

Hume's Functionalistic Theory of the Self

A Dissertation Presented

by

Sardar Hosseini

**Submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfilment of the requirements for a doctoral degree in
Philosophy**

Department of Philosophy

Arts

University of Ottawa

To the memory of my father and my first teacher,
Asadullah Hosseini

Abstract

The main claim of this dissertation is that Hume's theory of the self can be interpreted in terms of a causal or functional theory of mind. It is a thesis about Hume's identification of mental particulars—impressions and ideas—in terms of the kind of roles that each plays in the cognitive system that it is a member of. The true Humean idea of the human mind is to understand it as a system of different mental states and processes, which are linked together by the relation of cause and effect. Functionalism as such can be construed as both teleo-functionalism and psycho-functionalism. The former is rooted in his teleological characterization of the mind according to which the bundle of perceptions persists over time by maintaining functional continuity, whereas the main source of Hume's psycho-functionalism lies in his Representational Theory of Mind.

Hume, however, Hume expresses his strong dissatisfaction with his earlier treatment of the topic, and confesses that he now finds an inconsistency in his original account. He does not make clear in his recantation what he finds problematic in his earlier account. And although more than a dozen interpretations have been suggested, no consensus as to what Hume's worry is has emerged. I claim that Hume's functionalism, as presented in the main body of the *Treatise*, stores a problem for him and when he arrives at the Appendix he realises the problem and confesses that he is unable to resolve it. The problem that leads to the inconsistency has two main possible sources: First, the principles of constancy and coherence may successfully account for the arising belief in the idea of the continued and distinct existence of external objects and the idea of personal identity, but they fail to explain our belief in other minds (selves). Second, Hume's functionalism is circular because it presupposes personal identity. The central idea is that if Hume is right to say that something like functional continuity would suffice for persons to persist through time, then he must show that we can have a *complete* account of how one's mental states produce the idea of a persisting self without making assumption about the identity condition of their subject or bearer. And of course, psycho-functionalism, including Hume's, identifies a mental state in terms of its functional relations to other mental states that are the states of *the same person*. This is straightforwardly circular.

ACKNOWLEDGEMENTS

In acknowledging those who have made it possible for me to complete this dissertation, I must express deepest gratitude to my supervisor, Professor David Raynor, who first introduced me to David Hume and guided me through this work. He offered continuous encouragement and innumerable stimulating ideas. I also thank the members of my dissertation committee: Professor Andrew Sneddon, Professor Graeme Hunter, and Professor Patrice Philie. Finally, I thank my wife, Zahra, and my son, Erfan, for their patience and encouragements over the years. I am grateful to them both.

Table of Contents

Acknowledgments iii

List of Abbreviations v

Introduction 1

I. Hume's Functionalism: A First Approximation	19
I.1: The Lockean Background to Hume's Functionalism	23
I.2: Hume's Characterization of the Self as a Fiction	35
I.3: The Mind as a System of Causal Connection	44
I.4: Hume on the Nature of the Mind	55
II. Hume's Computational Theory of Mind	78
II.1: Hume's Mechanics of the Mind	80
II.2: The Imagination and Causal Inference	91
II.3: Hume on the Architecture of the Mind	97
II.4: Hume's Representation Theory of Mind	106
• Hume on Mental Representation	109
• Hume on Computation	113
• Hume on Meaning	120
III. Hume's Psychology of Identity Ascriptions	124
III.1: The Continued and Distinct Existence of External Objects	131
III.2: The Distinction between Constancy and Coherence	138
III.3: Hume's Principle of Constancy	163
III.4: Hume's Principle of Coherence	175
IV. Hume's Functionalism and Personal Identity	186
IV.1: Hume on Character Traits, Self-Concerns, and Personal Identity	189
IV.2: Functionalism and the Psychological Continuity Criterion of Personal Identity	196
IV.3: Functionalism Tells us Nothing about Personal Identity	204
IV.4: Hume's Reification of Perceptions and the Impersonal Description of the Self	214
IV.5: Tension in Hume's Functionalistic Theory of Personal Identity	219
V. Hume's Second Thoughts about Personal Identity	225
V.1: Interpretations of Hume's Second Thoughts	230
V.2: Hume's Criterion of Personal Identity: Constancy and Coherence	244
V.3: The Failure of Coherence and Constancy and the Belief in Other Minds	251

Notes and References 273

Bibliography 288

List of Abbreviations

- THN David Hume, *A Treatise of Human Nature*. References are given by page number to the edition of L. A. Selby-Bigge revised by P. H. Nidditch (Oxford, 1978).
- EHU David Hume, *An Enquiry Concerning Human Understanding*. References are given by page number to the edition of L. A. Selby-Bigge revised by P. H. Nidditch (Oxford, 1975).
- EPM David Hume, *An Enquiry Concerning the Principles of Morals*. References are given by page number to the edition of L. A. Selby-Bigge revised by P. H. Nidditch (Oxford, 1975).
- EST David Hume, *Of the Standard of Taste and Other Essays*. References are given by page number to the edition of The Library of the Liberal Arts edited by J. W. Lenz (Indianapolis, 1965).
- Abs David Hume, *An Abstract of a Book Lately Published, Entitled, A Treatise Of Human Nature, &c.* References are given by paragraph numbers. Cited From the Clarendon Edition of *A Treatise of Human Nature*, Oxford: Clarendon Press 2007.
- EMPL David, Hume, *Essays: Moral, Political, and Literary*, edited by Eugene F. Miller, Library of Classics: Indianapolis, Ind., 1987. References are given by page number.
- Essay John Locke, *An Essay Concerning Human Understanding*. References are given by page number to the edition of P. H. Nidditch (Oxford, 1975).
- AT C. Adam, P. Tannery (eds.), *CEuvres de Descartes*, revised ed., 12 vols. (Paris: Vrin/CNRS, 1964-76.
- CSM J. G. Cottingham, R. Stoothoff, and D. Murdoch (eds.), *The Philosophical Writings of Descartes*, 2 vols., Cambridge University Press 1985.
- CSMK J. G. Cottingham, R. Stoothoff, D. Murdoch, and A. Kenny (eds.), *The Philosophical Writings of Descartes: The Correspondence*, vol. III, Cambridge University Press 1991.

Introduction

The seventeenth century philosophers were still thinking about the self in the mode of the Ancient and Scholastic traditions. They still breathed and thought within the scheme of that tradition according to which the self was conceived as the soul, an immaterial substance. Descartes' substance dualism is said to be the most vigorous and spirited defence of such a view of the self at the very beginning of modern philosophy. He described the body as sheer mechanical processes which are subject to the laws of physics, whereas the unextended mind is situated completely outside the schema of the body and, thus, is not subject to the laws of physics. However, that view of the self was undergoing a massive transformation. Martin and Barresi ⁽¹⁾ have called the seventeenth century "a time of momentous and soul-shattering intellectual transformation." The natural philosophers of the century had already introduced the *experimental method* of reasoning into the sciences and had made great advances. Newton played a central role in showing that there could be a natural philosophy of the external world. On the other hand, the moral philosophers turned away from the dogmatic/scholastic conception of the soul as a mysterious immaterial substance and favoured to investigate the mind as a natural phenomenon. By the end of the eighteenth century the immaterial soul was replaced by the self as material mind, a dynamic natural system subject to general laws of natural philosophy. The new natural philosophy turned away from its predecessors' dogmatic assertions about the essence of the soul, and instead, it started investigating the mind's activities and states under the light of human experiences. So, the immaterial soul transformed into an empirical self, or mind. This transformation posited new possibilities about the nature of the self and the way we understand it as well as new challenges. Of course, it elevated our understanding of ourselves and deepened our perspective on how to explain our conscious mental activities, but it also caused serious philosophical controversies and existential terror. One major philosophical problem that this new understanding of ourselves caused was whether the self persists from moment to moment, or survives over time. This new emerging imaginary, or "fictitious," to use Hume's terminology, notion of the self was powerful enough to cause reaction among philosophers, which in turn, backfired and attracted some important eighteenth century philosophical figures to the idea of the mind as an immaterial substance. The outcome of

this soul-shattering transformation was that by the end of the eighteenth century, the mystery of the immaterial soul was replaced by the new mystery of the material self.

Hume plays a major role in this soul-shattering intellectual transformation. He describes his own philosophy as an attempt to introduce the Newtonian Experimental Method of Reasoning into moral subjects. This scientific method of reasoning that is founded on experience and observation is said to be “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” (THN, xvi). Hume’s theory of the self is radically new and controversial at the time of its appearance. There is a strong sense that Hume has pushed that soul-shattering transformation to its edge. One of Hume’s most important philosophical achievements is the demystification of the immaterial soul through the careful analysis of the psychological mechanism by which our mental activities are described.

This dissertation is about Hume’s account of the self, or to use the more professional term, personal identity. In the years preceding the publication of Hume’s *Treatise of Human Nature* (THN), the question of personal identity had already “become so great a question in philosophy, especially of late years in *England*” (THN, 259). Most notably, Locke had discussed the question of personal identity before Hume, and gave the first formulation of the problem in the chapter “Of Identity and Diversity” in his *Essay Concerning Human Understanding*. Shaftesbury in his *Characteristics* reacted to Locke’s memory-consciousness based account of the self and rejected it because he thought that it was insufficient to be ethically decisive (2). Moreover, George Berkeley in his *Alciphron* and Joseph Butler in his *Analogy of Religion* had discussed the question of personal identity. Furthermore, Clarke and Collins produced a six-part written debate concerning the materiality and immateriality of the soul. Ephraim Chambers, Isaac Watts and Vincent Perronet had debated personal identity as well (3). Outside the English philosophical tradition, Leibniz is the most important philosopher who has discussed personal identity in his *Discourse on Metaphysics* and in his *New Essays on Human Understanding*, in which in the second work he reacted to Locke’s account of the self. The debate about personal identity continued after the publication of Hume’s *Treatise*. The most important debate is Thomas Reid’s theory of personal identity in which he

rejected both Locke's and Hume's accounts of the self.

There are reasons for the re-examination of Hume's account of the self. Historical reasons aside, Hume's account of the self was central to his philosophy, particularly to his systematic teachings in the *Treatise*. It has direct bearings on his account of epistemology in Book I and his account of the passions in Book II and his account of morality in Book III of the *Treatise*. More importantly, the account of the self that he defends in Part IV of the Book I, in particular that of the section of "Of Personal Identity," is closely related to his so-called science of *human nature*. And the science of "Human Nature is the only science of man" (THN, 273). Hume's new science of human nature, which is based on the experimental method of reasoning, provides a complete system of sciences upon which other sciences depend, and it is in this science alone that we "can expect assurance and conviction" (THN, 6, 273). In Hume's view, understanding the mind is central to the understanding of all of his philosophy. Finally, Hume's account of the self may have significant relevance to contemporary debate on personal identity. It is true that there are not many philosophers today who are prepared to defend the Humean 'bundle theory' of the mind, but Hume's functionalistic account of the self, which I will defend in this work, is directly related to contemporary discussions on personal identity. In particular, contemporary philosophers of mind largely believe that theories about the traditional mind/body problem, including functionalism, have no ramifications for the question of personal identity. Sydney Shoemaker has challenged this common view and said that functionalism tells us important things about personal identity. It is partly my intention here to show that Hume may also contribute significantly to our understanding of this problem. Hume seems to think that functionalism implies a theory of personal identity. So in this sense a functionalistic interpretation of Hume's theory of the self might open some interesting prospects for the contemporary theories of the self.

The literature on Hume's theory of the self has expanded over the years. In particular, the debate and the controversy about the famous section of "Of Personal Identity" (THN, 251-263) and Hume's self-doubt in the Appendix (THN, 633-636) have mushroomed in the last three decades. In the section of "Of Personal Identity" Hume sets out his so-called 'bundle theory of the self.' However, in the Appendix he talks about an

inconsistency without ever exactly identifying the source of the inconsistency. The major preoccupation of recent works on Hume's theory of personal identity has been to understand the relation between these two conflicting views that he takes in the *Treatise*. Hume's scholars have almost exclusively focused on finding the source of the inconsistency, and have produced a dozen interpretations, but there are not as many works treating Hume's theory of the self in a broader context. Antony Pitson's *Hume's Philosophy of the Self*, which gives a comprehensive treatment of Hume's account of the self, is an exception. There have been scattered suggestions for a functionalistic construal of Hume's theory of personal identity here and there, but there is no all-encompassing treatment of it available at the moment. Norman Kemp Smith, for example, noted that Hume's comparison of personal identity with the identity which we ascribe to vegetables and animal bodies indicates that personal identity consists in "identity of function" (4). In this respect, personal and object identity is explained in terms of the function that the parts of each play in conspiring to a *common end* or purpose, and that this end endures total and radical transformation (THN, 257). John Hougeland has also named Hume as one of the sages of modern Artificial Intelligence theory (AI) (5). He claims that Hume is the first modern philosopher who undertakes a "mental mechanics" which does not invoke a homunculus. Hougeland ascribes a non-homuncular functionalism to Hume, which is greatly influenced by Newton's mechanical philosophy of nature. Furthermore, and most recently, Jerry Fodor, one of the great fathers of functionalism has attributed a Representational Theory of Mind (RTM) to Hume. He claims that "Hume's *Treatise* is the foundational document of cognitive science: it made explicit, for the first time, the project of constructing an empirical psychology on the basis of a representational theory of the mind" (6). Last but not least, John Biro has argued that not only Hume's science of man and the new cognitivism have lots in common, but also Hume's cognitive psychology is "in some respects anticipation of, and superior to, the current theories" (7) of contemporary cognitive science. Daniel Dennett, Don Garrett, Tony Pitson, and Simon Blackburn have also attributed some form of functionalism to Hume (8).

