the air about hunting Foxes in

has them often, "but not of an inventive class." But Darwin is really interested in how castles in the air seem to defy the active powers of the understanding: "The facility with which a castle in the air is interrupted & utterly forgotten—, so as to feel a severe disappointment. In real train of thought this does not happen." In a castle, Darwin writes, the "train cannot be discovered—is closely analogous to my Fathers positive statement that insanity is only cured by forgetfulness.— & the approach to believing a vivid castle in the air, or dreams real again explains insanity" (Barrett et al. 2009, pp. 527–528).[68] The Lockean themes here — castles in the air escaping consciousness, not being reformable, standing apart from healthy thought — are not coincidental, and the space between Locke's castles and Darwin's can suggest the contours for a more modest tradition within associationism — what we may call *medical associationism* — that flourished after the *Essay*.

In my introduction I argued that Hume takes hold of Locke's notion of association as a form of pathology and universalizes it, suggesting an unsteadiness at the foundations of human cognition, with the justification of knowledge a question left on the table for later philosophers such as Kant to pick up. Locke would have been horrified by this interpretation of association as a universal mechanism of cognition, and would have rejected it as an architectonic of madness. He would also have rejected David Hartley's Newtonian associationism, very different from Hume's (whom Hartley apparently was ignorant of, and vice versa). Hartley used the term

---

Shropshire, now Lamas in S America" (Letter 118, the Darwin Correspondence Project). In another letter to Henslow written in 1848, an older and more sober Darwin draws a similar contrast to Locke's, above: "I believe there exists," Darwin wrote, "& I feel within me, an instinct for truth, or knowledge or discovery, of something same nature as the instinct of virtue, & that our having such an instinct is reason enough for scientific researches without any practical results ever ensuing from them […] I feel that such study is better than castle-building" (Letter 1167, the Darwin Correspondence Project).

[68] All quotations from Darwin's notebooks are transcriptions, so contain spelling errors, cross-outs, etc.

"association" to refer to neurological processes that gave rise to ideas, and that could explain the relations between them. As we saw in Chapter 2, Locke had no interest in physiological speculations of this kind.

Most of all Locke would have been disturbed by the universality attributed to association by both Hartley and Hume, which left little role for the active powers of the understanding. Rather than being different in kind from healthily connected ideas, Hartley believed mad ideas were simply an extreme of the common sort of upset to the neural harmonics of the brain experienced in the course of everyday life — no line could be drawn between madness and sanity. Hartley draws a close connection between madness and personal identity, in notably Lockean terms; discussing madness he writes,

> "A particular Set of Ideas shall be extremely magnified, and, consequently, an unnatural Association of Sameness or Repugnancy between them generated, all other Ideas and Associations remaining nearly the same. Thus, suppose a Person, whose nervous System is disordered, to turn his Thoughts accidentally to some barely possible Good or Evil. If the nervous Disorder falls in with this, it increases the Vibrations belonging to its Idea so much, as to give it a Reality, a Connexion with Self. For we distinguish the Recollection and Anticipation of things relating to ourselves, from those of things relating to other Persons, chiefly by the Difference of Strength in the Vibrations, and in their Coalescences with each other." (1749, p. 401)

Nonetheless the intellectual labor so central to Locke's theory of personality has been replaced, here, by the vibrations of nerves. Madness is not the defeat of the active powers by chaotic, unintentional forces, but an extreme of the common mechanisms of cognition. Hume

sees a similar fluidity between sanity and madness, though he frames it in ideational, as opposed to neurophysiological, terms:

> "As a lively imagination very often degenerates into madness or folly, and bears it a great resemblance in its operations; so they influence the judgment after the same manner, and produce belief from the very same principles. When the imagination, from any extraordinary ferment of the blood and spirits, acquires such a vivacity as disorders all its powers and faculties, there is no means of distinguishing betwixt truth and falshood; but every loose fiction or idea, having the same influence as the impressions of the memory, or the conclusions of the judgment, is receiv'd on the same footing, and operates with equal force on the passions." (1978, p. 123)

Again, for Hume madness is a difference of degree, not kind — he compares it to the impassioned state of the poet who, "in the warmth of poetic enthusiasm […] has a counterfeit belief" (ibid.).

