Empiricist Roots of Modern Psychology

From the thirteenth through the sixteenth centuries, European philosophers were preoccupied with using their newfound access to Aristotle’s metaphysics and natural philosophy to develop an integrated account, hospitable to Christianity, of everything that was thought to exist, including God, pure finite spirits (angels), the immaterial souls of humans, the natural world of organic objects (plants, animals, and human bodies) and inorganic objects. This account included a theory of human mentality. In the sixteenth and early seventeenth centuries, first in astronomy and then, later, in physics, the tightly knit fabric of this comprehensive medieval world view began to unravel.

The transition from the old to the new was gradual, but by 1687, with the publication by Isaac Newton (1642-1727) of his *Principia Mathematica*, the replacement was all but complete. Modern physical science had fully arrived, and it was secular. God and angels were still acknowledged. But they had been marginalized. Yet, there was a glaring omission. Theorists had yet to expand the reach of the new science to incorporate human mentality. This venture, which initially was called “moral philosophy” and came to be called “the science of human nature,” became compelling to progressive eighteenth century thinkers, just as British empiricism began to seriously challenge an entrenched Cartesian rationalism.

Rationalism and Empiricism

The dispute between rationalists and empiricists was primarily over concepts and knowledge. In response to such questions as, where does the mind get its stock of concepts, how do humans justify what they take to be their knowledge, and how far does human knowledge extend, rationalists maintained that some concepts are innate, and hence not derived from experience, and that reason, or intuition, by itself, independently of experience, is an important source of knowledge, including of existing
things. They also maintained that one could have a priori knowledge of the existence of God. Empiricists, on the other hand, denied that any concepts are innate, claiming instead that all of them are derived from experience. They also tended to claim that all knowledge of existing things is derived from experience. And, as time went on, empiricists became increasingly skeptical, first, that one could have a priori knowledge of God and, later, that one could have knowledge of God at all.

Rene Descartes (1596-1650), who along with Galileo Galilei (1564-1642) was one of the founders of modern physical science, was the most influential rationalist of the seventeenth century. Even though when it came to the study of animal biology, Descartes was an avid experimentalist, in his abstract philosophy he elevated rational intuition over sense experience as a source of knowledge. He also claimed that humans have innate ideas, such as an idea of God, which do not come from experience. And he claimed that through reason alone, independently of appeal to experience, one could demonstrate the existence of God and the existence of immaterial souls—one such soul, intimately conjoined with a body, for each human person.

During the time that Descartes was making his major philosophical and scientific contributions, he had predecessors and contemporaries who were well known and highly influential empiricists. Chief among these were Francis Bacon (1561-1626), Pierre Gassendi (1592-1655), and Thomas Hobbes (1588-1679). However, Descartes' rationalism overshadowed the empiricism of his day—providing the framework for the most influential philosophy of the seventeenth century. It was not until close to the dawn of the eighteenth century, when John Locke (1632-1704) published his Essay Concerning Human Understanding (1690-94) that the tide began to turn against rationalism and toward empiricism.

In 1690, Aristotelean science was still firmly entrenched in the universities. Even so, in his Essay Locke not only expressed contempt for it, but generally dismissed it without much argument, taking it as obvious that it was on the wrong track. His main target, against which he argued at length, was Cartesian rationalism. In Britain especially, but also in France, Locke found an eager audience. He quickly became the
most influential empiricist of the modern era.

**Concepts**

One of Locke’s central ideas was that the human mind at birth is a *tabula rasa* (blank tablet) on which experience subsequently writes. He allowed that the mind might have innate capacities, such as the capacity to reason and to learn from experience, but he vehemently denied that it has any innate ideas (concepts). In trying to make this point, he taunted rationalists with the perhaps irrelevant observation that children, the mentally impaired, and “savages” lack many of the ideas that were said by rationalists to be innate. But his main thrust was to try to explain how humans *could have* acquired all of their concepts from experience, thereby making the appeal to innate ideas superfluous.

Throughout the eighteenth century many empiricists enthusiastically embraced Locke’s tabula rasa thesis, in whole or in part. These included George Berkeley (1685-1753), who allowed that humans have a *notion* (as opposed to an idea) of the self that is not derived from experience, and David Hume (1711-1776), who defended Locke’s view by refashioning a central component of the way Locke had supported it. Some other philosophers simply ran with Locke’s idea, including the French philosopher Étienne Bonnot de Condillac (1715-1780), who in his *Treatise on Sensations* (1754) claimed that external sensations by themselves could account not only for all human concepts, but for all mental operations as well. Using the example of a statue endowed with only the sense of smell, Condillac tried to explain how from this bare beginning attention, memory, judgment, and imagination—indeed, one’s entire mental life—might have developed. His views thus embodied a more extreme version of the tabula rasa perspective than can be found even in Locke.

In contrast to Condillac, many British empiricists after Locke had doubts about Locke’s explanations of the experiential origins of several of the concepts that he examined, including especially those of causation and of the self. Over time these more austere empiricists—Hume is the premier example—tended increasingly to agree that ideas as robust as the ones Locke assumed that we have could not have been derived from experience. But then, rather than rejecting Locke’s tablula rasa thesis, they
concluded that our ideas are not as robust as Locke had imagined. Thus, Hume developed his “bundle theory of the self” and his “regularity theory of causation” in order to fashion concepts of these notions thin enough that they actually could have been derived from experience. A question, then, was whether these thinner concepts were nevertheless thick enough to account for the ways humans meaningfully think about the world, especially in science.

The tabula rasa thesis played an important role in encouraging thinkers to speculate about how the mind becomes stocked with its simple ideas, how it then combines and augments these to form more complex ideas, and finally what the laws might be—the so-called principles of association—that govern how one idea leads to another in human thought. The tabula rasa thesis also put great pressure on the assumption that humans understand what it might even mean to have, or be, an immaterial self, let alone to know that one has, or is, one.