Although these philosophers have recognized and thus argued for a causal/functional Humean theory of the self (mind), the literature lacks a comprehensive treatment of Hume's involvement with functionalism. Kemp Smith's remark does not go

beyond a brief statement of Hume's teleological functionalism. Fodor's recent book, *Hume Variations*, contributes significantly to the causal/functional interpretation of Hume's philosophy of mind, and thus remedies the shortcoming to some extent, but it remains a limited defence and exposition of Hume's Theory of Ideas, or Hume's Representational Theory of the Mind (RTM). Fodor, for example, neither discusses Hume's psychology of identity ascriptions nor does he deal with the conditions of identity (criterion of personal identity) and its continuity over time. Nor he discusses Hume's teleo-functionalism. Biro's works are centered on Hume's cognitivism and the question of the architecture of the mind. He, like Fodor, does not discuss Hume's psychology of identity ascriptions and/or the conditions of personal identity.

This dissertation attempts to remedy defects of this kind. First, it highlights the significance of Hume's much neglected teleological functionalism and its place in his overall causal/functional theory of mind. Second, it tries to show how Hume's psychology of identity ascriptions, particularly the mechanism of constancy and coherence, proves to be an essential case study for his views about the sub-doxastic nature of the mind. Third, it will argue that the mechanism of constancy and coherence function as the Criterion (conditions) of Personal Identity. My aim is to deliver an interpretation of Hume's theory of the self, which to some extent addresses these concerns.

I need to point out three things at the very outset. First, it is not my intention to present a historical treatment of Hume's theory of the self. This work does not aim to investigate the historical roots of Hume's functionalistic theory of the self in his predecessors; rather, my chief concern is to show how Hume develops his account of the self in response to, and to a certain extent in continuation of, Locke's general theory of object identity. In particular, I want to show that in the tradition in which Hume was writing, deriving from Locke, the problem of personal identity was understood as that of giving an account of what *constitutes* personal identity. In contemporary philosophy of mind this issue is called the Criterion (condition) of Personal Identity. In Chapter I, I have briefly discussed Locke's account of the criterion of personal identity and its impact on Hume's theory of the self. This move is indispensably historical, however.

Second, I shall clearly state the main thesis that I am going to defend in this

project. The central idea of this dissertation is that Hume's theory of the self can be interpreted in causal and/or functional terms. In a very broad sense, functionalism characterizes a mental state like a thought, desire, or pain solely in terms of its function, or the role that it plays in the cognitive system that it is a member of; a mental state is identified in terms of what it does rather than by what it is made of. In other words, functionalism takes the nature or identity of a mental state to be determined by its causal relations to sensory inputs, other mental states, and behavioural outputs. The claim that Hume's theory of mind can be construed on causal and/or functional grounds, indeed, is a thesis about Hume's identification of mental particulars (impressions or ideas, or generally perceptions) in terms of the kind of roles that each plays in our cognitive system. Hume conceives the mind as a compound, or collection, or more accurately a 'bundle' of perceptions in which each one of these perceptions is ontologically independent and is capable of existing separately. However, these individual or particular perceptions are connected by the existence of certain natural relations, namely, resemblance, causation and contiguity. For Hume, the causal relation plays the central role in this interconnected system of perceptions. I argue that Hume's causally interconnected web of perceptions is similar to the central intuition that goes with the functionalist view, which maintains that the nature of mental states is determined in terms of the causal/functional roles that such states, or events, play in the overall operation of the mental system. Each particular perception in this Humean interrelated causal web operates in a similar fashion to the causal/functional role that functionalism assigns to each mental state.

We can discern two familiar and interrelated characteristics of functionalism in Hume's theory of mind. First, each perception, or mental state, plays a causal/functional role in the overall operation of the mind as a compound of ontologically distinct existences. Here, perceptions are determined in terms of the causal/functional roles each one of them is playing within the interconnected causal web of the mental world. This is similar to how contemporary functionalist theory of mind defines the nature of mental states in terms of their functional/causal roles—that is, it is the nature of mental states to be caused in certain ways, and to produce in connection with other mental states certain behaviours or other mental states. With respect to Hume, he seems to be defining

perceptions, which are the building blocks of his mental world, in terms of their causal/functional roles in the overall operations of our minds. Hume's functionalism in this sense may loosely be described as a version of *psych-ofunctionalism* (Chapter II).

Psycho-functionalism like other versions of functionalism—machine-functionalism and analytic functionalism—claims that mental states and processes are defined in terms of the roles that they play in an input-output system producing behaviors. But psycho-functionalism, unlike the other versions of functionalism, characterizes mental states and processes as part of a psychological explanation of human behavior. According to psychofunctionalism, “it is impossible for a system to have beliefs, desires, etc., except insofar as psychological theories true of us are true of it. Psychofunctionalism (so understood) stipulates that Psychofunctional equivalence to us is necessary for mentality” (9). Hence, psycho-functionalism purports to explain the nature of the underlying mental processes and mechanisms involved in perception, memory, and information processing, etc., as part of a cognitive psychological theory. Certainly, equivalency is an important aspect of psycho-functionalism, and Hume's theory of the mind seems to be in accord with it, at least, that is how I understand it. In particular, psycho-functionalism is constituted by the so-called Representational Theory of the Mind (RTM; henceforth). RTM consists of some main theses: (a) mental representations are intentional; (b) psychological explanation is law-like; (c) thinking is computation. The first thesis—intentional realism— says that mental states are about something; or that they are directed towards some objects. In this respect, every mental state is a *representation* of an object. At the heart of such a representation lies a relation, a relation between mental particulars (the relational nature of mental particulars). The second thesis claims that mental causation is the very basic feature of intentional explanation. In this regard, causal relations among mental states that are specified under intentional description instantiates psychological laws. Therefore, psychological explanation is law-like. Hume seems to be committed to both of these theses. Hume takes for granted a Theory of Ideas, according to which a perception is a mental particular and having a perception is a mental event. Likewise, he seems to believe that thinking is a mental process in which some mental events sufficiently cause others. Hence, for Hume, psychological laws are the ones that govern mental causation. Finally, Hume thinks that

mental processes, including thinking, are causal relations among mental representations. Hume seems to be committed to the basic idea of computationalism according to which token mental representations are symbols, and that “tokens of symbols are physical objects with semantic properties” (10). In short, computation is that causal relation between those tokens of symbols. Hume’s associationism could be understood as the prime example of such computational relation. Therefore, for Hume, association is a sort of causal relation among mental symbols. Hume’s psycho-functionalism consists in these main theses of RTM.

Second, the mind as the bundle of perceptions generates certain basic philosophical ideas that, while lacking derivative bases in our sense-perceptions, function as the most basic elements of the mental world. The idea of a persisting self, for example, is a functionally constructed entity that is generated by our psychological belief-determining mechanism, and at the same time it gives to the mind a sense of continuity and connectedness over time. In other words, the idea of the persisting self is a ‘fictional’ idea that comes into existence as a result of the operations within our perceptual systems (or mind). In a similar fashion, the idea of the continued and distinct existence of the external object is constructed by the same psychological belief-determining mechanism, particularly by Hume’s associative principles of the imagination that shapes our experiences of the external reality. Moreover, the same psychological mechanism operates to generate the ideas of necessary connection and substantial forms. All these ideas are ‘fictional’ in the sense that they cannot be derived from any particular sense-impression, though they are produced by the mind. These fictional ideas play certain functional roles in the overall life of the mind. The idea of a self, for example, plays the role of a simple substance persisting over time, which gives a sense of identity to its subject. So the self in this sense is simply a fictional entity that plays a functional role in giving us a sense of identity over time. Functionalism in this sense could be referred to as *teleofunctionalism* (Chapter I).

There is another sense that we could call Hume’s theory of the self teleofunctional. Hume, following Locke, characterizes object identity (particularly the organic objects and the artifacts) in terms of an underlying internal organization striving towards a *common end* or *purpose*. He argues that we ascribe identity to variable and

changing objects, like vegetables, animals, ships, and buildings, even though that any change of considerable part in a mass of matter, strictly speaking, destroys its identity. Hume claims that there is an *artifice*, which induces our imagination “to advance a step farther; and that is, by producing a reference of the parts to each other, and a combination to some *common end* or purpose” (THN, 257). For example, a ship, of which a considerable part has been changed by frequent reparations, is still considered the same. Even the total change of its materials does not hinder us from ascribing an identity to it. “The common end, in which the parts conspire, is the same under all their variations, and affords an easy transition of the imagination from one situation of the body to another” (THN, 257). An oak, grows from a small plant to a large tree, but it is still the same tree; an infant becomes a man, sometime fat, sometimes lean, without any change in his identity; a church that is built from brick, fell to ruin, and is rebuilt from free-stone, but it is still the same; and a river that is totally altered in less than twenty four hours, continue to be the same during several ages. Hume clearly defines the identity of these object (that are, variable and changing) in terms of an underlying teleo-functional principle. More importantly, Hume claims that in order to explain the nature of personal identity the same method of reasoning, which so successfully explained the identity of plants, animals, and ships, and houses, must be used. “The identity, which we ascribe to the mind of man,” Hume says, “is only a fictitious one, and of a like kind with that which we ascribe to vegetables, and animal bodies” (THN, 259). In this sense, for Hume, personal identity, like object identity, must be constituted by the facts about the purpose, or function, of such entities. If the identity of an oak tree, for instance, is constituted by the participation of its parts in order to maintain the same continued life over time, the identity of the human mind is also constituted by the participation of its members (impressions and ideas) to maintain its overall psychological continuity, or a sense of persisting self, over time.

My interpretation of Hume’s functionalistic theory of the self depends upon these two important senses of functionalism: psycho-functionalism and teleofunctionalism. There is a sense that Hume’s teleofunctionalism, however, is not intended to establish a full-dressed functionalistic theory of mind; rather, it wants to provide a *prima facie* case for his causal and/or functional account of the mind. Perhaps the reason that Hume’s

commentators have not discussed his so-called teleo-functionalism has to do with Hume's tendency to treat it as a *prime facie* case for his psycho-functionalism. However, I think that teleofunctionalism plays an important role in Hume's theory of the self, and emitting, or neglecting, such an important aspect of his theory of the self, at the very least, leaves a large portion of the text of the section of "Of personal Identity" unexplained.

The third point that I wanted to draw attention here is about the relation between personal identity, or its psychological continuity over time, and functionalism. I think that this question has significant bearing on Hume's second thoughts about personal identity in the Appendix. I shall briefly explain this. The main question is that if functionalism, which is a metaphysical theory of mind, implies personal identity. It is important to know, whether the metaphysics of mind has implications, if any, on its identity and psychological continuity over time. Many contemporary philosophers of mind think that functionalism does not tell us very much, if anything, about the question of personal identity. Eric T. Olson is the most prominent proponent of such a view. More specifically, it is commonly held that the mind/body problem has no ramifications on the question of personal identity. The traditional mind/body problem, which is a metaphysical problem, is concerned with determining the ontological status of mental states and their relations with the physical states. On the contrary, the question of personal identity, which is understood to be a psychological question about the relation of different stages of a person over time, is concerned with the specification of the continuity conditions of our identity alone. A materialist can subscribe to the psychological continuity criterion of identity, while a proponent of the physical continuity criterion of identity could hold a dualist position about the mind/body problem. Hence, traditionally speaking, philosophers believed that the mind/body problem has no, or very little, relevance for the question of personal identity. Among contemporary philosophers of mind only Sydney Shoemaker has argued that the mind/body problem has certain implications on the question of personal identity. In fact, he thinks that functionalism, which is a theory about the mind/body problem, implies a psychological continuity criterion of personal identity. I hold that Hume makes the same argument long before Shoemaker. He shows us how a causal/functional account of the

mind implies a psychological continuity theory of personal identity. In fact, he shows how deeply intertwined is the question concerning the metaphysics of mind and the question of its persistence through time. It is a well-received idea among many of Hume's scholars that he does bring psychology into play, and it is even suggested that he confuses psychology with metaphysics. It may well be that he confuses psychology with metaphysics, but it will hardly change the fact that the idea of personal identity, which is clearly an idea that is generated by a psychological belief-determining mechanism, somehow is entailed by his causal/functional account of the mind. However, Hume's project fails on a critical moment (more on this in a moment).

On the other hand, Olson has rejected the view that an account of the nature of mental states would imply a theory about our persistence over time. He thinks that Shoemaker's functionalism implies nothing about how mental properties are related to their bearers; the nature of mental properties does not tell us anything about the identity conditions of their bearers. Olson, indeed, denies that there is connection between functionalism, as a theory about the nature of mental states, and the identity conditions of the bearers of those mental states. The reason that he denies such a connection is that he does not think, like Shoemaker, that a person is a sort of substance. Shoemaker distinguishes between two different notions of substance, namely, substance in the sense of the "subject of properties" and substance in the sense of the "parcel of stuff." More specifically, he thinks that substances as subjects of properties have two distinct criteria: (a) the independence criterion of substantiality, and (b) the relatively autonomous self-perpetuator criterion of substantiality⁽¹¹⁾. Shoemaker believes that a person is a substance as an ontological independent existence upon which other entities, modes, affections, and states depend, and that it somewhat perpetuates itself over time. Shoemaker's understanding of these criteria is formed by a famous distinction that W. E. Johnson has made between "immanent" causation and "transeunt" causation⁽¹²⁾. The immanent causation is internal to something career or history that spontaneously causes its later stages, whereas the transeunt causation is the more ordinary sort of causation, when one thing or event brings about another thing or event. Shoemaker conceives the self as an independent substance that autonomously perpetuate itself; it does so by the sort of immanent causation internal to the nature of mind that continually and spontaneously

causes its later stages.