Locke's narrower conception of associated ideas as a distinct form of pathology that stands in contrast to healthy thought was abandoned, and his narrow diagnosis applied broadly. More generally the faculty psychology of Locke's day was eclipsed in the 18th century by Newtonian models of cognition inspired by wave theory, by hydraulics, and by theories of reflex action. Towards the end of the 18th century physicians of the mind like William Cullen, William Battie and John Haslam enthusiastically took up Hartley's model, and Thomas Arnold developed a new psychiatric nosology by combining the Lockean distinction between sensation and reflection and the Hartleian vibrational theory. Locke's emphasis on healthy connections between ideas, and warnings about the dire effects of faulty connections, became the basis for an

associationist theory of psychopathology within a general associationist theory of cognition. Missing was the theoretical apparatus provided by the *Essay* that differentiated healthy connections from associations on the basis of the involvement of the understanding.

As Robert Hoeldtke has pointed out, however, the passive and mechanical picture of Hume and Hartley found resistance among a pocket of philosophically-oriented physicians in the early 19th century. While Hoeldtke finds resonance between this opposition to the "school of philosopher-physicians stemming from Locke to Hartley" and the early modern Scholastic metaphysics they replaced (Hoeldtke 1967, p. 43), it is better seen as a *return* to Locke. The regrowth of interest among psychopathologists in the faculties was due to the looming influence of Thomas Reid over Scottish philosophy, and his powerful rejection of the replacement of the understanding's active powers with physiological processes. Philosophers sympathetic to Reid's critique of Hume who were nonetheless taken by the general connectionist picture integrated elements of the new faculty psychology into their associationism. Thomas Brown (1778–1820), for example, replaced the term "association" with "suggestion," to underscore the non-determinative nature of connections between ideas and the role of the active intellect in picking and choosing between them (ibid., p. 56). Dugald Stewart (1753–1828) treated association as an ability of the mind, that worked in conjunction with attention to shape cognition. Madness was the lack of this capacity:

> "In madness, the power of the will over the body remains undiminished, while its influence in regulating the train of thought is in a great measure suspended; either in consequence of a particular idea, which engrosses the attention, to the exclusion of everything else, and which we find it impossible to banish by our

efforts; or in consequence of our thoughts succeeding each other with such rapidity, that we are unable to stop the train." (quoted in Hoeldtke 1967, p. 57)

Others in this tradition include James Prichard, John Conolly, and John Abercrombie, the latter of whom presents perhaps the most philosophically sophisticated account of madness as a breakdown in the voluntary powers of the understanding. Abercrombie (1780–1844) couches this distinction in the language of consciousness; mad ideas, he argues, are formed in the absence of our usual conscious awareness. Abercrombie collected a compendium of examples of what he called "double consciousness," such as disorders of perception, somnambulance, fugue states, and madness, and explicitly tied such phenomena to pathologies of personal identity (Rieber 2006, p. 15). Hoeldtke persuasively argues that the tension between the active and passive faculties brought on by the integration of the associationist and faculty psychologies in the 19th century allowed for the theorization of the sub-conscious, bolstered by a shot of German idealist philosophy from the pens of William Hamilton and others (Hoeldtke 1967, p. 62).

These names — John Abercrombie, Thomas Brown, Dugald Stewart — are frequent in Darwin's notebooks, and their unique strain of medical associationism was also fed to him by his grandfather, the famous Scottish philosopher Erasmus Darwin, and his father, a physician.[69] Darwin noted down cases of split consciousness from his father's practice, and described the "castles in the air" — trains of ideas separate from the conscious self — that characterized mad

[69] "My father quite believe my grand F doctrine is true, that the only cure for madness is forgetfulness" (Barrett et al. 2009, p. 523) — and, in a list of items to follow up on scratched in the back of his "M" notebook, "My Father about double consciousness.—& somnambulism" (ibid., p. 156). There are many more such references throughout Darwin's notebooks to "my father's theory" and double consciousness in general.

people.[70] It is clear from the notebooks, however, that Darwin's interest in psychopathology is in the service of his current obsession, in the late 1830s — the mechanism behind the inheritance of characteristics.