Effectively doing away with the idea that to understand human nature one must understand the role of an immaterial self in human mentality was crucial to the emergence of a scientific psychology. In the eighteenth century, empiricism, and the tabula rasa thesis in particular, was at the forefront of this important initiative. More generally, the tabula rasa thesis encouraged an austere empiricist epistemology and metaphysics that inhibited acceptance of many common sense and even scientific assumptions about the reality of the external world and our epistemological access to it, as well as about the meaning of the concepts in terms of which we think about ourselves and the world. Not all empiricists embraced this entire program, but for those who did, which included most notably Hume, empiricism tended to lead to skepticism. This encouraged other thinkers—Immanuel Kant (1724-1804) is the premier example—to explore radically alternative ways to account for human knowledge, including new proposals about how the human mind might have come to be stocked with its concepts.

Today something like the doctrine of innate ideas, under the guise of what is called nativism, has become the prevailing orthodoxy among philosophers and psychologists. However, it was not until the second half of the twentieth century that
nativism gained this sort of ascendancy, at which time nativism’s rise was due initially, and perhaps primarily, to widespread acceptance of the approach to language acquisition championed by Noam Chomsky.¹ Once nativism had made this inroad the way was open for others to advance a variety of nativist theses—for instance, for Jerry Fodor to argue that since there is no viable empiricist theory of concept acquisition it is prima facie reasonable to believe that all concepts are innate.²

**Knowledge**

In addition to Locke’s making subsequent empiricists uncomfortable by conceding too much to common sense about the content of our ideas, he also muddied his empiricist credentials by agreeing with Descartes that we have a demonstrative knowledge of God’s existence and an intuitive knowledge of our own existence. Locke even claimed to believe that the self is an immaterial substance. However, he coupled these agreements with the wildly controversial observation that matter might think.³ And, even more threatening to the idea of the self as immaterial substance, he gave an empirical account of personal identity that made no appeal to anything immaterial.

Subsequently Berkeley and Hume denied that we have a demonstrative knowledge of God’s existence. Berkeley, however, claimed that we can know on empirical grounds that God exists. And he claimed that we have an intuitive knowledge of our own existence as an immaterial substance (privately he expressed doubt on the point). Hume, in the work that he published during his lifetime, eschewed any concession to the idea that God exists and even denied that we intuit our own existence, at least if it is conceived as robustly as Locke conceived it. In addition, Hume famously gave more empirically austere analyses of several of Locke’s key notions. Other empiricists, as we shall see, did not become so preoccupied with Locke’s tabula rasa thesis that they allowed their commitment to an austere empiricist epistemology to interfere with their contributions to the newly emerging science of human nature. Instead, they allowed themselves realistic assumptions about the material world and our epistemological access to it. David Hartley (1705-1757), Adam Smith (1723-1790), and Joseph Priestley (1733-1804) were in this group.
There was, thus, a major divide within the empiricist camp, not so much over whether Locke’s tabula rasa thesis is true, since few empiricists questioned it, but over the role that it and the austere empiricist epistemology that it encouraged should play in science, particularly in an empirical investigation of the human mind. But, due to the high visibility and persuasiveness of those empiricists who were preoccupied with the more austere approach, empiricism quickly became linked with skepticism, a reputation that it retained into our own times. As late as 1945, Bertrand Russell (1872-1970), himself a latter-day empiricist, wrote that Hume “developed to its logical conclusion the empirical philosophy of Locke and Berkeley, and by making it self-consistent made it incredible.” Hume, thus, represents, Russell continued, “a dead end”; in his direction “it is impossible to go further.” And, although “to refute him has been, ever since he wrote, a favourite pastime among metaphysicians,” Russell could “find none of their refutations convincing.” Russell concluded, “I cannot but hope that something less sceptical than Hume’s system may be discoverable.”

Such was the influence of the austere epistemology spawned by empiricism. But what Russell expressed is a philosopher’s worry. Whether it has much to do with how science should be conducted, and a science of psychology in particular, is a separate question. Hume, though, thought that it had a lot to do with how a science of human nature should be conducted. In his view, austere empiricism and science are inextricably linked. Hence, in his strictures about how a science of human nature should be pursued, psychology never escapes from the clutches of epistemology. That, as it turns out, was not the way forward.

**The Self**

Although Locke’s official view was that the self is an immaterial substance, he saw that for the purpose of developing a science of human nature, that idea was a non-starter. However, rather than challenge the immaterial self thesis directly, Locke turned to the topic of personal identity, where he had two main ideas, one negative and one positive. His negative idea was that the persistence of persons *cannot* be understood empirically as parasitic upon the persistence of any underlying substance, or substances, out of
which humans or persons might be composed. His positive idea was that the persistence of persons can be understood empirically in terms of the unifying role of consciousness.

Most of the time when Locke talked about consciousness in the context of talking about personal identity he meant remembers. His eighteenth century critics invariably attributed to him the view that a person at one time and one at another have the same consciousness, and hence are the same person, just in case the person at the later time remembers, from the inside, the person at the earlier time. Whether or not this is what Locke had in mind, his eighteenth century critics were right in thinking that the memory interpretation of personal identity that they attributed to him is vulnerable to decisive objections. However, almost all of them wanted to defeat what they took to be Locke’s memory view in order to retain the view that personal identity depends on the persistence of an immaterial soul.

For his part, Locke pointed out correctly that one can determine empirically whether someone retains the same consciousness over time, but not whether someone retains the same immaterial soul. As a consequence, he thought, the soul view is not only a wrong account of personal identity, but the wrong kind of account, whereas his own view, by contrast, is at least the right kind of account. As it happened, Locke was right: the kind of account he offered was riding the crest of a wave of naturalization that was about to engulf his critics.

An early indication of what was about to happen occurred soon after Locke’s death. Between 1706 and 1709 Samuel Clarke (1675-1729) and Anthony Collins (1676-1729) confronted each other in a six-part written debate. At the time, Clarke, who was Newton’s right hand man, was an enemy of empiricism and one of the most highly respected philosophers of the time, a status that he retained throughout the century. Collins, who in the last years of Locke’s life had been one of his most beloved and devoted disciples, was a relative unknown.