Shoemaker combines this view of the self as a substance, as specified by these criteria, with the functionalist account about the nature of mental states. So, central to Shoemaker's functionalist continuity account of the self is a notion of substance as an ontologically independent experiencing subject. By contrast, Olson holds that to think about a person as an individual substance as the subject of properties presupposes already the existence of personal identity, or some sort of psychological continuity over time. His objection is similar to the old circularity criticism that Butler brought against Locke saying that consciousness presupposes personal identity, and therefore, cannot constitute personal identity. Therefore, for Olson, "the psychological continuity view doesn't follow from functionalism alone, but only from functionalism together with an assumption about personal identity: the assumption that the persistence of thinking beings has something to do with psychology, for instance" (13). This is clearly circular.

Definitely, Olson finds Shoemaker's version of the substantial view of the self troubling, and it seems that his basic insight remains true: that any account of the self linked to substance may lead to circularity. Shoemaker believes that functionalism, a theory about the nature of mental states, implies a sort of psychological continuity criterion of personal identity. However, Olson seems justified in thinking that when Shoemaker ties functionalism with a notion of substance it becomes problematic. There is one way leading us out of this problem: functionalism without making commitment to substance. This seems to be Hume's project, at least as I understand it. Hume sees the same problem in Locke's theory of personal identity. What he does is that he provides a psychological continuity criterion of identity without relating it to any notion of substance. Locke explicitly talked about his theory of object identity (organic objects) in terms of their functional organizations, but he never applied this argument to his theory of personal identity. On the contrary, Hume explicitly characterizes his theory of the self in terms of the sort of functional similarity that exists between organic objects and minds, but he refuses to make any commitment to substance. It might be possible that something like Humean functionalism, which makes no commitment to substance, show us the way out of the difficulty that Olson has brought up against Shoemaker's functionalist continuity account of personal identity. Hume, unlike Locke and Shoemaker, specifies

the identity conditions of our persistence in terms of the mere mechanical operations of our perceptions (mental states) and the sort of causal powers that they impose on each other. Hume does not think that there exists another distinguished and separate entity like substance or an experiencing subject over and above the very causal/functional processes taking place within the mind. Of course, there is no place for Johnson's notion of immanent causation in Hume's philosophy.

Nevertheless, this will cause a problem for Hume, something that may have caused him to think twice about his original theory of the self when he raises self-doubts in the Appendix. As Olson has pointed out, functionalism may tell us how mental states are related to one another, but "it implies nothing about how they must relate to their bearers, except perhaps that they must have bearers. Functionalism would tell us nothing about personal identity over time" (14). Basically, if Hume is right to say that something like causal/functional continuity would suffice for persons to persist through time, then he must show that the characteristic causes and effects of a thing's mental states must always be states of the same person as the one in the original state. What indeed Hume's functionalist theory of the self provides us with is a momentary self that exists in this moment and in the next. The causal/functional theory of the self that he envisages may give us a feeling, or sense, of enduring self over some time, but as soon as our consciousness comes to a halt, for instance, in deep unconscious state of asleep, we lose the entire senses of ourselves. The reason is that Hume rejects the idea that there is a strong causal link between mental properties and their bearers. In fact, he rejects the idea that the self is an ontologically independent being (i.e., an individual substance) upon which mental properties would depend. As a result, his bundle of perceptions will turn out to be loose enough to avoid, for instance, troubles like the problem of *other minds*. Or, as Olson has suggested, the constraints that the causal/functional nature of mental properties put on the persistence conditions of their bearers is loose enough to avoid troubles with the "branching" and "fission" cases (15). The fission cases, or branching, tell us that if both of my hemispheres are transplanted into two different empty heads, the two recipients will each be psychologically continuous with me. It implies that I will be identical with two different beings, and that cannot be. The point here is that the fission cases are similar to the problem of other minds in the sense that they both fail to secure a

robust causal relation between the bearer (the subject of) mental properties and such mental properties themselves. My interpretation of Hume's self-doubt (inconsistency) in the Appendix results from considerations of this nature. I shall briefly explain it now.

In the Appendix, Hume, first, claims that all his hopes vanish when he comes to explain the principles that unite our successive perceptions in our thought or consciousness (THN, 636). Second, he asserts that there are two principles that he "cannot render consistent; nor is it in [his] power to renounce either of them, viz. *that all our distinct perceptions are distinct existences, and that the mind never perceives any real connexion among distinct existences*" (THN, 636). It is widely acknowledged that these two principles are not inconsistent and, indeed, the second principle derives from the first one. Third, Hume immediately points out that if our perceptions either inhere in something simple and individual, or did the mind perceive some real connexion among them, there would be no difficulty in the case (THN, 636). Now, if we put, I think, all these three statements together the result will be quite revealing. What they reveal is that Hume's main worry is about the principle, or principles, that unite our successive perceptions in the bundle. Hume's worry cannot be about the inconsistency of the two aforementioned principles because they are not simply inconsistent. Hume also reveals that if our perceptions inhered in something simple and individual (a substance) or if the mind perceived some real connexion among perceptions, then there would not be any difficulty in the case. But since his original characterization of personal identity does not allow for a notion of substance or real connexion among successive perceptions, it follows that what bothers Hume has to do with the principle, or principles, that unite our successive perceptions in the mind. But what are those principles?

In the section of "Of Personal Identity," Hume explicitly asserts that the principles that unite our successive perceptions in the mind consist of *causation*, *resemblance* and *contiguity* (THN, 260). And in the section of "Of Scepticism with Regard to the Senses" he talks about the imagination and subsequently the principles of *constancy* and *coherence* as the psychological mechanism that *unite* our perceptions in the mind (THN, 195). This shows that, more specifically, for Hume, constancy and coherence function as the principles that unite our successive perceptions in the mind. If this is true, I suggest that, indeed, in the Appendix Hume complains about constancy and

coherence insofar as they fail unite our successive perceptions or to secure a real connection among them. Consequently, according to my interpretation, Hume's problem in the Appendix is about the failure of constancy and coherence. But what made Hume to think that constancy and coherence fail to unite our successive perceptions in our thought or consciousness?

I think that Hume might have thought one of the followings: First, he might have thought that constancy and coherence cannot secure a real connection between successive perceptions. According to Hume's account of constancy and coherence in the section of "Of Scepticism with Regard to the Senses," the imagination and constancy and coherence unite a series of successive perceptions by a process of imaginative postulation, or supplementation, of an unperceived perception in the broken or fragmentary series. This entails that the imaginative postulation of unperceived perception to fill the gaps does not amount to producing real connexions among successive perceptions in our thought or consciousness. Therefore, constancy and coherence are insufficient to secure the unity of the bundle. And lack of real connections among successive perceptions leaves us with a "loose" or "floating" bundle, and this may in turn leads to the problem of other minds. In that case, we cannot talk about a unified mind because perceptions, at least qualitatively identical perceptions, can occur in other minds. Without a doubt, Hume's account of sympathy in Book II of the *Treatise* indicates that we actually share the mental states of other people around us. In addition, insofar as constancy and coherence fail to unite our successive perceptions into a single bundle, they both fail to provide any criteria for the conditions of personal identity (Chapter IV). Second, Hume might have detected circularity in his original characterization of personal identity, something similar to what Olson has levelled against Shoemaker. Hume defines a perception (mental state) in terms of its causal/functional role in combination with other perceptions (mental states) that are members or states of *the same person*, and the effects (outputs) that such mental states produce belong to *the same person* or mind ⁽¹⁶⁾. When we characterise mental properties (perceptions) in terms of their membership of a single mind, we indeed, invokes or presupposes a notion of personal identity; we cannot have a complete account of what those properties are without making assumption about the persistence conditions of their bearers. Thus, any attempt to define personal identity in terms of causal/functional

relations among perceptions (states) will lead to circularity. If Hume holds that the causal chains of perceptions belong to the same person, then it clearly presupposes a notion of personal identity.

Supposedly, Hume had the two following options at his disposal: (a) either he had to accept that all perceptions belong to the same single mind; or (b) that they are not necessarily members or states of one unified mind. If he holds the former it leads to circularity, and if he holds the latter it leads to the problem of other minds and incoherence. Hume has no choice, but to plead the privilege of a sceptic. I contend that worries of this nature are powerful enough to make Hume despair.

To sum up, Hume, like Shoemaker, thinks that not only the right theory about the nature of mental states (perceptions) is a version of functionalism, but also that functionalism tells us important things about personal identity. He not only explains the basic tenets of a functionalistic theory of mind, but he also explains how that theory produces some of our most key philosophical concepts. As far as personal identity is concerned, he holds that functionalism as a theory about the nature of mental states implies a theory about the identity conditions of persons over time. However, as I explained above, he does not succeed to carry out his project, and that seems to be the cause of his worry in the Appendix.

It is worth noting that the current work is an interpretive one; I will try to show how Hume's conception of the mind fits the functionalist account of the mind, and that how he thinks such a functionalist theory implies personal identity. Although I maintain that Hume ultimately fails to account for the idea that personal identity follows from the functionalist theory of mind, I do not want to dispense with his psychological /functionalist continuity approach to the question of personal identity. There I am just trying to explain Hume's self-criticism in the Appendix. I believe that Hume, as well as Shoemaker, is in the right direction insofar as they think that functionalism tells us important things about personal identity. However, if Hume, like Shoemaker, had conceived the self as an independent substance that perpetually and spontaneously causing its successor states, then he would not have found himself in a situation to denounce the idea that the functionalist theory of mind implies personal identity.

In Chapter I, I set out the groundwork for my functionalist interpretation of

Hume's theory of the self. I, first, present the Lockean background to Hume's theory of the self. I want to show that Hume is influenced by Locke's account of personal identity in which the main question is the problem of the criterion of personal identity. Second, I argue that teleological functionalism provides an explanation, at least in a *prime facie* sense, for Hume's theory of personal identity. Third, I sketch out Hume's arguments for his fictional characterization of the idea of the self. He identifies the mind as a system of causal connection in which perceptions are defined by their causal/functional roles. This constitutes Hume's psycho-functionalism. Finally, I will examine Hume's views on the question of the nature of mind. I ascribe a version of 'property dualism' to him and conclude that such a position is compatible with his functionalist theory of mind.

In Chapter II, I discuss Hume's computational approach to the study of the mind. In particular, I try to show that his computationalism consists in a Representational Theory of Mind (RTM). I also discuss Hume's mechanics and the architecture of mind and their relations to his computationalism. The imagination plays a central role in his mechanical conception of the mind. And the question of the architecture of mind centres on the idea that the cognitive capacities and performances of the mind are automatic, modular, and sub-doxastic.

In Chapter III, I examine Hume's psychology of identity ascriptions. My main concern is to describe how the idea of personal identity as a kind of 'fictional' entity arises in the mind. There, I will explain a psychological belief-determining mechanism responsible for the generation of the ideas of both the continued and distinct existence of external objects and the self. Specifically, I discuss the roles that the principles of constancy and coherence play in such mechanism. Hume thinks that these principles are fundamental for the production of the fictional ideas of the self.

In Chapter IV, I argue that Hume thinks that his functionalism implies a psychological-functional continuity account of personal identity over time. In this respect, I will show that Hume's theory of personal identity "as it regards our passions or the concern we take in ourselves," and specially his views on character traits, and self-concern function as a mechanism that provides a causal basis for the continuity, or durability, of a person's intentions, interests, and concerns over time. I argue that Hume's

account is very similar to Shoemaker's functional continuity account of personal identity. However, it is my contention that Hume's account eventually fail because it is circular.

In Chapter V, I consider Hume's worry about an inconsistency in his theory of the self. I contend that Hume's functionalism, as presented in the main body of the *Treatise*, stores a problem for him and when he arrives at the Appendix he realises the problem and confesses that he is unable to resolve it. I will argue that Hume's account of constancy and coherence, as laid out in T,I.iv.2, establishes that the principles of constancy and coherence unite our series of successive perceptions in the mind by a process of imaginative postulation of an unperceived perception. These principles may successfully account for the arising belief in the idea of the continued and distinct existence of external objects and the idea of personal identity, but they fail to explain our belief in other minds (selves). In fact, Hume's account of *sympathy* in Book II of the *Treatise* establishes that we *actually* share the mental states of other minds (selves) around us. This shows that Hume treats the belief in other selves, like the beliefs in external objects and persons, as an important part of his psychology of identity ascriptions. Therefore, he must use the same psychological belief-determining mechanism, particularly the principles of constancy and coherence, to account for our belief in other minds. Nevertheless, coherence and constancy do not play any role whatsoever in the explanation of the arising of belief in other minds. Further, the absence of constancy and coherence in the explanation of the belief in other minds shows that not only the belief in other minds is not explained adequately by Hume's psychological belief-forming mechanism, but also the absence of constancy and coherence undermines even their roles as the necessary and sufficient conditions for the beliefs in objects and personal identity. This problem becomes even more pressing, if we note that Hume's theory of mind as defended in Book I of the *Treatise* is incompatible with his account of the self in Book II.