Darwin was especially preoccupied with instinct. His proto-evolutionary notion was that instincts were like habits, but were developed over generations rather than the lifespan of the individual. But once he admitted to "Thought (or desires more properly) being heredetary," he had to further confess to himself that "it is difficult to imagine it anything but structure of brain heredetary" and goes on to chide himself: "Oh you Materialist!" (Barrett et al. 2009, p. 291). Darwin solves the problem in Notebook C, written just before Notebook M. A couple hundred pages into the volume he begins to refer to "My view of instinct," and emphasizes his belief in the heritability of instinctual behaviors: "Man by effort of Memory can remember how to swim having once learnt," he writes, "& if that was a regular contingency the brain would become webfooted & there would be no act of memory." Instinct is therefore "memory transmitted without consciousness" (Barrett et al. 2009, p. 292).

For evidence Darwin draws on his father's patients. One case, a Miss Cogan, was completely senile, but still able to recall the songs of her childhood. Darwin compares her humming to the song of a bird, writing, "Miss C. memory cannot be called memory because she did not remembered, it was an habitual action of thought-secreting organs, brought into play by morbid action" (Barrett et al. 2009, p. 521). In his reflections Darwin is fascinated by the

---

[70] Indeed one of Darwin's earliest memories, recorded in his autobiography, is of visiting the house of a family friend whose wife was deranged. "The poor creature," Darwin writes, "as soon as she saw me, was in the most abject state of terror that I ever saw, weeping bitterly and asking me over and over again, 'Is your father coming?'" The doctor, it turned out, had purposefully so terrorized the woman that the mere mention of his name could stop her violent behavior" (Darwin 1958, p. 40). This example of association is almost identical to one Locke offers in II.xxxiii.14.

varieties of trains of thought that could literally capture the imagination of the individual, overriding their ability to think freely. Miss Cogan was only able to communicate one of the trains of memory that existed in her brain, one that was of the embedded sort Darwin chooses to call instinctual — a castle in the air. In the healthy brain, he believes, different trains are woven together; a combination of the traces of inherited habits, the habitual actions acquired in the individual's own lifetime, and the actively manufactured chains of ideas created in moments of conscious reasoning, during which mental energy is expounded.

Darwin admits that he can read a novel for hours, but any length of time spent producing work as demanding as his recently-completed paper on the geography of Glen Roy seems "analogous to muscle in one position great fatigue. May explain excessive labour of inventive thought" (ibid., p. 541). In other words, "Perhaps one cause of the intense labour of *original inventive* thought is that none of the ideas are habitual, nor recalled by obvious associations. As by reading a book" (ibid., p. 540). It is precisely the capacity for this kind of mental work that is diseased in the mentally ill.

It is in this sense that Darwin contrasts "the labour of invented thought" with the building of "castles in the air," in which the mind first flits from idea to idea in patterns of connection that the castle-builder might not be critically aware of, such that, emerging from one's castle, one might be at a loss for how one arrived there in the first place. The only way to recover the fantasy is "by going to the beginning of the castle," because otherwise the "train cannot be discovered." Any student who has ever day-dreamed can sympathize with Darwin's observation that "the facility with which a castle in the air is interrupted & utterly forgotten" can lead to "severe disappointment," as well as the contrast he draws with a real train of thought, which (for better or worse) one cannot lose track of because of "papers, &c &c round one" (ibid., p. 527).