Clarke and Collins’ point of departure was the question of whether souls are naturally immortal, where by “soul,” they agreed to mean "Substance with a Power of
Thinking" or "Individual Consciousness." Clarke, who had a sophisticated understanding of Newtonian science and was revered throughout the century for his opposition to empiricism, defended the traditional Platonic idea that souls are immaterial. Collins countered that the soul is material.

Both men agreed that individual atoms are not conscious. Their dispute, thus, turned on the question of whether it is possible that a system of matter might think. Clarke argued that it is not possible, Collins that matter does think. Throughout their debate Clarke played the part of the traditional metaphysician. He argued largely on a priori grounds. Collins, though not always consistently, played the part of the empirical psychologist. His faltering, but often successful, attempts to reformulate traditional metaphysical issues empirically embodied the birth pangs of a new approach, one that grew steadily throughout the century. The Clarke-Collins debate is, thus, a poignant record of two thinkers' struggles to cope with a rapidly changing intellectual climate, Clarke by hanging onto the old, Collins by groping for the new.

Although Collins' approach was the progressive side of Locke's, he went beyond Locke, first, in espousing materialism, and second, in replacing Locke's metaphysically awkward same-consciousness view of personal identity with a more defensible connected-consciousness view. Throughout Collins said that he sought, and that Clarke should have been seeking, an empirical account consciousness. Collins repeatedly criticized Clarke for trying to settle by verbal fiat what could only be settled empirically. Clarke countered by reiterating a priori dogma. For instance, he claimed that strictly speaking, consciousness is neither a capacity for thinking nor actual thinking, "but the Reflex Act by which I know that I think, and that my Thoughts and Actions are my own and not Another's." He also claimed that "it would necessarily imply a plain and direct Contradiction, for any power which is really One and not Many . . . to inhere in or result from a divisible Substance." However, he conceded that his own "affirming Consciousness to be an individual Power," was neither "giving an Account of what Consciousness" nor "intended to be so." It is enough, he concluded, that "every Man feels and knows by Experience what Consciousness is, better than any Man can
explain it." As it turned out, however, this was not enough.

It soon became clear to subsequent thinkers that while intuition might be a sufficient basis to resist the reduction of the mental to the material, it was impotent as a source of explanations of mental phenomena. Collins returned to this point again and again, even claiming to be able to explain how consciousness could be transferred from a material system of the brain initially composed of certain particles to one subsequently composed of other particles, without changing the individual subject of consciousness whose brain is involved. By our current standards, his explanation is crude, but it was a genuine scientific explanation and Clarke had nothing comparable to offer.

Throughout the eighteenth century the Clarke-Collins debate was well known to subsequent theorists. Yet even though Collins’ orientation was directly toward the development of a science of psychology of a sort that would be familiar to psychologists in our own times, the extent of his influence is unclear. However, even among those who sided with Clarke there was a gradual awakening to the idea that at least for scientific purposes the self had to be understood empirically. Thus, Clarke’s bravado in his debate with Collins contrasts with the subsequent defensiveness of Berkeley and Joseph Butler (1692-1752), a few decades later, as well as with the reluctance of most immaterial soul theorists after Hume even to do battle on the issue. And whereas toward the beginning of the century, it was enough simply to defend the immateriality of the soul and related a priori doctrines, such as the reflexivity of consciousness (the view that necessarily if one is conscious, then one knows that one is conscious), without also contributing to the emerging science of human nature, eventually soul theorists tended to bracket their commitment to the immaterial soul in order to conduct meaningful empirical research. Thus, while the immateriality of the soul is crucial to Berkeley’s metaphysics, it is almost irrelevant to his inquires into vision; and although Hartley, Thomas Reid (1710-1796), and Abraham Tucker (1705-1774) remained committed to the existence of the immaterial soul, each of them segregated that commitment from their empirical inquiries.

As a consequence, in debates among theorists about the nature of the mind, it
tended to matter less and less as the century wore on what one’s view was of the immaterial soul. Toward the end of the century, Hartley, the dualist, was regarded as an ally by Priestley, the materialist, while Reid, the dualist, attacked both. And while the main influences on Tucker, the dualist, were Locke, Clarke, and Hartley, it was not Locke and Hartley’s dualism that most impressed Tucker, but their more scientific pursuits. It is only a slight exaggeration to suggest that Priestley could have put forth the very same views he did, even if, like Hartley, he had been a dualist; and Reid could have put forth most of his views, even if he had been a materialist.

This bracketing of commitment to the immaterial soul, which was reinforced later in a different context by the methodological strictures of Kant, arguably was one of empiricism’s two greatest contribution to the eventual emergence of a science of psychology. The other was their contributions to formulating the principles of association. In both cases the basic message was that from the point of view of developing a science of human nature, the only ontological commitments that matter are those that can be tracked empirically and the only theories that matter those that can be confirmed or refuted empirically. Rationalists never quite got this, but it was central to the approach of empiricists. Unfortunately empiricists, for their part, tended not to get that for the purpose of doing science, it was more productive to make realistic assumptions about the world than to ground every claim in an empirically austere epistemology and metaphysics.

Self-Constiution

In empiricist traditions, it was not only the *immaterial* self that came under a cloud of suspicion, but even the *empirical* self. To see how this happened, one has to go back again to Locke, who in the *Essay* sometimes used the words *person* and *self* interchangeably, but more often used *self* to refer to a momentary entity and *person* to refer to a temporally extended one. Locke even defined the two terms differently. His definition of *person* highlighted that persons are *thinkers* and, as such, have reason, reflection, intelligence, and whatever else may be required for trans-temporal self-reference. His definition of *self* highlighted that selves are *sensors* and as such feel
pleasure and pain, and are capable of happiness, misery and self-concern.