I

Hume's Functionalism: A First Approximation

The true idea of the human mind, is to consider it as a system of different perceptions or different existences, which are link'd together by the relation of cause and effect and mutually produce, destroy, influence, and modify each other....In this respect, I cannot compare the soul more properly to any thing than to a republic or commonwealth, in which the several members are united by the reciprocal ties of government and subordination, and give rise to other persons, who propagate the same republic in the incessant changes of its parts. And as the same individual republic may not only change its members, but also its laws and constitutions; in like manner the same person may vary his character and disposition, as well as his impressions and ideas, without losing his identity. Whatever changes he endures, his several parts are still connected by the relation of causation. (THN, 261)

Hume's explicit identification of the mind as a teleofunctional system (or functional organization, to use Locke's terminology) occurs in his argument in the section of "Of Personal Identity." Hume characterizes object identity, particularly the organic objects and the artifacts, in terms of an underlying internal organization striving towards a *common end* or *purpose*. He argues that we ascribe identity to variable and changing objects, like vegetables, animals, ships, and buildings, even though that any change of considerable part in a mass of matter, strictly speaking, destroys its identity. Hume claims that "there is, however, another *artifice*, by which we may induce the imagination to advance a step farther; and that is, by producing a reference of the parts to each other, and a combination to some *common end* or purpose. A ship, of which a considerable part has been chang'd by frequent reparations, is still consider'd as the same; nor does the difference of the materials hinder us from ascribing an identity to it. The common end, in

which the parts conspire, is the same under all their variations, and affords an easy transition of the imagination from one situation of the body to another” (THN, 257). An oak, grows from a small plant to a large tree, but it is still the same tree; an infant becomes a man, sometime fat, sometimes lean, without any change in his identity; a church that is built from brick, fell to ruin, and is rebuilt from free-stone, but it is still the same; and a river that is totally altered in less than twenty four hours, continue to be the same during several ages. Hume clearly defines the identity of these object (that are, variable and changing) in terms of an underlying teleofunctional principle.

More fundamentally, Hume claims that in order to explain the nature of personal identity (or the self), the same method of reasoning, which so successfully explains the identity of plants, animals, and ships, and houses, must be used. For Hume, personal identity, like object identity, must be constituted by the facts about the purpose, or function, of such existences. If the identity of an oak tree, for instance, is constituted by the participation of its parts in order to sustain the same continued life over time, the identity of the human mind is also constituted by the participation of its members (impressions and ideas) to maintain its overall psychological continuity or a sense of persisting self over time. Hume’s view of the self, as described, is based upon a metaphysical belief, which denies the need for granting a special status for the human mind. According to Hume, the mind is a compound, or more accurately, a bundle of perceptions that constantly undergoes alteration. Change and interruption is the most fundamental characteristics of human mind; the mental world, as Hume describes, is interrupted, variable, fragmentary, and broken. He contends that there is no ground to believe that either reason or senses can prove that the self is a simple unchanging substance. In fact, he claims that “the identity, which we ascribe to the mind of man is only a fictitious one, and of a like kind with that which we ascribe to vegetables, and animal bodies” (THN, 259). At the heart of Hume’s notion of fictional self lies the idea that, strictly speaking, there is no substantial continuity in personal identity; rather, what gives us the sense of identity and persistence over time is the *continuity of function*. In the self, as well as in vegetables and animals, the parts of each conspire to a common end or function, and this function persists even throughout the most radical transformation. In the case of vegetables and animals, the continuity of life is the kind of end or function

that gives those objects identity over time. In the case of artifacts, functional continuity is the kind of common end that provides those objects with the conditions of identity over time. In the case of the self, the psychological continuity of is the typical kind of end or function that provides us with a sense of identity over time. Hume's notion of a fictional self is precisely this continuity of function playing out over time.

Historically speaking, Locke's general theory of identity has significant bearings on Hume's theory of the self; his influence on Hume's theory of personal identity is undeniable. Roland Hall has argued that Hume wrote the section "Of Personal Identity" with Locke's *Essay* in front of him. As he compellingly argued, the verbal similarities between Locke's and Hume's texts are remarkably conspicuous. It is very possible that Hume wrote the section on personal identity with an eye on Locke's section of "Of Identity and Diversity" ⁽¹⁾. Locke provided separate accounts for the identity conditions of organic object, artifact, and person. In the case of person, he claimed that consciousness (memory) alone makes personal identity. And in the case of organic objects and works of art, he identified the continuity of function as their identity conditions. Hume adopts this Lockean theme, and extends it to the case of personal identity as well. Locke held that there are different criteria (conditions) of identity for different objects, whereas Hume holds that there is a *single* condition of identity for all objects, including the human mind. Locke is committed to the long-lasting tradition of thought that assigns special status to minds, but Hume departs from that tradition and attributes no special status to minds. He insists that our perceptions must be understood like other physical objects. In this sense, perceptions, as the objects of minds, have the same ontological status as the physical objects, because, in his view, our minds are as experienceable as the physical objects surrounding us. Hume does not see any substantial difference between mental and non-mental beings. We can experience our minds as much as we are able to experience the physical objects around us. Moreover, it is worth noting that Hume also criticizes and emphatically rejects the metaphysical doctrine of an immaterial and indivisible substantial self. He grants perceptions ontological independence implying that perceptions do not require any support from other things. Hume seems to conceive substance as an independent existence upon which other things depend, and since perceptions do not need any support to exist, it follows that there is no

need for perceptions to inhere in some simple and individual substance. Hume's criticism of substance, along with his denial of special status to the mind, enables him to put forward a single account for the identity conditions of both persons and objects. An underlying psychological belief-determining mechanism, which produces the belief in the identity of persisting selves and the belief in the continued and distinct existences of external objects functions as the *single* identity condition that accounts for both personal and object identity (Chapter III). Locke was the first philosopher, who shifted our attention from the substantial account of the self towards a psychological account of its persistence. However, it is Hume who for the first time explains that psychological mechanism that account for personal and object identity.

In this chapter, I will explain Hume's characterization of the self as a teleofunctional entity. I will show how the identity (continuity) of function gives rise to psychological continuity, which this in turn generates the idea of personal identity. The analogy of object identity, i.e., the identity that we ascribe to vegetables, animals, ships, and houses, plays a vital role in our understanding of Hume's teleofunctional theory of the self. Hume's teleofunctionalism is almost entirely neglected by his commentators, as if he never wrote anything about it. Nevertheless, Hume spends almost half of the section of "Of Personal Identity" hammering that the same method of reasoning that "so successfully explain'd the identity of plants, and animals, and ships, and houses, and of all the compounded and changeable productions either of art or nature" (THN, 259) must be used to explain personal identity. He takes the analogy to be vital to his understanding of personal identity. One may suggest that Hume's teleofunctionalism does not intend to establish a full-dressed functionalistic theory of mind. Rather, Hume wants to provide a *prima facie* case for his causal and/or functional account of the mind. Perhaps this may be the case, and the reason that Hume's commentators have neglected it has to do with his tendency to treat it as a *prime facie* case for his psycho-functionalism. However, I want to point out that Hume's tendency to treat his teleofunctional characterization of the self as a *prime facie* case might have been influenced by his predecessors' widely use of teleofunctional characterization of the mind. Locke's teleofunctional identification of organic objects and artifacts, for instance, has greatly influenced Hume's theory of the self. I contend that the understanding of Hume's teleofunctionalism is as important as his

psycho-functionalism in our overall conception of his functionalistic theory of the self, and there is substantial textual support indicating that Hume himself considers his teleofunctionalism a key part of his theory of the self.

My objective is not to provide a historical treatment of Hume's functionalistic theory of the self. Nor I intend to vindicate my interpretation of Hume's theory of the self by appealing to Locke's teleofunctionalism. Rather, I want to establish a case for the fruitfulness of interpreting Hume's theory of the self with Locke's notion of teleofunctionalism in view. Hume takes Locke's teleofunctionalism to be a genuine issue, for he thinks that it can be applied equally to personal identity. In this sense, it provides an excellent background to Hume's functionalistic theory of the self.

In I.1, I will provide a brief historical background to Hume's teleofunctionalism. In particular, I will contextualize his teleofunctional characterization of the self by showing its Lockean themes and concerns, and his reaction to them, that stands as the background of his theory of the self. In I.2, I will outline Hume's characterization of the self as a fiction and then compare it with his teleofunctional characterization. In I.3, I give a detailed analysis of Hume's idea of the self. There, I will show that there is no a single *impression* from which the idea of the self is derived. In I.4, I will discuss Hume's conception of the mind as a system of causal connection. And, finally, in I.5, I will examine Hume's view on the nature of the mind. I will ascribe a version of 'property dualism' to him and then argue that such view of the nature of mind is compatible with his functionalist theory of mind.

I.1: The Lockean Background to Hume's Functionalism

Locke is considered to be the precursor of modern personal identity theory. Historically speaking, he gave the problem its first formulation in the chapter of "Of Identity and Diversity" in *An Essay Concerning Human Understanding*. The historical significance of Locke's treatment of personal identity is best captured by Hume's reflection, almost fifty years later, where he states that the question of personal identity "has become so great a question in philosophy, especially of late years in *England*⁽²⁾"(THN, 259). Locke is still considered an influential figure in the current debate on personal identity; particularly the Psychological Continuity account is deep-rooted in his simple consciousness-memory theory of personal identity. There are, indeed, many contemporary philosophers who

describe themselves as Lockean or neo-Lockean in their understanding of the problem.

Locke's denial of substantial identity is the most distinctive feature of his account of personal identity. Instead, he characterizes personal identity in terms of psychological continuity, or continuity of consciousness-memory over time. Locke's scepticism concerning the possibility of our knowledge of substances is the source of his empirical analysis of the origin of the idea of the self. In his view, the human mind lacks the capacity to have any clear and distinct idea of substances, and thus lacks the knowledge of its essences. Locke's rejection of Descartes' doctrine about the intuitive knowability of the indivisible soul along with his denial that we can have knowledge of its *immortality* led him to pose the question of personal identity in this new mode. Locke is explicit in his rejection of Descartes' theory of substantial self. "It is not therefore unity of substance that comprehends all sorts of identity, or will determine it in every case," he argues, "but to conceive and judge of it aright, we must consider what idea the word it is applied to stands for" (*Essay*, 445). He clearly states that what gives rise to the notion of identity is a system or collection of simple ideas producing complex ideas, including the idea of substance. According to Locke, "our specific ideas of substance are nothing else but a collection of a certain number of simple ideas, considered as united in one thing" (*Essay*, 405). Consequently, Locke thinks that an empirical theory of ideas could explain our knowledge about the immaterial mind. In fact, he uses his theory of ideas and the constraints that that theory places upon the human understanding in order to show (a) how his theory of ideas gives rise to the idea of an immaterial mind and (b) that we have no knowledge about the substance of things. For Locke an immaterial mind, or soul, is a complex idea which is made up from simple ideas received originally from reflection. In the same manner, our ideas of material substances are derived from our simple ideas of sensation. The former gives rise to the idea of an immaterial mind, or soul, and the latter gives rise to the idea of a body (*Essay*, 414). Therefore, all our knowledge is primarily derived from simple ideas, and simple ideas do not provide us with any knowledge about the substance of mind. So, further enquiry into the nature of substance leads us into "darkness and obscurity, perplexedness and difficulties" (*Essay*, 314). Furthermore, on the basis of his empirical theory of ideas, he even denies the possibility of our knowledge about the immortality of the soul. Since we have no knowledge about the essence of

mind, it is possible that “it being in respect of our notions, not much more remote from our comprehension to conceive that GOD can, if he pleases, superadd to it another substance with a faculty of thinking” (*Essay*, 193). These passages show that Locke is firmly committed to an empirical conception of the mind, and that his empirical account of the human mind does not allow that our idea of the self inheres in a simple substance. Therefore, an account of the self that rests on substance, be it immaterial or material, is simply inadequate. Put more sharply, it is impossible for us to know anything enlightening about the nature of an immaterial thinking substance. This required an altogether new account of personal identity, distinct from the substance view of the self.

The philosophical significance of Locke’s novel account of the self does not lie “in what it achieved, but in what it attempted, not in the solution which it provided, but in the further problems which it raised” (3). Locke sets the stage for a whole new philosophy of the self; he sowed the seeds for a new wave of naturalization of the mind, in which Hume later played *the* central role. Perhaps the most significant element in Locke’s theory of the self, which became so fundamental in Hume’s work was his replacement of the Cartesian substantial view of the self, in which the identity of person depended on the identity of an immaterial substance, with a *relational* account of the self, in which personal identity depends upon the existence of a psychological relationship that binds together earlier and later stages of a person (4). Locke’s central teaching in this regard is that the persisting self cannot be understood as parasitic on the persistence of any underlying substance(s) constituting persons (5). Instead, personal identity consists in a kind of psychological continuity, which is made possible by the relations of ideas. In particular, Locke’s psychological continuity account of the self is derived from his Theory of Ideas. According to that theory, all ideas are derived from experience and our experiences are either sensory or reflective, the former being caused by the external objects and the latter being furnished by the operations of the mind on itself. “These, when we have taken a full survey of them, and their several Modes, Combinations, and Relations,” Locke states, “we shall find to contain all our whole stock of *Ideas*; and that we have nothing in our Minds, which did not come in, one of these two ways” (*Essay*, 106). If there is nothing in our minds but simple and complex ideas, it follows that what constitutes identity and diversity is essentially relations among these ideas and the ways

of comparing these ideas. In this sense, Locke's theory of ideas according to which complex ideas are formed from simple ideas is the basis of his distinctive notion of the self. This approach indicates a total departure from Descartes and finds new and radical treatment in Hume's bundle theory of the self.