Insofar as a castle in the air is not accessible to the analyzing intellect, even though the castle-builder is vividly aware of its construction, Darwin concludes that the lost associations that constitute the architecture of a castle resemble the thoughts of a madman. The isolated train of thought that makes up an abandoned castle is "analogous to my Father's positive statement that insanity is only cured by forgetfulness.-- & the approach to believing a vivid castle in the air, or dreams real again explains insanity" (ibid.). Lunacy will lift if the invalid is able to neglect a pathological train of memory for a healthier one, much in the way a scholar cutting off a daydream abandons his fantasy to devote himself to hard analysis. Darwin's thesis is that the more intelligent a man (or animal) the more he is able to focus and avoid slipping into instinctual action. As is its wont, his mind races ahead, sniffing out evidence to support his idea: "Do people of weak intellects easily fall into *habits*. Get facts about instincts of mongrel dogs" (ibid., p. 596).

It is worth noting that to the extent that Darwin clarifies what he means by consciousness, it is a Lockean view. Scribbling on a sheet he stuffed into a folder called "Old & Useless Notes" Darwin writes, "Consciousness is *sensation* […] with memory added to it" (ibid., p. 601). In other words, consciousness for Darwin is merely the ability to connect immediate experience with previous ideas. Darwin wrote in the margin of Abercrombie's *Inquiries Concerning the Intellectual Powers and the Investigation of Truth* that the "insane man has perfect *consciousness* – somnambulism has not" (Barrett et al. 2009, p. 607). The problem with madness, as with dreams, Darwin declares in his notebook, is the madman has two or more *incommunicable* consciousnesses: "parallel trains of thought necessary heirs to every action, & always running on in mind [are absent so] one [can] not compare the castle with them, and therefore [can] not *doubt* or *believe.*" (ibid. p. 547). Unlike the somnamulist, however, the madman may have a perfect

grasp of the trains of thought that tell him about the external world, with merely some aberrations that form an additional and pathological train of thought. Rather than a lack of sensory input from the external world, he is troubled by a disease of ideas, the "unfolding & generalizing of the means by which an instinct is transmitted" (ibid., p. 576).

Darwin's new sketch of the madman is of one plagued with "double individuality," in which the brain has "whole trains of thoughts, feeling & perception separate, from the ordinary state of the mind." Here Darwin finds a powerful new vocabulary to explain to himself his idea of instinct. "Now if memory of a tune & words can thus lie dormant," he muses in considering the case of Miss Cogan, "during a whole life time, quite unconsciously of it, surely memory from one generation to another, also without consciousness, as instincts are, is not so very wonderful" (ibid., p. 521). By the end of Notebook M Darwin is ready to argue that all innate ideas are merely such inherited trains of thought. Just as the experiences of the individual leave striations on the brain which can continue to weave in and out of conscious experience in the form of memory, the experiences of one's ancestors remain as acquired characteristics.

Darwin even considers (this is his metaphysical notebook, after all) the implications of his theory of the inheritance of ideas for church doctrine:

> "The possibility of the brain having whole train of thoughts, feelings & perception separate, from the ordinary state of mind, is probably analogous to the double individuality implied by habit, when one acts unconsciously with respect to more energetic self […] Agrees with insanity, as in Dr Ash's[71] case, when he struggled as it were with a second & unreasonable man. –If one could remember all ones father's actions, as one does in second childhood, or when drunk they would not

[71] Dr. Ash was another patient of Darwin's father's.

be more different, & yet they would make one's father & self one person--& thus eternal punishment explained." (ibid., p. 538)

These are very Lockean thoughts, not only because of their forensic valence but also for their engagement with the question of innate knowledge and inborn depravity. Darwin would have of course have learned about the debate surrounding innate ideas at Cambridge, and he reflects on it in his journal after reading a review of the works of Coleridge in the *Westminster Review* (1840). Darwin paraphrases the author — writing anonymously, but now known to scholars to be the young John Stuart Mill — as follows:

> "[He] says the great division amongst metaphysicians – the school of Locke, Bentham, & Hartley, &. The school of Kant to Coleridge, is regarding the sources of knowledge.—whether 'anything can be the object of our knowledge except our experience.'—is this not almost a question of whether we have any instincts, or rather the amount of our instincts—surely in animals according to usual definition, there is much knowledge without experience. So there may be in men." (ibid., p. 610)

This foray into professional philosophy is rare for Darwin, but not unique. He engages directly with the question of innate knowledge in one other place I have found, writing, "Plato […] says in Phaedo that our '*necessary ideas*' arise from the preexistence of the soul, are not derivable from experience.—read monkeys for preexistence." The remainder of Notebook M is preoccupied with field observations of the expression of emotion in apes and men, as Darwin turns theory to praxis and begins to investigate the natural history of human reason in the living fossils of our next of kin.

With his original account of instinct, a key step in the development of his revolutionary theory of natural selection, Darwin answered an anxious prompt written (also in a private notebook) by Locke centuries earlier: "How birds imitate sounds if they doe not heare and retrieve the idea of those sounds in their memorys, and consequently have sence, is hard for me to conceiv" (Dewhurst 1963, p. 130).[72] Darwin's solution is revolutionary: habits can, outside of the ambit of consciousness, become hereditary, such that while all our ideas are a result of experience, we nonetheless already come into the world with some preformed. Darwin's solution would have been unthinkable without the particular strain of medical associationist thought that re-enlivened Locke's original theory of association as consciousness-splitting and passive, in contrast (if not in conflict) with the active powers of the understanding. We may find rather uncharitable, then, Darwin's note that "He who understands the baboon would do more towards metaphysics than Locke" (Barrett et al. 2009, p. 539).

[72] It seems Locke only admitted this anxiety to himself, laughing it off in public. In an anecdote of Yeats's retold by Craig (2007, p. 130), "when Locke's French translator Coste asked him now, if there were no 'innate ideas,' he could explain the skill shown by a bird in making its nest, Locke replied, 'I did not write to explain the actions of dumb creatures,' and his translator thought the answer 'very good, seeing that he had named his book *A Philosophical Essay upon Human Understanding*."

BIBLIOGRAPHY

Anstey, Peter. 2011. *John Locke and Natural Philosophy*. Oxford: Oxford University Press.

Anstey, Peter, and John Burrows. 2009. John Locke, Thomas Sydenham, and the authorship of two medical essays. *Electronic British Library Journal*, Article 3.

Atherton, Margaret. 1998. Locke and the issue over innateness. In *Locke*, ed. Vere Chappell, 48–59. Oxford: Oxford University Press.

-----. 2008. Locke's theory of personal identity. *Midwest Studies in Philosophy* 8(1): 273–93.

Ayers, Michael. 1993. *Locke: Epistemology and Ontology*. London: Routledge.

Bacon, Francis. 2000. *The New Organon.* Ed. Lisa Jardine and Michael Silverthorne. Cambridge: Cambridge University Press.

Barrett, Paul H, Peter J Gautrey, Sandra Herbert, David Kohn, and Sydney Smith. 2009. *Charles Darwin's Notebooks, 1836–1844*. Cambridge: Cambridge University Press.

Balibar, Etienne. 2013. *Identity and Difference: John Locke and the Invention of Consciousness.* London and New York: Verso Books.

Behan, David P. 1979. Locke on persons and personal identity. *Canadian Journal of Philosophy* 9 (1): 53–75.

Brown, Theodore M. 1970. “The College of Physicians and the Acceptance of Iatromechanism in England, 1665–1695.” *Bulletin of the History of Medicine* 44 (1): 12–30.

-----. 1987. Medicine in the shadow of the *Principia*. *Journal of the History of Ideas* 48 (1): 629–658.

Bynum, C W. 1990. Material continuity, personal survival, and the resurrection of the body: a Scholastic discussion in its medieval and modern contexts. *History of Religions* 30 (1): 51–85.

Casson, Douglas John. 2011. *Liberating Judgment: Fanatics, Skeptics, and John Locke's Politics of Probability*. Princeton: Princeton University Press.