We know how, in Locke’s view, humans come into being. It is a biological process. How do selves (or persons) come into being? His answer was that is a psychological process that begins with an organism’s experience of pleasure and pain, which gives rise, first, to the idea of a self–its own self–that is the experiencer of pleasure and pain, and then to concern with the quality of that self’s experience (each of us wants more pleasure, less pain). Then the momentary self thus constituted (or perhaps the organism) thinks of itself (or its self) as extended over brief periods of time (say, the specious present); finally, through memory and the appropriation ingredient in self-consciousness, it thinks of itself as extended over longer periods of time. Locke, thus, thought of the constitution of the self as at least being capable of being analyzed into an ordered, multi-step process. He may or may not have thought that the prior phases of this process temporally precede the subsequent phases.

Whatever Locke’s view on this question of timing, he clearly thought that self-constitution involves appropriation—a kind of self-declaration of ownership—and that appropriation and accountability go hand in hand. A person, he said, is "justly accountable for any Action" just if it is appropriated to him by his self-consciousness. He regarded the appropriation ingredient in self-consciousness as a natural relation between the organism and its present and past, which then is the basis for a non-natural relation of moral ownership.

Joseph Butler, more than any other eighteenth century critic of Locke, took Locke’s observations about the role of appropriation in self-constitution seriously. It is “easy to conceive,” Butler said, “how matter, which is no part of ourselves, may be appropriated to us in the manner which our present bodies are”. But, he continued, where there is appropriation, there must be an appropriator. Locke had an appropriator in “man,” which he distinguished from “person” and allowed might be merely a material organism. Butler thought that he (Butler) had already shown that the appropriator must be something simple and indivisible, and, hence, could not possibly be a material organism. This simple, indivisible appropriator, he assumed, is who we truly are. But
what this being appropriates, he went on to explain, is not thereby part of itself, but, rather, something it owns. Butler had learned from Locke that, for all we know, the thinking principle in us may be material. So, he astutely conceded that the appropriator might be a simple material entity.\(^{17}\) In his view, it is our simplicity, not our immateriality, that ensures our survival. He thereby adapted the Platonic argument for immortality to the purposes of an age in which materialism was on the rise, recasting the a priori in an empirical mold.

When Butler turned to the topic of personal identity per se, he argued that on a relational view such as that of Locke or Collins, people would have no reason to be concerned for the future life of the person who they nominally regard as themselves, for if our being were just to consist in successive acts of consciousness, then it would be a mistake “to charge our present selves with anything we did, or to imagine our present selves interested in anything which befell us yesterday” or will befall us tomorrow “since our present self is not, in reality, the same with the self of yesterday, but another like self or person coming in its room, and mistaken for it: to which another self will succeed tomorrow.”\(^{18}\)

In response to what Butler saw as the dangers of empirical analysis, he proposed that we take as primitive the idea of personal identity, which he said defies analysis. Like Clarke, he maintained that we can determine intuitively that we have persisted, not just in “a loose and popular sense” such as we might employ in saying of a mature oak that it is the same tree as one that stood in its spot fifty years previously, even though it and that former tree have not one atom in common, but in “the strict and philosophical sense” which requires sameness of substance.\(^{19}\) On Locke’s view, he claimed, we would have to consider ourselves to be selves and persons not really, but only in a fictitious sense. He thought that such a consequence refutes Locke’s view. And, like Clarke, he admitted that he thought this not because he thought that he could show Locke’s view to be false (he admitted that he could not), but rather because “the bare unfolding this notion [that selves are merely fictitious entities] and laying it thus naked and open, seems the best confutation of it”.\(^{20}\) Empiricists continued to struggle
with this issue throughout the nineteenth century.

One who did so was John Stuart Mill (1806-1873), who claimed that the self-knowledge that humans unquestionably have must be based on an intuitive belief in our own continued existence that comes with our ability to remember past states of mind as our own. Self and memory, Mill said, are “merely two sides of the same fact, or two different modes of viewing the same fact.” He explained that when a person—i—remembers something, “in addition” to the belief that I have “that the idea I now have was derived from a previous sensation” there is “the further conviction that this sensation was “my own; that it happened to my self.” He continued, “I am aware of a long and uninterrupted succession of past feelings, going back as far as memory reaches, and terminating with the sensations I have at the present moment, all of which are connected by an inexplicable tie, that distinguishes them not only from any succession or combination in mere thought, but also from the parallel succession of feelings” which are had by others. “This succession of feelings, which I call my memory of the past, is that by which I distinguish my Self. Myself is the person who had that series of feelings, and I know nothing of myself, by direct knowledge, except that I had them. But there is a bond of some sort among all the parts of the series, which makes me say that they were feelings of a person who was the same person throughout and a different person from those who had any of the parallel successions of feelings; and this bond, to me, constitutes my Ego.”

William James (1842-1910) later criticized Mill for having fallen back “upon something perilously near to the Soul,” quoting as evidence Mill’s remark that it is “indubitable” that “that there is something real” in the tie which is revealed in memory when one recognizes a sensation’s having been felt before, and thereby “connects the present consciousness with the past one of which it reminds me.” This tie, Mill said, “is the Ego, or Self.” Mill continued, “I ascribe a reality to the Ego—to my own mind—different from that real existence as a Permanent Possibility, which is the only reality I acknowledge in Matter.” This Ego, he concluded, “is a permanent element.” James remarked that “this 'something in common' by which they [remembered feelings] are
linked and which is not the passing feelings themselves, but something 'permanent,' of which we can 'affirm nothing' save its attributes and its permanence, what is it but metaphysical Substance come again to life?"  

James concluded that Mill here makes "the same blunder" that Hume had earlier made: "the sensations per se, he thinks, have no 'tie.' The tie of resemblance and continuity which the remembering Thought finds among them is not a 'real tie' but 'a mere product of the laws of thought;' and the fact that the present Thought 'appropriates' them is also no real tie." But, James continued, whereas Hume was content "to say that there might after all be no 'real tie,' Mill, unwilling to admit this possibility, is driven, like any scholastic, to place it in a non-phenomenal world."

In James' own approach to the self, the spirit of traditional empiricism burned brightly, but was now linked with a newfound interest both in physiology and in social interaction. From this perspective James claimed that the core of personhood is "the incessant presence of two elements, an objective person, known by a passing subjective Thought and recognized as continuing in time."  