Cartesians held that an immaterial simple substance would specify logically necessary and sufficient conditions for a person identified at t_1 to be the same person at t_2 (6). Locke rejected the Cartesian criterion of personal identity (7). He describes his criterion of identity as the following: "When we see anything to be in any place in any instant of time, we are sure ...that it is that very thing, and not another, which at that same time exists in another place, how like and undistinguishable soever it may be in all other respects: and in this *consists* identity" (*Essay*, 328: emphasis added). Locke concludes that no two entities of the same kind can occupy the same place at the same time. In this, he thinks, consists the identity of different objects. A page later, Locke again specifies his criterion of identity when he states his own account of the *principium of individuationis*, namely, that it "is existence itself, which determines a being of any sort to a particular time and place, incommunicable to two beings of the same kind" (*Essay*, 330).

Traditionally, the principle of individuation was understood as a logico-metaphysical issue concerning the conditions of identity (8). The Scholastics and the Cartesians treated the principle of individuation mainly as a logical and/or ontological and/ or metaphysical problem. They never thought of the principle as an epistemic and/or semantic issue. Locke also primarily understood the principle as a logico-metaphysical one (9). He defines it as "*existence* itself, which determines a being of any sort to a particular time and place, incommunicable to two beings of the same kind" (*Essay*, 330: emphasis added). This definition explicitly characterizes the principle of individuation as a logico-metaphysical issue. Most Scholastic philosophers conceived the principle of individuation in the context of the doctrine of substantial forms (10). However, Locke's theory of identity is to be understood as an objection to this Scholastic doctrine of substantial forms and qualities (11). And the reason for his objection against this doctrine is rooted in his intention to provide a foundation or a defence for corpuscular mechanism (12). Locke joined Boyle in attacking scholasticism and buttressed the 'mechanical philosophy' of his time (13). The upshot of this for Locke's theory of identity is that he is

no longer able to appeal to substantial forms to account for the conditions of identity. Locke, like Boyle, rejects the doctrine that the substantial form of an entity is the causal basis of its essential properties and the unifying principle of its matter. Conceived in this way, the doctrine of the substantial form is at variance with Locke's mechanical philosophy. However, as Noonan has noted "[a]lthough Locke's notions of substance and matter are so manifestly unaristotelian, something like Aristotle's *substantial form* holds a prominent place in his thought, at least with respect to living creatures" (14). In fact, for Locke, the continuity of the same organized life (functional continuity) functions as a kind of Aristotelian *substantial form*. In this sense, Locke's notion of 'organized life,' that is, functional continuity, is "meant to fill the role the substantial form plays in making for the identity of organisms, but in a way totally different from the way forms do it" (15). In short, in the same way that the substantial form played the role of the conditions of individuation among the Scholastics, functional continuity plays the role of the criterion of identity for Locke. According to Locke, functional continuity provides logically necessary and sufficient conditions for the continuity of organic objects and artifacts.

However, the principle of individuation, and subsequently the criterion of identity, can be conceived as an epistemic or semantic/linguistic issue as well (16). It seems that Locke was the first philosopher who began to distinguish between the metaphysical and the epistemological aspects of the conditions of identity. The epistemic search for the necessary and sufficient conditions of identity involved the determination of the conditions that make possible the knowledge of individual identity. The epistemic aspect draws our attention to the *relation* between a knower and the object of his knowledge, an entirely new approach to the problem of individuation. Accordingly, it suggests that any solution to the problem of individuation has to depend to a certain extent on the nature of knower and the object of his knowledge (17). This *relational* view of individuation will have certain implications for both Locke's and Hume's accounts of personal identity. Apparently, Locke thought that the metaphysical view of the conditions of identity was inadequate to produce a theory of identity that is based on, or at least, incorporates some of our fundamental psychological-natural tendencies involving our belief fixation of personal and object identity. Locke thought about the principle of

individuation in both metaphysical and epistemic senses. The result of such an understanding of the principle of individuation was that he offered different sets of conditions for personal and object identity. Locke proposed different *sorts* of identity conditions for persons and objects. He argued that what an organism consists in is different from what constitutes the self ⁽¹⁸⁾. Yet, what constitutes an atom or a mass of matter is different from what constitutes the organic objects and the self. In fact, what constitutes the identity of plants, animals, and man is the internal organization and the functions that continue over time, whereas what constitutes the self is memory as consciousness. What constitutes atoms and masses of matter is the unchanging nature of atoms as the particles composing masses of matter. Now that Locke's thoughts on the conditions of identity are in place, I shall briefly explain his theory of identity.

The simplest forms of existence are atoms. An atom is "a continued body under one immutable Superficies, existing in a determined time and place; it is evident, that considered in any instant of its existence, it is in that instant the same with itself. For, being at that instant what it is, and nothing else, it is the same, and so must continue as long as its existence is continued; for so long it will be the same, and no other" (*Essay*, 330). Atoms are physically indivisible, and thus, they exclude all other matters from the space that they occupy. Since they have no parts, it follows that they have no empty space within them that would permit interpenetration. They are also insusceptible to changes, and therefore do not undergo changes in their figure, bulk, number, and unity. But they experience changes in their state of motion, position, and situation, which somehow affect the integrity of the bulk of mass of matter. The continued existence of an atom is dependent on it continuing to occupy space. Therefore, the identity condition of an atom is determined by its spatio-temporal continuity and incommunicability. The identity of atoms, which is unchanging and uninterrupted, is primary and *perfect*.

In the case of masses of matter, what constitutes them is the composition of the objects—that is, the "cohesion of particles of matter any how united" (*Essay*, 330). Their identity is determined by the continuity of their atomic elements. It is said that "if two or more atoms be joined together into the same mass, every one of those atoms will be the same, by the foregoing rule: and whilst they exist united together, the mass, consisting of the same atoms, must be the same mass, or the same body, let the parts be ever so

differently jumbled. But if one of these atoms be taken away, or one new one added, it is no longer the same mass or the same body” (*Essay*, 330). So the same criterion of spatio-temporal continuity could be applied to masses of matter. Masses of matter undergo change with respect to their qualities, i.e., figure, bulk, state of motion and situation, but remain identical as long as their atomic particles remain the same. Hence, the Ship of Theseus, when its old planks are replaced by new planks, is not the same ship insofar as it loses both its original simple substances and the cohesion of its particles any how united. But a statue made of clay that is smashed and reproduced from the same mass of matters is the same statue.

It is worth stressing that Locke’s criterion to individuate inanimate objects is a metaphysical one. Strictly speaking, it claims that the existence, and continuation, of atoms as simple substances are both necessary and sufficient to individuate inanimate objects over time. The metaphysical account seeks to determine the identity conditions of objects on the basis of the existence of a certain ‘principle’ or ‘cause’ regardless of the way that minds may know something as individual. For Locke, atoms, as simple substances, are that very ‘principle’ or ‘cause’ that determine the identity of inanimate objects. Indeed, unlike the epistemological account, which involves the conditions under which something is known as individual *qua* individual, the metaphysical account claims that the mere existence of atoms is both necessary and sufficient to secure the continuity of a coherent unified mass of matter.

In the case of organic objects, Locke claims that ‘sameness of life’ or functional continuity constitutes their identity conditions. Locke’s account of identity involving animate objects—plants and animals—departs from his metaphysical account involving the identity of inanimate objects. Locke notices that these objects undergo changes, and because they frequently experience changes it means that the atomic particles that they are made of, or the cohesion of particles of matter any how united, do not remain the same over time. As a result, the identity condition of inanimate objects is inadequate to account for the identity of animate objects. He writes:

That being then one plant which has such an organization of parts in one coherent body, partaking of one common life, it continues to be the same plant as long as it partakes of the same life, though that life be communicated to new particles of matter vitally united to the living plant, in a like continued organization conformable to that sort of plants. For this organization, being at any one instant in any

one collection of matter, is in that particular concrete distinguished from all other, and is that individual life, which existing constantly from that moment both forwards and backwards, in the same continuity of insensibly succeeding parts united to the living body of the plant, it has that identity which makes the same plant, and all the parts of it, parts of the same plant, during all the time that they exist united in that continued organization, which is fit to convey that common life to all the parts so united. (*Essay*, 331)

Locke also thinks that the case is not essentially different in the identity conditions of animals. The identity conditions of both plants and animals consist in fulfilling the functions of nourishment and growth that constitute their vegetative and animal lives. The continuity of a plant or an animal's existence is the continuity of their proper functioning in order to maintain the common life which vitally unites its parts. Locke, like Boyle, compares the functionings of plants and animals with the functionings of a machine. He uses this analogy to illustrate his opinion. He claims:

The case is not so much different in brutes but that any one may hence see what makes an animal and continues it the same. Something we have like this in machines, and may serve to illustrate it. For example, what is a watch? It is plain it is nothing but a fit organization or construction of parts to a certain end, which, when a sufficient force is added to it, it is capable to attain. If we would suppose this machine one continued body, all whose organized parts were repaired, increased, or diminished by a constant addition or separation of insensible parts, with one common life, we should have something very much like the body of an animal. (*Essay*, 331)

Clearly, Locke's account of identity involving animate objects, and artifacts, brings up the notion of functioning as a necessary and sufficient condition for the identity of these objects. Furthermore, Locke extends the scope of his analogy and likens the identity of organic objects and machines to the identity of man, as a biological kind. He asserts that "this also shows wherein the identity of the same man consists; viz., in nothing but a participation of the same continued life, by constantly fleeting particles of matter, in succession vitally united to the same organized body" (*Essay*, 331). The main idea here is that manhood, like plants and animals, is constituted by its functional organization. Of course, this organization is instantiated at any given time by a collection of atomic particles. But the organization can persist over time and experience gradual changes in the particles of which it is made up. However, despite the gradual and constant changes in the particles of the body as such, the functions that organisms perform continue over time. In the case of organic objects, including man, the most important of these functions

is the continuation of the same life. For Locke, ‘sameness of life’ or functional continuity is the condition (criterion) of identity for these kinds of objects, be it an oak or a horse, or a man.

In the case of person, Locke has famously said that consciousness makes personal identity alone. Locke characterizes consciousness as the following:

For, since consciousness always accompanies thinking, and it is that which makes everyone 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 a 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; it is the same self now it was then; and it is by the same self with this present one that now reflects on it, that that action was done. (*Essay*, 335)

Locke is unequivocal here: personal identity consists in the sameness of consciousness. He never, however, specifies precisely what the sameness of consciousness consists in. Roughly, it says that a person P_1 at time t_1 and P_2 at t_2 is the same person only if the person at t_2 is conscious of having experienced what P_1 had experienced at t_1 . Locke’s commentators have interpreted the sameness of consciousness in terms of the continuity of memory or the capacity to remember. Hence, a person P_2 at time t_2 is the same person as P_1 at time t_1 if and only if P_1 and P_2 are both persons and P_2 can *remember* at t_2 (his doing) what P_1 did or experienced at t_1 . The basic idea is that a person’s identity with its former selves consists in just what that person remembers the former self experienced or did. Locke is explicit in the above passage that it is memory that retraces the spatio-temporal history of every single individual consciousness. Indeed, memory guarantees the fact that the identity of person consists in the backward extension of consciousness to any past action or thought. This is the simplest version of Locke’s memory theory of personal identity.

It is also pointed out that central to Locke’s theory of personal identity “is the idea that consciousness is reflexive and that it plays a dual role in self-constitution: it is what unifies a person not only *over* [diachronic unity] time but also *at* [synchronic unity] a time” (19). The self- reflexive (20) nature of consciousness is best captured in the passage where Locke defines a person as:

[A] thinking intelligent being, that has reason and reflection, and can consider itself as itself, 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: it being impossible for any one to perceive without perceiving that he does perceive. When we see, hear, smell, taste, feel, meditate, or will anything, we know that we do so. (*Essay*, 335)

Consciousness conceived in the self-reflexive sense poses a more inclusive account of personal identity than the simple memory theory. We cannot see, hear, smell, taste, feel, and meditate without knowing that we do so. It is this immediate self-consciousness that constitutes the nature of the self. It is ‘inseparable from thinking’ and ‘essential to it.’ Consciousness as such includes something far beyond remembering. It includes (a) the fact that one *remembers* having the experience of doing or feeling something, and (b) the awareness that one is having those experiences. When one remembers having had an experience or having performed an action, say, feeling distressed or being angry, one is not just aware that it is *him*, who, in the present, is having the experience of remembering, but one is also aware that it was *him*, who had the experience of feeling distressed or being angry. Thereby, one can claim the ownership of both remembering and having had the original experience, say, of feeling distressed, and appropriating (claiming the ownership of) that experience.