Chappell, Vere. 2011. Locke's theory of ideas. In *The Cambridge Companion to Locke*, ed. Vere Chappell, 26–55. Cambridge: Cambridge University Press.

Charland, Louis C. 2008. Alexander Crichton on the psychopathology of the passions. *History of Psychiatry* 19 (3): 275–296.

Clericuzio, Antonio. 2009. From Van Helmont to Boyle: a study of the transmission of Helmontian chemical and medical theories in seventeenth-century England. *The British Journal for the History of Science* 26 (3): 303–334.

Corneanu, Sorana. 2012. *Regimens of the Mind: Boyle, Locke, and the Early Modern Cultura Animi Tradition*. Chicago: University of Chicago Press.

Craig, Cairns. 2007. *Associationism and the Literary Imagination: From the Phantasmal Chaos*. Edinburgh: Edinburgh University Press.

Darwin, Charles. 1958. *The Autobiography of Charles Darwin, 1809–1882*. Ed. Nora Barlow. London and Glasgow: Collins Clear-Type Press.

Descartes, René. 1983. *Oeuvres De Descartes*. Ed: Charles Adam and Paul Tannery. Paris: Librairie Philosophique J. Vrin.

-----. 1985. *Volume 1 of The Philosophical Writings of Descartes*. Ed. John Cottingham, Robert Stoothoff, and Dugald Murdoch. Cambridge: Cambridge University Press.

Dewhurst, Kenneth. 1963. *John Locke, 1632–1704, Physician and Philosopher*. London: Wellcome Historical Medical Library.

Garrett, Aaron. 2013. Seventeenth-century moral philosophy: self-help, self-knowledge, and the devil's mountain. In *The Oxford Handbook of the History of Ethics*, ed. Roger Crisp, 229–279. Oxford: Oxford University Press.

Gaukroger, Stephen. 2002a. *Descartes' System of Natural Philosophy*. Cambridge: Cambridge University Press.

-----. 2002b. *The Soft Underbelly of Reason: The Passions in the Seventeenth Century*. London: Routledge.

Goodey, C F. 1994. John Locke's idiots in the natural history of mind. *History of Psychiatry* 5 (18): 215–50.

Guyer, Paul. 2011. Locke's philosophy of language. In *The Cambridge Companion to Locke*, ed. Vere Chappell, 115–145. Cambridge: Cambridge University Press.

Hankinson, R J. 2011. Stoic epistemology. In *The Cambridge Companion to the Stoics*, ed. Brad Inwood, 59–84. Cambridge: Cambridge University Press.

Harrison, John R., and Peter Laslett. 1965. *The Library of John Locke*. Oxford: Oxford University Press.

Harrison, Peter. 2007. *The Fall of Man and the Foundations of Science*. Cambridge: Cambridge University Press.

Hartley, David. 1749. *Observations on Man: His Frame, His Duty and His Expectations*. London: S. Richardson.

Hoeldtke, R. 1967. The history of associationism and British medical psychology. *Medical History* 11 (1): 46–65.

Hume, David. 1978. *A Treatise of Human Nature*. Ed. Peter H Nidditch and L. A. Selby-Bigge. Oxford: Oxford University Press.

-----. 1938. An abstract of a treatise of human nature 1740: a pamphlet hitherto unknown by David Hume." Ed. J.M. Keynes and P. Sraffa. Cambridge: Cambridge University Press.

James, Susan. 1997. *Passion and Action: the Emotions in Seventeenth-Century Philosophy*. Oxford: Clarendon Press.

Kallich, Martin. 1945. The association of ideas and critical theory: Hobbes, Locke, and Addison. *English Literary History* 12 (4): 290.

Lähteenmäki, Vili. 2008. The sphere of experience in Locke: the relations between reflection, consciousness, and ideas. *Locke Studies* 8 (1): 59–100.

Leites, Edmund. 1988. Casuistry and character. In *Conscience and Casuistry in Early Modern Europe*, ed. Edmund Leites, 119–133. Cambridge: Cambridge University Press.