He resolved to use the word *me* for "the empirical person" and *I* for "the judging Thought." Since the "me" is constantly changing: "the identity found by the I in its me is only a loosely construed thing, an identity 'on the whole,' just like that which any outside observer might find in the same assemblage of facts." The I of any given moment is a temporal slice of "a stream of thought," each part of which, as "I," can "remember those which went before, and know the things they knew" and "emphasize and care paramountly for certain ones among them as 'me,' and appropriate to these the rest." The core of what is thought to be the "me" "is always the bodily existence felt to be present at the time."  

Remembered-past-feelings that "resemble this present feeling are deemed to belong to the same me with it." And "whatever other things are perceived to be associated with this feeling are deemed to form part of that me's experience; and of them certain ones (which fluctuate more or less) are reckoned to be themselves constituents of the me in a larger sense," such as one's clothes, material possessions, friends, honors, and so on. But while the "me" is "an empirical aggregate of things
objectively known,” the “I” which “knows them cannot itself be an aggregate.” Rather, “it is a Thought, at each moment different from that of the last moment, but appropriative of the latter, together with all that the latter called its own.” In other words, what one calls “the I” is constantly changing. The I as a persisting thing is a fiction.

Closely related to the questions of how the self is constituted and whether anything so constituted could be a real thing was the question of how humans acquire a self-concept. Descartes had maintained that in order for anyone to be conscious one would have to know ( or be conscious) that oneself is conscious. But to know that oneself is conscious, one would have to already be in possession of a self-concept. Thus, in such a view there is no room for conscious beings to gradually develop a self-concept; they must already have one in order to be conscious in the first place. Eighteenth century rationalists, such as Clarke, continued to accept this view, and even Locke accepted it. It was not until the end of the eighteenth century that empiricists explicitly abandoned it.

The moment came in William Hazlitt’s (1778-1830) first work, An Essay on the Principles of Human Action (1805), which was the culmination of a kind of perspective on human mentality that had begun with Locke and been developed by Collins, Hume, and Priestley. According to Hazlitt people are naturally concerned about whether someone is pleased or suffers as a consequence of their actions. This is because "there is something in the very idea of good, or evil, which naturally excites desire or aversion." But, he wrote, before the acquisition of self-concepts, people are indifferent about whether those who may be pleased or suffer are themselves or others: "a child first distinctly wills or pursues his own good," he said, “not because it is his but because it is good." As a consequence, he claimed, "what is personal or selfish in our affections" is due to "time and habit," the rest to "the principle of a disinterested love of good as such, or for it’s own sake, without any regard to personal distinctions.”

Hazlitt asked why, if people connect to the future through imagination, which does not respect the difference between self and other, the force of habit is almost invariably on the side of selfish feelings. His answer involved his trying to account for
the growth of selfish motives in humans by appeal to their acquisition of self-concepts. In his view, when very young children behave selfishly it is not because they like themselves better, but because they know their own wants and pleasures better. In older children and adults, he thought, it is because they have come under the control of their self-concepts, which is something that happens in three stages. First, young children acquire an idea of themselves as beings who are capable of experiencing pleasure and pain. Second, and almost "mechanically" (since physiology insures that children remember only their own pasts) children include their own pasts in their notions of themselves. Finally, imaginatively, they include their own futures.\textsuperscript{30}

In the first half of the eighteenth century, the possibility of a developmental account of the acquisition of self-concepts that Locke may have seen dimly was invisible to most of his readers. As commonsensical as the idea of this possibility may seem to us today, it did not begin to emerge in the views of eighteenth century thinkers until mid-century. Hartley had formulated a developmental, associational account of the mind, but he focused on the development of the passions and did not consider the acquisition of self-concepts. Jean Jacques Rousseau (1712-1778), especially in \textit{Emile}, was sensitive to developmental concerns, but not particularly with respect to the acquisition of self-concepts. Reid, late in the century, had a developmental psychology, but because of his commitment to the immateriality of the soul and the reflexive nature of consciousness, he may actually have made an exception in the case of the idea of self. Priestley, largely under the influence of Hartley, accepted the possibility of a developmental account of the acquisition of self-concepts, but did not elaborate.

Hazlitt thought that to progress through all three of the development stages that he distinguished in the acquisition of self-concepts, a child has to differentiate its own mental activities from those of others. In his view, this involves "perceiving that you are and what you are from the immediate reflection of the mind on it's own operations, sensations or ideas." He then raised the question of how a child's formation of self-concepts is related to its development of empathy and sympathy. No one previously had asked this question.
In Hume’s emotional contagion model of human sympathy, humans infer from external behavior, facial expressions, and the like that others are in some particular mental state. Then, the resulting idea that humans form of another’s state becomes converted in their own minds into an impression, so that now they too are in the same state, though perhaps less vivaciously. In explaining how this conversion from idea to impression occurs, Hume appealed to the idea’s “proximity” in one’s mind to the impression one has of oneself, which he said is “so lively” that “it is not possible to imagine that any thing can in this particular go beyond it.” But, then, he added not a word of explanation about how people acquire their super-lively self-impressions.

Two decades later, Adam Smith gave an unusually thorough account of the role, in sympathy, of shifts from one’s own to another’s point of view. Yet Smith never attempted to explain how people acquire their ideas of the distinction between self and other. Aside from the applications of his ideas to ethical theory, Smith’s gaze was fixed on the importance of point of view as a feature of adult minds, not on the psychogenetics of point of view in our mental development. In explaining how sympathy is possible, it did not occur to him to explain how the conceptual apparatus that makes it possible came to be acquired in the first place.

Hazlitt speculated that young children imaginatively include only their own futures and not the futures of others in their ideas of self because the "greater liveliness and force" with which they can enter into their future feelings “in a manner identifies them” with those feelings. He added that once the notion of one’s own personal identity is formed, “the mind makes use of it to strengthen its habitual propensity, by giving to personal motives a reality and absolute truth which they can never have.” This happens, he thought, because "we have an indistinct idea of extended consciousness and a community of feelings as essential to the same thinking being," as a consequence of which we assume that whatever "interests [us] at one time must interest [us] or be capable of interesting [us] at other times.”