At the heart of Locke’s account of personal identity is the idea that the identity of ‘immaterial substance’ is neither necessary nor sufficient for the identity of a person. He denied that personal identity requires the continuous identity of any underlying material or immaterial substances. He separated the *cogito* from the concept of a thinking substance (21). Instead, consciousness alone constitutes the self, regardless of what substance (matter) it is composed of. The underlying substance could be “spiritual, or material, simple or compounded, it matters not” (*Essay*, 341). Locke’s rejection of the substantial view of the self has a huge ramification for his account of personal identity. It functions as the backdrop for Locke’s analogy between personal identity and the identity of organisms. The analogy between ‘sameness of life’ or functional continuity and consciousness supplied the main framework for Locke’s argument that the identity of persons is conceptually independent of any particular theory about the nature and what underlies the substance of the mind. The analogy between life and consciousness shows that in the same way that life is an organizing principle which unites a variety of fleeting or ever-changing parts into one continuing organism, consciousness is also a principle

that unites a variety of fleeting, interrupted, and discontinued perceptions into a unified mind. In the same way that life *functions* as the principle of unity of living organisms, consciousness also *functions* to unite interrupted perceptions, sensations, thoughts, and actions into a coherent self. Locke writes:

Different Substances, by the same consciousness (where they do partake in it) being united into one Person; as well as different Bodies, by the same life are united into one Animal, whose *Identity* is preserved, in that change of Substance, by the unity of one continued life. For it being the same consciousness that makes a man be himself to himself, *personal Identity* depends on that only, whether it be annexed only to one individual Substance, or can be continued in a succession of several Substances (21). (*Essay*, 336)

This analogy shows that personal identity may be preserved in a succession of different thinking substances (perceptions). As ‘sameness of life’ or functional continuity was essential to *organism*, consciousness is essential to *person*. In the same way that the identity of an organism is preserved by the unity of one continued life while its material substances undergo constant alterations, personal identity is also preserved by the same consciousness uniting different thinking substances into one person. Or as functional continuity unites various vital processes such as nutrition, respiration and growth into a single and coherent individual, consciousness unites various mental states and processes such as deliberation, reasoning, and most importantly, memory, into a single person (22). This analogy plays an important role in Locke’s account of personal identity. He used it against those who held that personal identity consists in an immaterial simple substance. “Those who place thinking in an immaterial substance only,” Locke argued, “must show why personal identity cannot be preserved in the change of immaterial substance... as well as animal identity is preserved in the change of material substances” (*Essays*, 335). Accordingly, successive mental substances can give rise to personal identity, and as a result, there is no need for personal identity to inhere in an immaterial simple substance.

To sum up, Locke thought consciousness plays a *quasi*-functional role by unifying different mental substances or perceptions into a single mind. His analogy between life and consciousness shows that in the same way that life unites different material substances into a persisting being, consciousness also unites different perceptions into a persisting self. Locke thought about the self as a collection of different psychological properties that is consistently and constantly interrupted, though united by

consciousness. Although there is no textual evidence showing that he ever thought about the self along the line of the bundle theory of perceptions, it would be surprising if Hume did not take notice of Locke's functional explanation of living organisms and his analogy between life and consciousness. In fact, it would be very surprising, if Hume did not notice the time-sliced nature of Locke's notion of consciousness-memory and his theory of personal identity for that matter. I think that Hume saw a potential bundle theory of mind in Locke's philosophy. However, he does not think that personal identity deserves a different treatment from the explanation given to the identity of plants, animals, and artifacts. Hume uses the same method of reasoning to explain the identity of organic objects, artifacts, and the human mind.

Hume, like Locke, firmly believes that the question of personal identity can be determined without recourse to the substantial view of the self. He forcefully argues that the identity of person consists in certain modes of psychological continuity, but he rejects the idea that consciousness or memory makes personal identity alone. In particular, Hume holds that one must utterly rid oneself of Descartes' doctrine of self-reflexive consciousness in order to be able to construct a true theory of the self. On his account, there is neither an abiding self, or perfect identity in the mind, nor any *real* connection between the different perceptions that constitute the mind. The fictional identity that we ascribe to ourselves is merely a quality which we attribute to the mind because of the union of the ideas in the imagination (THN, 260). Personal identity consists in a psychological relation *tout court*. Undoubtedly, Locke's aim was essentially different from Hume's. Locke wanted to determine the limits of moral responsibility whilst Hume's central aim is to describe the origin of one's natural belief in a universal but mistaken belief in the persistence of self over time. Locke was preoccupied with the idea that consciousness, or memory, can guarantee the continuity of a person in the afterlife (23), whereas Hume is mainly concerned to specify the psychological causes, or the psychological mechanism, which generates the natural but mistaken belief in personal identity over time. Accordingly, the two philosophers' theories of the self stem from different inspirations. Nonetheless, the similarities between them on the question of identity are immense. Hume adopts Locke's functionalistic scheme and extended it to explain personal identity as well.

The crucial shift that takes place in Hume's account of the self is rooted in Locke's basic idea that it is the abstract ideas of things, their nominal essences (24), which give rise to the notion of identity. Hume takes Locke's abstract ideas to be the mere construction of the perceiving mind. My functionalistic reading of Hume's theory of the self relies to a great extent on this premise. According to Hume, it is our psychological mechanism that is solely responsible for the ascription of identity to all kinds of objects, persons or otherwise. Strictly speaking, Hume holds that identity does not consist in something *outside* of the mind; rather, identity is something that the mind constructs itself. Hence, the notion of identity is a mind-dependent idea altogether. This is rooted in his epistemological approach to the question of personal identity. Hume seems to be doing away completely with the metaphysics of identity; instead, he thinks that the epistemic view provides an adequate and coherent general theory of identity, including personal identity. He begins to think that the main problem concerning identity, particularly personal identity, is the psychological mechanism that triggers the feeling that produces our belief in the identity of different objects.

I.2: Hume's Characterization of the Self as a Fiction

Hume discusses the self and its identity in two main places: in the section of "Of Personal Identity" in Part IV of Book I of the *Treatise* and in the Appendix published a year later with Book III. In the Appendix, he expresses his dissatisfaction with his earlier treatment of the topic, and confesses that he now finds an inconsistency in his original characterization of the self (Chapter V). His discussion about the continuity and identity of external objects in section of "Of Scepticism with Regard to the Senses" and his argument in the section of "Of the Immateriality of the Soul" have direct, though implicit, implications on his theory of the self. In the former, he establishes that a psychological belief-forming mechanism (Chapter III,) (25) gives rise to the ideas of object and personal identity, and in the latter he criticises a presupposition of both sides of the debate between materialists and immaterialists, that the soul is a substance in which perceptions inhere. Hume's rejection of this presupposition is the foundation of his characterization of the self as a fiction in the section of "Of Personal Identity."

The section entitled "Of Personal Identity" contains the fundamentals of Hume's theory of the self. Roughly, it can be divided into two main parts. In the first part, starting

from page 251 and ending in page 255, Hume provides us with his identification of the self as a fiction. There, I think, Hume offers a brief account of his psycho-functionalism, and in the second part he explains his teleofunctional theory of the self. In what follows, I will sketch out Hume's argument for the characterization of the self as a fiction, and then briefly discuss his argument for a teleofunctional account of the self. His argument for the characterization of the self as a fiction has five main steps (26):

I) Hume argues that the human mind lacks "perfect identity."

- 1) Perfect identity is the invariability and uninterruptedness of an object, thro' a supposed variation of Time. (THN, 201)
- 2) Impressions always give rise to real ideas (The Copy Principle).
- 3) There is no impression invariable and constant.
- 4) There is no idea of the self as invariable and constant (from 1-3).

Therefore, there is no such conception of the self considered as something having "perfect identity."

II) Hume argues that that to which we ascribe perfect identity is a "bundle or collection of different perceptions."

- 5) All our particular perceptions are "different, and distinguishable, and separable from each other, and may be separately consider'd, and may exist separately, and have no need of any thing to support their existence." (THN, 252) (From the Separability Principle)
- 6) When I enter most intimately into what I call *myself*, I always stumble on some particular perception or other...I never can catch myself at any time without a perception, and never can observe any thing but the perception." (THN, 252)
- 7) The self is nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement." (THN, 252) (from 4-6)

Therefore, "there is properly no *simplicity* in [the mind] at one time, nor *identity* in different." (THN, 252)

III) He argues that there must be some relations that constitute the 'fictitious' or 'imperfect' identity' of the mind.

- 8) We have a great propensity to ascribe an identity to the successive perceptions.
- 9) All objects, to which we attribute identity, without observing their invariableness and uninterruptedness, are such as consist of a succession of related objects.

10) The identity that we ascribe to the self is an 'imperfect' or 'fictitious' one. (From 1, 7-9)

Therefore, only the successive perceptions constitute the mind.

IV) Hume argues that what gives us the propensity to attribute identity to this flux of successive perceptions is the mistaken supposition that we have invariable and uninterrupted existence through the entire course of our life. He maintains that there cannot be any "real connexions" discoverable by the *understanding*, but only mere association of our perceptions in the *imagination*.

11) The identity, which we ascribe to the self, is similar to that which we ascribe to vegetables and animal bodies.

12) The identity of vegetables and animal bodies derives from certain operations (functions) of the imagination.

13) The identity of the self must proceed from the same operation (function) of the imagination.
(From 11-12)

14) The identity which we ascribe to the self, however perfect we may imagine it to be, is not able to run the several different perceptions into one...every distinct perception, which enters into the composition of the mind, is a distinct existence, and is different, and distinguishable, and separable from every other perception, either contemporary or successive. (THN, 259)
(From 5, 10)

15) The 'fictitious' relation of identity, which unites our perceptions, is either something that *really* binds our several perceptions together, or only associates their ideas in the imagination.

16) The understanding never observes any real connexion among objects, and even the union of cause and effect, when strictly examin'd, resolves itself into a customary association of ideas.
(THN, 260)

17) Identity is nothing really belonging to these different perceptions, and uniting them together; but is merely a quality, which we attribute to them, because of the union of their ideas in the imagination, when we reflect upon them. (THN, 260) (From 10 and 13-16)

Therefore, the relation that constitutes personal identity cannot be "real connexions" discovered by the *understanding*, but rather by the *imagination*.

V) Hume claims that these relations are resemblance and causation.

18) The only qualities, which can give ideas a union in the imagination are ...resemblance, contiguity, and causation.

19) Contiguity has little or no influence on the present case (the self).

20) The self is a bundle of perceptions related by resemblance and causation. (From 17-19)

Therefore, what unites perceptions into a bundle is resemblance and causation.

Hume thus arrives at a theory of personal identity that is strictly explained in terms of the relations of causation and resemblance. Hume generally downplays the role of *contiguity in space* in this case, but thinks that *contiguity in time* may play some role. He describes the mind as a bundle of perceptions to which we attribute identity and simplicity in virtue of the association of its perceptions. Hume rejects the idea of an internal impression as the origin of a simple and unchanging self. He claims that, when we reflect upon ourselves, we only find succession of ordinary but related impressions and ideas. We can never catch ourselves at any time without perceptions; we never encounter a self, but only some perceptions. However, we are inclined to believe that we are simple and identical selves persisting over time. According to him, the idea of the self is a ‘fiction’ that results from the confusion of the succession of related perceptions (objects) with that of the uninterrupted and invariable objects. In fact, what gives us such a great natural propensity to ascribe an identity to these successive perceptions, and to suppose ourselves possessed of an invariable and uninterrupted existence through the whole course of our lives, is that we generally confound or mistake ‘perfect identity’ with the ‘succession of related perceptions.’ Hume argues that resemblance and causation are the cause of this confusion or mistake that make us substitute the notion of identity with diversity (THN, 254). So we “feign the continu’d existence of the perceptions of our senses, to remove the interruption: and run into the notion of a *soul*, and *self*, and *substance*, to disguise the variation” (THN, 254). Our propensity to this mistake is so powerful that we fall into it before we are even aware of it; and though we may correct this mistake, and return to a more reflective method of thinking, “we cannot sustain our philosophy, or take off this bias from the imagination” (THN, 254). In short, in Hume’s view, the mind is a collection of ontologically distinct existences, each separable and distinguishable, lacking ‘real connexions’ among them; and we feel connexions between them because of the effects of the associative principles, viz., resemblance, causation, and continuity in time. The source of the idea of the self is a “smooth and uninterrupted progress of the thought along a train of connected ideas” (THN, 260). He concludes that identity is not something inherent in the perceptions or in the relations connecting them together, but merely a quality that we attribute to them because of their union in the imagination. Therefore, the idea of the self is neither derived from a simple and unchanging impression nor it is

constituted by some ‘real connexions’ among perceptions. Such an idea is a fictional existence that we merely attribute to the succession of our related perceptions. This is Hume’s official position in the main text of the *Treatise*, and he never wavers until he arrives at the Appendix. In the remainder of this section, I want to develop a more in-depth analysis of Hume’s characterizations of the self as a fiction.

Hume conceives the mind as a succession of ontologically distinct but related perceptions that are associated by certain ‘qualities’—resemblance, contiguity, causation (THN, 11). This is called the *relational* view of the self. Hume’s relational view of the self opposes both the Cartesian conception of the mind as a simple substance and Locke’s memory-consciousness theory of personal identity. This constitutes the negative phase of his theory of personal identity. The issue that Hume is immediately concerned with in the section of “Of Personal Identity” is to reject the existence of any idea of the self as simple and identical. On his account, in order to have a simple and identical idea of the self we must either locate the corresponding impression from which the idea of the self is derived or to show that perceptions inhere in a simple substance. Hume’s empiricism implies that any idea of the self as simple and identical must be traceable to its original impression, but since there is no such impression empirically known to us that we could derive the idea of the self from, it follows that the simple substance view of the self must be rejected altogether. In rejecting this view, Hume makes the following introspective observation:

[W]hen I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe anything but the perception. When my perceptions are remov’d for any time, as by sound sleep, so long as am I insensible of *myself*, and may truly be said not to exist. (THN, 252)

The conclusion that Hume draws here is that we fail to locate any corresponding impression from which the simplicity and identity of mind is derived. Therefore, we must affirm that “what we call a *mind*, is nothing but a heap or collection of different perceptions, united together by certain relations, and suppos’d, tho’ falsely, to be endow’d with a perfect simplicity and identity” (THN, 207). This ‘bundle’ of different perceptions, “which succeed each other with an inconceivable rapidity” is in perpetual flux and movement. “The mind is a kind of theatre,” Hume says, “where several perceptions successively make their appearance; pass, re-pass, glide away, and mingle in

an infinite variety of postures and situation. There is properly no *simplicity* in it at one time, nor *identity* in different; whatever natural propension we may have to imagine that simplicity and identity” (THN, 253). Hume immediately warns that “[t]he comparison of the theatre must not mislead us. They 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” (THN, 253). This cautionary remark is intended to avoid misconceptions about the existence of any ‘real connexion’ among perceptions and the mistaken view that these perceptions belong to a mind with perfect simplicity and identity. Hume stresses this concern more bluntly in the *Abstract* where he says “it must be our several particular perceptions, that compose the mind. I say, *compose* the mind, not *belong* to it. The mind is not a substance, in which the perceptions inhere” (THN, 658: emphasis original). For Hume, the idea of the mind is only that of particular perceptions causally related to each other, without the existence of any ‘real connexions’ among perceptions or perceptions belonging to a mind as a substance inhering them.