Locke, John. 1824. *The Works of John Locke in Nine Volumes.* London: J. Rivington.

-----. 1960. *Two Treatises of Government*. Ed. Peter Laslett. Cambridge: Cambridge University

Press.

-----. 1974. *An Essay Concerning Human Understanding*. Ed. Peter H Nidditch. Oxford: Clarendon Press.

-----. 1979. *The Correspondence of John Locke: in Eight Volumes*. Ed. E. S. De Beer. Oxford: Clarendon Press.

-----. 1996. *Some Thoughts Concerning Education and, of the Conduct of the Understanding*. Ed. Ruth W Grant and Nathan Tarcov. Indianapolis: Hackett Publishing.

LoLordo, Antonia. 2012. *Locke's Moral Man*. Oxford University Press.

Malebranche, Nicholas. 1997. *The Search After Truth: with Elucidations of the Search After Truth*. Ed. Thomas N. Lennon and Paul J. Olscamp. Cambridge: Cambridge University Press.

Meynell, G. G. 2006. John Locke and the preface to Thomas Sydenham's *Observationes Medicae*. *Medical History* 50 (1): 93–110.

Milton, J R. 2001. Locke, medicine, and the mechanical philosophy." *British Journal for the History of Philosophy* 9 (2): 221–43.

MacDonald, M. 1981. Insanity and the realities of history in early modern England. *Psychological Medicine* 11: 11–25.

More, Henry. 1656. *Enthusiasmus Triumphatus*. London: I. Flesher.

Olivecrona, Karl. 1974. Locke's theory of appropriation. *The Philosophical Quarterly* 24 (96): 220–234.

Passmore, J A. 1978. Locke and the ethics of belief. *Proceedings of the British Academy* 64: 185–208.

Porter, Roy. 1983. The rage of party: a glorious revolution in English psychiatry? *Medical History* 27 (1): 35–50.

Rapaport, David. 1974. *The History of the Concept of Association of Ideas*. New York: International Universities Press.

Rickless, Samuel. Forthcoming. "Locke's 'sensitive knowledge': knowledge or assurance? *Oxford Studies in Early Modern Philosophy*, vol. 7.

Rieber, Robert W. 2006. *The Bifurcation of the Self: The History and Theory of Dissociation and its Disorders*. Springer Science & Business Media.

Romanell, Patrick. 1984. *John Locke and Medicine*. Amherst, NY: Prometheus Books.

Schmidt, Jeremy. 2004. Melancholy and the therapeutic language of moral philosophy in seventeenth-century thought. *Journal of the History of Ideas* 65 (4): 583–601.

Sellars, John. 2013. *The Art of Living*. Edinburgh: A&C Black.

Sena, John F. 1973. Melancholic madness and the Puritans. *Harvard Theological Review* 66 (3): 293–309.

Smith, Jad. 2006. Custom, association, and the mixed mode: Locke's early theory of cultural reproduction. *English Literary History* 73 (4): 831–53.

Sprat, Thomas. 1702. *The History of the Royal Society of London, for the Improving of Natural Knowledge*. London: Scot.

Spellman, W M. 1988. *John Locke and the Problem of Depravity*. Oxford: Clarendon Press.

Sutton, John. 1998. *Philosophy and Memory Traces: From Descartes to Connectionism*. Cambridge: Cambridge University Press.

-----. 2010. Carelessness and inattention: mind-wandering and the physiology of fantasy from Locke to Hume. In *The Body as Object and Instrument of Knowledge*, ed. Charles T. Wolfe and Ofer Gal, 243–63. Dordrecht and New York: Springer Verlag.

Suzuki, Akihito. 1992. Mind and Its disease in Enlightenment British medicine. Thesis submitted to the Philosophy Department, University College London.

Sydenham, Thomas. 1966. *Dr. Thomas Sydenham, 1624–1689*. Ed. Kenneth Dewhurst. Berkeley: University of California Press.