Hazlitt claimed that a bias in favor of ourselves in the future could never "have gained the assent of thinking men" but for "the force" with which a future-oriented idea
of self "habitually clings to the mind of every man, binding it as with a spell, deadening its discriminating powers, and spreading the confused associations which belong only to past and present impressions over the whole of our imaginary existence." However, whereas a host of previous thinkers—Descartes, Locke, Berkeley, Butler, and others—thought that people have intuitive knowledge of their own identities, Hazlitt rejected as "wild and absurd" the idea that people have any sort of identity that could be available to be intuited. We have been misled, he claimed, by language: by "a mere play of words." In his view, both children and adults fail to look beyond the common idioms of personal identity and as a consequence routinely mistake linguistic fictions for metaphysical realities. To say that someone has a "general interest" in whatever concerns his own future welfare "is no more," he insisted, "than affirming that [he] shall have an interest in that welfare, or that [he is] nominally and in certain other respects the same being who will hereafter have a real interest in it." No amount of mere telling "me that I have the same interest in my future sensations as if they were present, because I am the same individual," he claimed, can bridge the gulf between the "real" mechanical connections I have to myself in the past and present and the merely verbal and imaginary connections that I have to myself in the future.\(^33\)

**Toward a Science of Human Nature**

When Locke published his *Essay*, he was eager to launch a science of human nature. Four decades later, when Hume published *A Treatise of Human Nature* (1739), he assumed that a science of human nature had not only been launched, but had already taken a wrong turn.\(^34\) He was intent on setting things right, which he thought involved having the science of human nature assume its rightful position among the sciences. In his view, that position was at the *foundation* of a mighty edifice of human knowledge. Whereas today we tend to think of physics as the most fundamental science, Hume thought of the science of human nature as the most fundamental since only it would build an account based on experience (rather than things), which for Hume was our ultimate source both of evidence and meaning. “There is no question of importance,” Hume said, “whose decision is not comprised in the science of man; and there is none,
which can be decided with any certainty, before we become acquainted with that science.” In explaining “the principles of human nature,” he continued, “we in effect propose a complete system of the sciences, built on a foundation almost entirely new, and the only one upon which they can stand with any security.”

How, then, to proceed? The first step, Hume thought, was to reveal the basis on which any genuine science of human nature must be built. That, he said, is “experience and observation,” by which he meant the ultimate impressions (what twentieth century philosophical empiricists would call sense data) on the basis of which all of a human’s more complex ideas (concepts) would have to be wholly constructed. As it happened, however, for psychology to find its feet as a science it had to abandon such epistemological and metaphysical pretensions. Its practitioners had to realize that it was not their job, qua psychologists, to get to the absolute bottom of things. Happily, that task could be left to philosophers. Rather, it was their job, as psychologists, to explain human behavior. To do that, they had to take certain things for granted that in a more philosophical frame of mind could be seen to be deeply questionable. This was the approach that Hartley followed and that Hume’s friend and confidant Adam Smith followed in his early work on “the moral sentiments” (mainly human sympathy). It is also the approach that Hume himself often followed, in spite of his methodological manifesto.

This contrast between an austere empirical philosophical approach and a more realistic scientific approach is especially poignant in Hume’s account of self and personal identity. In Book I of the Treatise, the heart of his account is his argument that belief in a substantial, persisting self is an illusion. More generally, he was intent on showing that belief in the persistence of anything is an illusion. This is what today we would call philosophy, rather than psychology. However, in the remainder of Book I, Hume addressed the task of explaining why people are so susceptible to the illusion of self. And in Book II he explained how certain dynamic mentalistic systems in which we represent ourselves and others actually work, such as those systems in us that generate sympathetic responses to others. In these more psychological projects, Hume often seems to have taken for granted things that in Book I he had subjected to
withering skeptical criticism.

In Hume’s view, since all ideas arise from impressions and there is no impression of a “simple and continu’d” self, there is no idea of such a self. This critique of traditional views led him to formulate his alternative “bundle” conception of the self and also to compare the mind to a kind of theatre in which none of the actors—the “perceptions [that] successively make their appearance”—is either “simple” at a time or, strictly speaking, identical over time. Hence, none is the traditional self. Beyond that, Hume claimed, humans do not even have minds, except as fictional constructions. Thus, in his view, a crucial respect in which minds are not analogous to real theatres is that there is no site for the mental performance, or at least none of which we have knowledge; rather, there “are the successive perceptions only, that constitute the mind; nor have we the most distant notion of the place, where these scenes are represented, or of the materials, of which it is compos’d.”

With these philosophical preliminaries out of the way, Hume turned to the psychological task of explaining how objects that are constantly changing, including the materials out of which we ourselves are constructed, nevertheless seem to persist. His answer, in one word, was: resemblance. When successive perceptions resemble each other, he said, it is easy to imagine that the first simply persists. In fact, “our propensity to this mistake” is so ubiquitous and strong “that we fall into it before we are aware.” And even when we become aware of our error “we cannot long sustain our philosophy, or take off this biass from the imagination.”

Hume may have thought that a crucial difference between Locke and himself on the question of personal identity is that whereas Locke thought that there is a fact of the matter about whether a person persists, Hume thought that there is a fact of the matter only about the circumstances under which the illusion of persistence is nourished. In his capacity as a psychologist, Hume tried to explain what those circumstances were. But he did not stop there. As soon as he moved on to the largely psychological concerns that dominate Book II of the *Treatise*, he became deeply involved in what today we would call social psychology of the self. He, thus, completed a transition from skeptical
philosophy to the most general sorts of associational issues, and then to specific psychological hypotheses about how self-representations function in our mental economy, as for instance in his explanation of how sympathy works.