Hume’s rejection of the substantial account of the self constitutes the negative phase of his theory of personal identity. On such an account, the attempt to locate the impression from which the idea of the self is derived is entirely unsuccessful. “Self or person is not any one impression,” he argues, “but that to which our several impressions and ideas are suppos’d to have a reference” (THN, 251). Supposedly, even if there was any single impression that gave rise to the idea of the self “that impression must continue invariably the same, thro’ the whole course of our lives; since self is suppos’d to exist after that manner. But there is no impression constant and invariable” (THN, 252). Therefore, there is no impression that the idea of the self is derived from and “consequently there is no such idea” (THN, 252). The core of Hume’s negative teaching on the self consists in his impression hunting and the failure to find one.

Hume, however, offers a positive philosophy about the idea of the self that is “cultivated by true metaphysics” (THN, 11). Hume’s main concern here is to account for the idea of the self as something to which we ascribe an identity over time. Hume makes a distinction between two different senses of identity, namely, ‘perfect identity’ and ‘improper identity.’ To have a ‘perfect identity’ is to have an idea of an enduring or

identical self that “remains invariable and uninterrupted thro’ a suppos’d variation of time (THN, 253). This means that, for instance, there should be a single impression from which the idea of the self is derived, and that such an impression remains invariable and uninterrupted over time. Hume’s verdict here is that there is no such an impression, and thus, there is no ‘perfect identity.’ Hume asserts that, however, “we attribute identity, in an improper sense, to variable or interrupted objects” (THN, 255). He claims that we have a distinct idea of several different perceptions existing in succession, though connected together by a close relation. This, he says, “to an accurate view affords as perfect a notion of *diversity*, as if there was no manner of relation among objects [perceptions]” (THN, 253). The idea of ‘improper identity’ results from a propensity of the imagination to confound, or mistake, the idea of ‘perfect identity’ with that of diversity. And this common way of thinking that “makes us substitute the notion of [perfect] identity, instead of that of related objects” (THN, 254) is so powerful “that we fall into it before we are aware; and tho’ we incessantly correct ourselves by reflection, and return to a more accurate method of thinking, yet we cannot long sustain our philosophy, or take off this bias from the imagination” (THN, 254). Thus, by mistake we ascribe identity to a succession of related perceptions. At this point, I want to take a short detour in order to explain Hume’s distinction between two kinds of philosophy.

In Hume’s view, there are two kinds of philosophy: (a) a kind of dogmatic philosophy (metaphysics) that claims certainty about the ultimate nature of things, and (b) a real philosophy, ‘true metaphysics’ that claims knowledge within the bounds of sense-experience and observation. The former leads us into “sophistry and illusion” (THN, 163), whereas the latter provides us with a “new scene of thought” (27)—a new experimental philosophy of human nature—“which will not be inferior in certainty, and will be much superior in utility to any other of human comprehension” (THN, xix). Hence, for Hume, philosophy (metaphysics) has two different faces; (a) that which indulges itself in ‘abstruse,’ and ‘presumptuous and chimerical’ speculation (THN, xvii), which “arise[s] either from the fruitless efforts of human vanity, which would penetrate into subjects utterly inaccessible to the understanding, or from the craft of popular superstitions, which being unable to defend themselves on fair ground, raise these entangling branches to cover and protect their weakness” (EHU, 11); and (b) that which

“enquire[s] seriously into the human understanding, and show[s], from an exact analysis of its powers and capacity, that it is by no means fitted for such remote and abstruse subjects...[We] must cultivate true metaphysics with some care, in order to destroy the false and adulterate” (EHU, 12). Indeed, delineating the “proper province of human reason” (EHU, 12), requires adopting a negative and critical approach, which rebukes dogmatic metaphysics. It is only through this negative approach that we are able to rid ourselves from the popular, superstitious, abstruse, and unjustified reasoning of the school metaphysics. By rejecting this speculative philosophy we are engaged in an “accurate and just reasoning” (EHU, 12-13). Accordingly, “there are many positive advantages” (EHU, 13) to this accurate and just reasoning. In fact, the positive phase of Hume’s program is deeply embedded in his empirical analysis of the origins of some of our fundamental philosophical ideas, i.e., the idea of necessary connection, object identity and personal identity. In sum, Hume’s new philosophical program has two different, but related dimensions: the repudiation of metaphysics (the negative phase) and the establishment of an empirical/experimental science of human nature (the positive phase).

Returning, then, to what Hume has said about the idea of the self, it becomes clear that “abstruse philosophy” fails to account for the idea of the self as something simple and identical. However, he notes that we all, as ordinary people, believe in a persisting self over time. Hume claims that “we have a great propensity to ascribe an identity to these successive perceptions, and to suppose ourselves possess of an invariable and uninterrupted existence thro’ the whole course of our lives” (THN, 253). He identifies this position with what he calls the *vulgar* view, a non-philosophical view, which is the position we all hold at one time or another. Hume’s “most profound metaphysics” (THN, 189) is to give a satisfactory explanation of the origins of the idea of a persisting self as conceived by the vulgar. On this account, the idea of a persisting self, as held by ordinary people, is viewed essentially as a product of the association of ideas that are united by the imagination. Hume parallels his treatment of the self with that of the continued and distinct existence of external objects. In both cases, there are certain natural relations existing among our perceptions that lead us to confound the idea of identity with diversity. In the case of external objects, the vulgar fails to distinguish between perceptions—as the immediate contents present to our consciousness—and objects as

independent and distinct from the mind. The generality of mankind are unable to make a distinction between the objects and their representations in the mind; they recognize only one single existence (THN, 202). But philosophers are able to differentiate between the mind-dependent perceptions and the mind-independent objects. Philosophers are able to explain the vulgar belief in the existence of external objects as an idea that arises from features of our sense-impressions. As Hume tells us, the vulgar confounds the succession of related objects (perceptions) with something which possesses a perfect simplicity and identity. More specifically, he argues that there are certain relations among our perceptions, particularly, the relation of resemblance, that make us mistake, or confuse, the idea of diversity with identity (THN, 204). The succession of related perceptions places the mind in the same disposition as when it is experiencing an uninterrupted and invariable object. The passage between successive but resembling perceptions is “so smooth and easy, that it produces little alteration on the mind, and seems like the continuation of the same action; and as the continuation of the same action is an effect of the continu’d view of the same object, ’tis for this reason we attribute sameness to every succession of related objects. The thought slides along the succession with equal facility, as if it consider’d only one object; and therefore confounds the succession with the identity” (THN, 204). But since philosophical reflection makes us aware of the interrupted nature of successive but resembling perceptions we find ourselves involved in ‘contradiction.’ However, the power of the propensity to ascribe a perfect identity to what is interrupted and variable is so compelling that we have no alternative except to “feign the continu’d existence of the perceptions of our senses, to remove the interruption; and run into the notion of a *soul*, and *self*, and *substance*, to disguise variation” (THN, 254). Hence, Hume’s ‘most profound metaphysics’ arrives at the notion of a *fictional* soul, or self. In fact, Hume’s ‘real philosophy’ intends to achieve two different but related things: (a) to rebut the sophistry and illusion of the popular metaphysics regarding the simplicity and identity of the self by way of explaining the fictional nature of their doctrine, and (b) to explain how we come to ascribe an identity to persons and objects in the face of interruption and disunity. Philosophy, as Hume understands, is intended with revealing the true nature of our fictional beliefs in the existence of external objects and the identity and simplicity of the mind. “The most perfect philosophy of the natural kind only staves

off our ignorance a little longer,” he says, “[a]s perhaps the most perfect philosophy of the moral or metaphysical kind serves only to discover larger portions of [our ignorance]. Thus the observation of human blindness and weakness is the result of all philosophy, and meets us, at every turn, in spite of our endeavours to elude or avoid it” (EHU, 31).

Before turning to the next section, I want to point out that, for Hume, consciousness is the same as the very occurrence of perceptions in the mind; it is not something distinct, or over and beyond perceptions themselves (28). Memory plays an important role, though a secondary one in Hume’s bundle theory of the self. Basically, memory is the only means through which we become acquainted with the continuance and extent of the succession of our perceptions as the only source of a fictitious notion of the self (THN, 261). “Had we no memory, we never shou’d have any notion of causation,” Hume says, “nor consequently of that chain of causes and effects, which constitutes our self or person” (THN, 262). Nonetheless, he concludes that “memory does not so much *produce* as *discover* personal identity, by showing us the relation of cause and effect among our different perceptions” (THN, 262). The reason that Hume rejects the memory account of identity is that the idea of personal identity involves something more than what memory provides. It involves, for instance, resemblance, causal connexion among perceptions, the imagination, etc. Therefore, in Hume’s view, the Lockean memory-consciousness condition of personal identity is insufficient to produce the idea of the self. I turn now to the second part of Hume’s argument in the section of “Of Personal Identity.”

I.3: The Mind as a System of Causal Connection

In the second part, beginning from page 255 and ending in page 262, Hume argues that his notion of fictional self, as characterized in the first part, can be proven “to the satisfaction of every fair enquirer, [that] to show from daily experience and observation, that the objects, which are variable or interrupted, and yet are suppos’d to continue the same, are such only as consist of a succession of parts, connected together by resemblance, contiguity, or causation”(THN, 255). Accordingly, the idea that personal identity is a fiction or that our ascription of identity to persons over time is a mistake or a false belief, can be vindicated by our “daily experience and observation.” Hume argues that our common experience tells us that we typically ascribe identity to changeable

objects— plants, animals, and artifacts—while we so *palpably* know that their form, size, and matter are going to totally alter over some time. The identity that we ascribe to these objects is only in an ‘improper sense.’ That is, their identity is fictional. For Hume, the problem of accounting for our mistaken belief in the existence of persisting self is the same problem of accounting for our palpably false belief in the identity of plants, animals, houses, and ships. In fact, he argues that the same psychological mechanism of the imagination which account for our ascriptions of identity to plants, animals, houses, etc., can equally account for the ascriptions of identity to persons over time. Thus, he claims that “[t]he identity, which we ascribe to the mind of man, is only a fictitious one, and of a like kind with that which we ascribe to vegetables, and animal bodies. It cannot, therefore, have a different origin, but must proceed from a like operation of the imagination upon like objects” (THN, 259). Thus, for Hume, the psychological mechanism of the imagination that generates our belief in the fiction of personal identity is the same mechanism that produces our common sense belief in the identity of organic objects and artifacts. This identity-ascribing mechanism operates on the basis of a propensity or disposition of the mind which confounds identity with the succession of related but distinct perceptions.

In general, Hume wants to establish that the question concerning the identity of persons is no different from the question about the identity of objects, and that the identity which we ascribe to both of them is palpably fictional. He claims that there is a single psychological mechanism that produces their fictional ideas. In this respect, the self is no different from an oak tree, and thus it should not be treated differently either. The identity of an oak tree is accounted for by the kind of teleofunctional continuity that underlies it. Likewise, the identity of the self must be explained by the same kind of teleofunctional continuity that constitutes it. I need to examine these in some details.

We have seen already that, on Hume’s account, certain features of our perceptions make us ascribe an identity to the successive but related perceptions despite their interruption, and we disguise the interruption by forming the idea of a ‘real existence’ that connects them throughout (THN, 199). Basically, Hume thinks of the idea of identity in terms of a relation which belongs to an object viewed in different times. As a result, he distinguishes between the notion of identity, on the one hand, and the ideas of unity and

number, on the other. A single object at any particular moment of its history is said to convey the idea of unity, not identity. Neither multiplicity nor numerical identity conveys the idea of identity. He concludes that since number and unity are incompatible with the relation of identity, it must lie in something else.