Thomson, Ann. 2008. *Bodies of Thought: Science, Religion and the Soul in Early Englightenment*. New York: Oxford University Press.

Tully, James. 2009. Governing conduct. In *Conscience and Casuistry in Early Modern Europe*, ed. Edmund Leites, 12–71. Cambridge: Cambridge University Press.

Tuveson, Ernest Lee. 1960. *The Imagination as a Means of Grace: Locke and the Aesthetics of Romanticism*. Berkeley: University of California Press.

Walmsley, Jonathan. 1998. John Locke's natural philosophy (1632–1671). Thesis submitted to the Philosophy Department, King's College London.

Walmsley, Jonathan. 2000. *Morbus*-Locke's early essay on disease. *Early Science and Medicine* 5 (4): 367–93.

-----. 2013. Review: Anstey, *John Locke and Natural Philosophy. Locke Studies* 12.

Walsh, Julie. 2015. Locke and the power to suspend desire. *Locke Studies* 14 (February): 121–57.

Warren, Howard Crosby. 1921. *A History of the Association Psychology*. New York, Chicago and Boston: Charles Scribner's Sons.

Waxman, Wayne. 2005. *Kant and the Empiricists: Understanding Understanding*. Oxford: Oxford University Press.

Weinberg, Shelley. 2011. Locke on personal identity. *Philosophy Compass* 6 (6): 398–407.

Williams, Bernard. 1970. The self and the future. *The Philosophical Review* 79 (2): 161–80.

Willis, Thomas. 1684. *A Medical-Philosophical Discourse of Fermentation, or, of the Intestine Motion of the Particles in Every Body*. London: H. Clark for T. Dring.

Williston, Byron. 2002. Self-deception and the ethics of belief." *Philo* 5 (1): 62–83.

Winkler, Kenneth. 1991. Locke on personal identity. *Journal of the History of Philosophy* 29 (2): 201–26.

Wolfe, Charles T. 2010. "Empiricist Heresies in Early Modern Medical Thought." In *The Body as Object and Instrument of Knowledge*, Ed: Charles T Wolfe and Ofer Gal. Dordrecht and New York: Springer Verlag. 333–44.

-----. 2011. Why was there no controversy over life in the Scientific Revolution?" In *Controversies Within the Scientific Revolution*, ed. Marcelo Dascal and Victor D Boantza. 187–219. Amsterdam: John Benjamins Publishing Company.

-----. Forthcoming. From Locke to materialism: Empiricism, the brain and the stirrings of ontology. In *18th-Century Empiricism and the Sciences*, ed. A. L. Rey S. Bodenmann.

Wolfe, David E. 1961. Sydenham and Locke on the limits of anatomy. *Bulletin of the History of Medicine* XXXV (3): 193–220.

Wolterstorff, Nicholas. 1996. *John Locke and the Ethics of Belief*. Cambridge: Cambridge University Press.

Wood, Paul. 2003. Science in the Scottish Enlightenment. In *The Cambridge Companion to the Scottish Enlightenment,* ed. Alexander Broadie. Cambridge: Cambridge University Press. 94–116.

Wright, John P. 1987. Association, madness, and the measures of probability in Locke and Hume. In *Psychology and Literature in the Eighteenth Century*, ed. Christopher Fox, 103–27. New York: AMS Press.

-----. 1991. Locke, Willis, and the seventeenth-century epicurean soul. In *Atoms, Pneuma, and Tranquillity: Epicurean and Stoic Themes in European Thought,* ed. Margaret J Osler, 239–258. Cambridge: Cambridge University Press.

Yaffe, Gideon. 2000. *Liberty Worth the Name: Locke on Free Agency*. Princeton: Princeton University Press.

Yolton, John W. 1956. *John Locke and the Way of Ideas*. Oxford: Oxford University Press.

Young, R M. 1968. Association of ideas. In *Dictionary of the History of Ideas*, ed. Philip P. Weiner. New York: Charles Scribner's Sons.