Subsequently Reid, who in spite of his own empirical investigations was a virulent opponent of empiricist epistemology, criticized Hume for denying that there is anything more to mind than a “succession of related ideas and impressions, of which we have an intimate memory and consciousness.” Reid asked “to be farther instructed, whether the impressions remember and are conscious of the ideas, or the ideas remember and are conscious of the impressions, or if both remember and are conscious of both? and whether the ideas remember those that come after them, as well as those that were before them?” His point was that since ideas and impressions are passive, they cannot do anything, whereas Hume implied that the “succession of ideas and impressions not only remembers and is conscious” but also “judges, reasons, affirms, denies,” even “eats and drinks, and is sometimes merry and sometimes sad.” Reid concluded, “If these things can be ascribed to a succession of ideas and impressions in a consistency of common sense, I should be very glad to know what is nonsense.” In Reid’s view, if in accounting for the mind substance were to have no place, then agency would have no place either. Since Reid thought it would be absurd to deny agency, substance had to be retained. But what Reid might instead have concluded from his criticism is that in order to conduct a science of human nature one has to make realistic assumptions about the mind.

**Associationism**

The theory that complex ideas in the human mind are constructed out of simple components and that the succession in the mind of (mostly) complex ideas can be explained by appeal to their similarity with each other and their repeated juxtaposition had been around since classical times. However, this theory not only resurfaced in the modern era, but became a preoccupation of empiricists. In the seventeenth century, Hobbes used it to explain the succession and coherence of ideas:

The *cause* of the *coherence* or consequence of one conception to another, is
their first coherence or consequence at that time when they are produced by sense; as for example, from St. Andrew the mind runneth to St. Peter, because their names are read together; from St. Peter to a stone, for the same cause; from stone to foundation, because we see them together; and for the same cause from foundation to church, and from church to people, . . . [and thus] the mind may run almost from anything to anything.⁴⁰

In the eighteenth century, such appeals to association acquired renewed vitality, due primarily to the influence of Locke, Hume, and Hartley, all of whom gave association a central role in their accounts of experiential phenomena. But neither Locke nor Hume appealed to association to speculate on the physiological underpinnings of empirical phenomena. That task was left to Hartley.

Philosophically Hartley was a dualist, but methodologically he was a materialist. Differing in this respect from Collins before him and Priestley after, Hartley believed that “man consists of two parts, body and mind,” where the mind “is that substance, agent, principle, &c. to which we refer the sensation, ideas, pleasures, pains, and voluntary motions.” But Hartley accepted Locke’s concession that it is possible, for all we know, that matter thinks. And he doubted that either problems with materialism or prescientific intuitions we may have about the so-called unity of consciousness could be used to prove that the soul is immaterial, confessing that “it is difficult to know [even] what is meant by the Unity of Consciousness.” He claimed that there is a problem with materialism in that “that Matter and Motion, however subtly divided, or reasoned upon, yield nothing more than Matter and Motion still.” But it was, he said, “foreign to [his] Purpose” to pursue the issue.

In addition to being a dualist, Hartley was a theist. But he never allowed his metaphysical and theological views to interfere with his attempt to establish a deterministic associationist psychology. Inspired by Newton’s suggestion in *Principia Mathematica* that vibrations of corpuscles of light might cause vibrations in the retina of the eye, which would then be transmitted to the brain where they would produce the sensation of sight, and by some intimations of associationism in John Gay’s (1699-
Dissertation Concerning the Fundamental Principles of Virtue or Morality (1731), Hartley proposed a “physics of the soul” in which physical vibrations in the brain, spinal cord, and nerves are the basis of all sensations, ideas, and motions of men and animals. In his view, the “higher” the mental function–images and ideas, for instance, are higher than sensations–the more delicate the vibrations with which it is associated. And when mental functions are similar, as in the case of images and ideas that faithfully replicate sensations, it is due to a correspondence in the vibrations.

All learning, Hartley claimed, including that involved in perception, memory, imagination, emotion, and language, is the consequence of repetitive juxtapositions of corpuscular vibrations and mental associations that produce habits in accordance with a pleasure-pain principle, a view that he illustrated especially by appeal to the study of how children learn languages. Hartley thereby produced the the first truly general account of human and animal psychology, which was an association based, mechanistic, deterministic, physiological psychology.

In France, the physician Julien Offray de la Mettrie (1709-1751), in his Natural History of the Soul (1745) and his Man a Machine (1748) developed Hartley’s approach by arguing that human beings are merely physiological machines. Subsequently, Condillac laid the groundwork for a association-based psychophysiological account of human nature that became influential on the continent in the nineteenth century. Meanwhile, in Britain, Priestley encouraged the acceptance of Hartley’s ideas in his Hartley’s Theory of the Human Mind on the Principle of Association of Ideas (1775). Priestley thought that the sentient and thinking principle in man must be "a property of the nervous system or rather of the brain," insisting that it is scientifically useless to postulate an immaterial substance to account for any aspect of human mentality or behavior. Priestley saw the differences between humans and other animals as differences of degree, rather than kind, and held that human infants begin like other animals and only gradually learn adult human modes of thinking, including the ability to conceptualize themselves.

In British philosophy, where empiricism still held sway in the nineteenth century,
interest in associationism gathered strength. Thomas Brown (1778-1812), in his three volume, *Lectures on the Philosophy of the Human Mind* (1820), importantly elaborated associationist theory by distinguishing primary and secondary laws of suggestion (his word for *association*). And James Mill (1773-1836), in his *Analysis of the Human Mind* (1829), sketched a general view of the mind in which it was little more than mere machinery for the associate process, a view that many psychologists came to regard as an important advance on Hartley’s approach.

John Stuart Mill, James Mill’s son, became an enthusiastic follower of the Positivism of Auguste Comte (1798-1857), but criticized Comte’s negative attitude toward psychology: Comte “rejects totally, as an invalid process, psychological observation properly so called, or in other words, internal consciousness, at least as regards our intellectual operations.”43 To fill this gap, Mill made detailed comments on and refinements to his father’s thoughts, ultimately arguing for his own associationist system of “mental chemistry.” However, J. S. Mill’s own contributions to psychology, while extremely attentive to internal consciousness, were primarily epistemological. Like Hume, he thought that his own phenomenalism, which he called *the psychological theory*, was a kind of foundational psychology. In Mill’s view, material objects are “permanent possibilities of sensation,” and other minds are inferred to exist based on an analogy with one’s own case, which he presumed one knows directly. He claimed that like objects in the external world, minds too are just actual and possible sensations. Subsequent psychologists tended to regard his psychology as too philosophical to be responsive to their own interests.

Meanwhile Alexander Bain (1818-1903) revived and greatly developed Hartley’s interest in a physiological approach to the understanding of human mentality. In *The Senses and the Intellect* (1855) and *The Emotions and the Will* (1859), Bain drew upon Hartley and others to work out a sensory-motor associationism that marked a turning point in the history of associationist psychology. Before his work associationists like Hume and J. S. Mill were committed to experience as the primary source of knowledge. Bain, in a more realist mode, accepted movement and social interaction as primary,
which he then used to explain higher mental functions, including self-attributions. He claimed, for instance, that when attention is turned inward upon oneself as a personality “we are putting forth towards ourselves the kind of exercise that properly accompanies our contemplation of other persons.”

Bain’s more sophisticated psychophysiology was distinctive, first, for its realism, in that he began by assuming the existence of the physical world, including as items in it other people and himself; second, by the primacy he gave to social observation, in that we first make judgments about others, and only later think of ourselves as one “other” among many; and, third, by his suggestion that this progression from others to self not only explains the origin of the notion of self, but also our ability to feel toward the self emotions that originally we felt toward others. Ultimately J. S. Mill would praise Bain’s account as the highest point yet reached by the empiricist tradition.

Concurrent with such philosophical and psychological developments there was in the nineteenth century a growing spirit of naturalized science, typified by the work of Charles Darwin (1809-1882), but independently including inquiry into the development of self concepts and the physiology of the brain. In 1855, the same year in which Bain published *The Senses and the Intellect*, Herbert Spencer (1820-1903) published *The Principles of Psychology*, which grounded psychology in evolutionary biology. Subsequently William James would build on both of these contributions.

James followed Bain, who had defined belief as a rule or habit of action, and Charles Sanders Pierce (1839-1914), who had claimed that the point of theory is not to represent reality, but to enable us to act more effectively, in turning partly away from empiricism toward what came to be known as Pragmatism. In some ways, James was the last philosopher/psychologist and arguably the last psychologist of importance in whom a sort of empiricism that could be traced back directly to Locke and Hume still resonated strongly. Increasingly, in the twentieth century, philosophy and psychology tended to go their separate ways. Throughout the first half of the century empiricism, particularly in its incarnation in epistemology, continued to be a potent force in philosophy, but was much less so in psychology. There the influence of empiricism
tended to be supplanted by a newfound preoccupation with behavior and with the social
dimensions of mental development.

Endnotes

*Throughout the present paper I have drawn freely from two books that I co-authored
with John Barresi: Naturalization of the Soul: Personal Identity in the Eighteenth Century
(London: Routledge, 2000) and The Rise and Fall of Soul and Self: An Intellectual
History of Personal Identity (New York: Columbia University Press, 2006). I am grateful
to John for allowing me to draw upon work that is as much his as mine. I am also
grateful to him, to Marya Schechtman, and Michael Mathias for comments on an earlier
draft.

http://www.rep.routledge.com/article/W028..

http://www.rep.routledge.com/article/U014SECT1,


5. See Marya Schechtman’s contribution to the present volume.
v.3, pp. 720-913.
7. Ibid., p. 750.
8. Ibid., pp. 769-73.
10. Ibid., p. 790.
11. Ibid., pp. 809, 870.
15. For a competing view, see Michael Ayers, Locke, 2 vols.(London: Routledge, 1991),
v. 2, pp. 266-67.

17. Ibid., pp. 87-8.
18. Ibid., pp. 328, 331-2.
20. Ibid., pp. 322, 325.

22. Ibid., pp. 174-5.

25. Ibid., p. 373.
26. Ibid., p. 400.
27. Ibid., pp. 400-1.

29. Ibid., pp. 33-4.
30. Ibid., pp. 34-5.
31. Hume, op. cit., p. 317
32. Hazlitt, op. cit., pp. 10-1, 140.
33. Ibid., pp. 6, 10-1, 27-9.

35. Ibid., Introduction.

37. In and of itself, Hume suggested, our supposing that objects persist is not so bad. But “in order to justify to ourselves this absurdity,” we make up a story, often one in which the principle character is the notion of substance; that is, we invent the fictions of “soul, and self, and substance to disguise the variation” in our perceptions. When, as in
the case of “plants and vegetables,” we cannot fool ourselves into believing that the persistence of an underlying substance accounts for the persistence of the organism, we invent an equally “unknown and mysterious” surrogate—presumably, “life”—to connect the successive and different perceptions, *ibid.*, pp. 254-55.


41. Hartley’s formulation of associationism states that “Any sensation A, B, C, etc. by being associated with one another a sufficient number of times, gets such a power over the corresponding ideas a, b, c, etc. that one of the sensations A, when impressed alone shall be able to excite in the mind b, c, etc. the ideas of the rest.” In David Hartley, *Observations on Man* (1749), reprinted as two volumes in one, with an introduction by T. L. Huguelet (Gainesville, Florida: Scholars Facsimiles and Reprints), vol. I, prop. 10.

42. Priestley’s rootedness in science, together with the matter of factness of his materialistic approach and his unproblematic commitment to realism, differed radically from the epistemologically oriented versions of empiricism championed by Locke, Berkeley, and Hume. Because of it Priestley did not think that his style of empiricist epistemology led to skepticism about the external world, as Reid had claimed, or indeed to skepticism about anything, and he more cleanly separated philosophy from science than Hume, in particular, had been able to do. Priestley, *Disquisitions Relating to Matter and Spirit and the Doctrine of Philosophical Necessity Illustrated*, 1777, reprinted at New York: Garland Publishing Inc., 1976, p. 163