It is the idea of time or duration that forms the idea of identity as a medium between the ideas of unity and number (THN, 201). Consequently, he formulates his *principium individuationis*, or the principle of identity, as follows: “the invariableness and uninterruptedness of any object, thro’ a suppos’d variation of time” (THN, 201). Seemingly, Hume’s principle of individuation entails ‘perfect identity’ when we ascribe identity to an object whose parts continue uninterruptedly and invariably the same despite observing any motion or change of place either in the whole or in any of the parts (THN, 255). Hume here comes closest to Locke’s metaphysical principle of individuation. At first sight, it seems that he, like Locke, holds that the identity of, for instance, a mass of matter, consists in the continuity of its minute particles over time in spite of the fact that its atomic structure or figure may change. This position is metaphysical in the sense that it allows for the possibility of having identity as something distinct and independent from our perceptual experience (29). Nevertheless, it seems to me that Hume refuses to accept the Lockean metaphysical principle of individuation. He rejects Locke’s principle of individuation because it does not distinguish between the idea of identity and the idea of unity. Locke said that an object “considered in any instant of its existence, it is in that instant the same with itself... and so must continue as long as its existence is continued; for so long it will be the same, and no other” (*Essay*, 330). In reply to this, Hume says that the proposition that an object is the same with itself conveys the idea of unity and not identity (THN, 200). For example, a single perception (object) in any particular moment of its existence is said to convey the idea of unity. We can only talk about identity when time and duration is involved; a perception, P_1 , at t_1 , is identical with P_1 at t_2 , if it continues to be the same without any interruption or variation. Whereas Locke seems to be saying that an object, P_1 , at t_1 , is identical, if it is the same with itself in that moment. It seems that this is sufficient to guarantee the continuity of the object. Of course, Locke allows identity over time, but that continuity strictly consists in the continuity of unity over time. So it seems that Hume is right that Locke confuses the idea of unity with that

of identity. He is also right that the principle of individuation cannot consist in the unity—the mere existence of an object in a particular time and place—of an object. For Hume, identity is essentially a relation that consists in the succession of interrupted but related objects in the mind of a perceiver. Strictly speaking, the relation of identity is generated by our perceptual experiences. Unlike Locke who individuated objects in terms of their existences outside the mind as mind-independent entities, Hume individuates objects in terms of their representations in the mind as mind-dependent objects. Indeed, the reason that Hume equates ‘objects’ with ‘perceptions’ indicates the fact that, for him, it is exclusively those contents that generate the principle of identity. For Hume, the principle of identity is essentially an epistemic one. This explains why Hume departs from Locke’s principles of identity as a metaphysical one.

The consequence of this departure is very important for Hume’s theory of identity. He applies his theory of identity equally to persons, inanimate objects, and plants and animals. In all these cases, we ascribe identity to the things that are interrupted and variable. Philosophical reflections make us aware of the fact that all we have is, indeed, the idea of *diversity*, i.e., the succession of interrupted but related objects, filling the place of the idea of identity. It is the action of the imagination through which we come to believe that an uninterrupted and invariable object is the same as the succession of interrupted but related objects. The feeling, or disposition of mind, in both states are almost the same, and this resemblance is the cause of the confusion, or mistake, which makes us substitute the idea of identity with that of diversity (THN, 254). This is why we confound identity with diversity. The reason that Hume, unlike Locke, applies his theory of identity equally to all kinds of objects, including the self, is directly related to the argument defended in the last paragraph. Simply put, for Hume, the idea of identity, which is a natural relation generated by the imagination, is entirely a mind-dependent idea. The core concept is that the idea of identity is something that exists in the mind, not in the objects, and that the ideas of object and personal identity are formed by nothing but the determination of our thoughts. To put it differently, the mind projects the idea of identity on objects outside the mind as well as on itself. Or, the mind constructs the idea of personal identity and object identity, inanimate objects or otherwise, by the operation of a singular psychological belief-determining mechanism. The consequence of this is

that the notion of identity is something that is constructed by the mind, and as such it is completely irrelevant to what kinds of objects it is applied. It is not the properties of objects, or their kinds, structures, figures, place, etc., that determine their identity; rather, it is the ways of our perceptual system, or ways of ideas, that determine the identity of different objects. That is why Hume has a single theory of identity that is equally applicable to all kinds of objects and persons. Let me explore this further.

There is a canonical passage in which Hume asserts that understanding the question of personal identity, in its deepest sense, is connected to our understanding of the identity of objects. “To explain it perfectly we must take the matter pretty deep,” he says, “and account for that identity, which we attribute to plants and animals; there being a great analogy betwixt it, and the identity of a self or person” (THN, 253). A few pages later he stresses the same point and concludes: “And here ’tis evident, the same method of reasoning must be continu’d, which has so successfully explain’d the identity of plants, and animals, and ships, and houses, and of all the compounded and changeable productions either of arts or nature” (THN, 259). He goes on to say, “the identity, which we ascribe to the mind of man, is only a fictitious one, and of a like kind with that which we ascribe to vegetables and animal bodies. It cannot, therefore, have a different origin, but must proceed from a like operation of the imagination upon like objects” (THN, 259). The analogy is strikingly clear. But what is not so clear is in what sense personal identity is analogous to object identity. Hume tries to account for this analogy in the rest of this section of “Of Personal Identity.” There he explicitly discusses different cases of plants, animals, and artifacts, and shows how a kind of “*common end* or purpose” (THN, 257) fills the role of an identifying principle. Put otherwise, for Hume, the identity of these kinds of objects consists in a kind of *teleological* functioning of all parts of the object pertaining to a certain end. It is said that despite the fact that plants, animals, and artifacts undergo total changes, we still attribute identity to such changing objects because their different parts conspire to attain a general purpose or end. Similarly, the mind, as a collection of distinct perceptions—all the perceptions that enter into the composition of the mind remain to be distinguishable and separable from each other—is united by the relation of identity according to which all the distinct existences conspire to produce a unifying conceptions of the mind. For instance, an oak that grows from a small plant to a

large tree continues to be the same oak despite the fact that all its particles of matter alter over the course of its life (THN, 257). Similarly, the mind continues to be the same entity in spite of the fact that new perceptions constantly enter into its composition and alter its contents. In the same way that all parts and particles of an oak conspire to achieve a certain end, say, an organized life, all the perceptions in a mind also strive towards performing certain *functions* achieving identity of mind.

Hume's strategy concerning the identity of objects is to show that we ascribe identity to them despite their continuous alteration. We consider a mass of matter, say, a mountain that undergoes *small* or *inconsiderable* changes, as being the same, though, strictly speaking, these changes absolutely destroy its identity. When the change is produced *gradually* and *insensibly* the mind "feels an easy passage from the surveying its condition in one moment to the viewing of it in another, and at no particular time perceives any interruption in its actions. From which continu'd perception, it ascribes a continu'd existence and identity to the object" (THN, 256). The conclusion that Hume draws here is that "since this interruption makes an object cease to appear the same, it must be the uninterrupted progress of the thought, which constitutes the imperfect identity" (THN, 256). But even when the changes are observed to become considerable we attribute identity to such changing objects. Here, Hume introduces the notion of an "artifice, by which we may induce the imagination to advance a step further; and that it, by producing a reference of the parts to each other, and a combination to some *common end* or purpose" (THN, 257). The notion of 'artifice' that Hume brings up here palpably refers to the functional nature of mind. The idea of common end, or functioning, is an 'artifice' of the mind because the mind constructs that idea; it is the workmanship, to use Locke's language, of the human mind. We attribute identity to changing objects simply because such changes, either gradual or total, cannot destroy or override the *function* of a system as long as the *function* or "common end, in which the parts conspire, is the same under all their variations, and affords an easy transition of the imagination from one situation of the body to another" (THN, 257). The notion of 'artifice' clearly indicates that the identity that we ascribe to a changing object is intrinsically an *artificial* or *fictional* identity. That is to say, the mind produces, or constructs, an artificial/fictional identity for varying objects; "A ship, of which a considerable part has been chang'd by

frequent reparations,” Hume argues, “is still consider’d as the same; nor does the difference of the materials hinder us from ascribing an identity to it” (THN, 257). The reason for ascribing identity to it is that even though the materials composing the ship are almost entirely altered, the ship as an entity performing certain functions remains the same under all the variations.

Hume argues that there is an additional element, which would explain the functional nature of mind in a more direct and lucid way. This additional component is the remarkable “*sympathy* of parts to their *common end*,” or “the reciprocal relation of cause and effect in all their actions and operations” (THN, 257). The causal/reciprocal relation is primarily observable in plants and animals; “where not only the several parts have a reference to some general purpose, but also a mutual dependence on, and connexion with each other” (THN, 257). “An oak, that grows from a small plant to a large tree, is still the same oak,” he argues, “tho’ there be not one particle of matter, or figure of its parts the same. An infant becomes a man, and is sometimes fat, sometimes lean, without any change in his identity” (THN, 257). Hume does not explain why the causal relation is only observable in plants and animals, but tells us that the effect of such a causal-reciprocal relation, or mutual dependence of several parts, is so strong that we attribute identity to variable and interrupted objects despite the fact that such objects endure total change in which their form, size, and substance are entirely altered. The notion of ‘artifice’ is at work here as well. Hume compares the fictional notion of a *substance*, or *soul*, or *self*, which arises as a result of confounding identity with variation, to the notion of artificial/fictional identity of plants and animals, which arises as a result of “our propension...to imagine something *unknown* and *mysterious*, connecting the parts, beside their relation” (THN, 254-55: emphasis added). In the same way that we suppose the continued existence of our successive perceptions to remove the interruption, which results in feigning the notion of self, or soul, we also feign ‘something unknown and mysterious’—a general purpose or function—to disguise the changing objects and, thus, attribute identity to such compound and changeable objects.

Although Hume explicitly states that the causal/reciprocal relation is observable only in plants and animals, he later broadens the scope of its application and employs it in the case where we confound numerical and specific identity. He brings up the case of a

breaking noise where “a man, who hears a noise, that is frequently interrupted and renew’d, says, it is still the same noise”; [and] tho’ ’tis evident the sounds have only a specific identity or resemblance, and there is nothing numerically the same, but the cause, which produc’d them” (THN, 258). Or when a church, for instance, “which was formerly of brick, fell to ruin, and that the parish rebuilt the same church of free-stone, and according to modern architecture... [in which] neither the form nor the materials are the same, nor is there any thing common to the two objects, but their relation to the inhabitants of the parish; and yet this alone is sufficient to make us denominate them the same” (THN, 258). Finally, he says that despite the fact that a river alters completely in less than twenty four hours, “this hinders not the river from continuing the same during several ages” (THN, 258). In all these cases, we confound the numerical identity, say, the continuous and uninterrupted noise for a period of time, with that of specific identity, i.e., the interrupted noises that have specific identity or resemblance. Likewise, we confound the rebuilt church with the old one because of the confusion between the numerical and specific identity. The same is also true of the river. Hume clearly states that in the cases of rivers and noises they have nothing numerically the same except the causes that produced them. In these cases, the first object is completely annihilated before the second comes into existence, except that the causes or sources of them remain the same. That is why we attribute identity to them regardless of the fact that they are numerically distinct objects. The case of the church is slightly different in the sense that there is no material or formal causes that are in common between the two objects, but only the way that they are related to the inhabitants of the parish—that is, the function or common purpose of a church in a community of people.

Hume’s most ingenious and resourceful case in relation to the functional nature of the mind occurs when he openly compares the mind to a republic or commonwealth. Accordingly, to understand the true idea of the human mind—that is, “to consider it as a system of different perceptions or different existences, which are link’d together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other”(THN, 261)—is to properly compare it to a “republic or commonwealth, in which the several members are united by the reciprocal ties of government and subordination, and give rise to other persons, who propagate the same republic in the incessant changes

of its parts” (THN, 261). In the same way that a commonwealth may change its members, laws and constitution, an individual’s characters, dispositions, as well as his impressions and ideas, may alter him without losing his identity. As a consequence, whatever changes these two systems may endure, their varying parts are still connected by the relation of causation (THN, 261).

This analogy is reminiscent of Hobbes’ functionalistic theory of the State. He then models his functional theory of the State, or the artificial man as he calls it, after his functional theory of the mind⁽³⁰⁾. Hobbes, in his Introduction to the *Leviathan*, where he likens a commonwealth or a republic to an individual human, claims:

Why may we not say that all automata (engines that move themselves by springs and wheels as doth a watch) have an artificial life? For what is the *heart* but a *spring*; and the *nerves* but so many *strings*, and the *joints* but so many *wheels* giving motion to the whole body, such as was intended by the artificer. (*Leviathan*, 3)

He goes on to say that in that commonwealth:

The *sovereignty* is an artificial soul, as giving life and motion to the whole body; the *magistrates* and other *officers* of judicature and execution, artificial *joints*, *reward* and *punishment* ... are the *nerves*, that do the same in the body natural; the *wealth* and *riches* of all the particular members are the *strength*; *salus populi* (the people’s safety) its *business*; *counsellors*, by whom all things needful for it to know are suggested unto it, are the *memory*; *equity* and *laws*, an artificial *reason* and *will*; *concord*, *health*; *sedition*, *sickness*; and *civil war*, *death*. Lastly, the *pacts*, and *covenants* by which the parts of this body politic were at first made, set together, and united, resemble that *fiat*, or the *let us make man*, pronounced by God in the creation. (*Leviathan*, 3)

This architectonic passage clearly states the scheme of a functionalistic theory of the State which is modelled on a functional conception of the mind. Hume’s use of the notion of a commonwealth as a functional entity constituted by the causal/reciprocal relations among its citizens is fundamentally a Hobbesian concept. Hobbes’ artificial conception of the State is most vividly stated when he asserts that in the same way that God declares *let us make man*, we also declare *let us make State*. Hence, in his view, a commonwealth is an artificial construction of the will of its members manifested in the formation of different bodies of government, i.e., institutions, ministries, etc. A commonwealth, as depicted here, exists in the minds of its members, and, thus, is a mind-dependent existence. Hume utilizes Hobbes’ artificial conception of the State to explain the functional nature of the mind. Similarly, the mind is a collection of ontologically distinct

existences that are tied by the reciprocal/causal relations among its members.

Generally speaking, for Hume, it is the causal relations among our perceptions that produce a sense of continuity for objects and persons. He maintains that the mind is not an entity or substance distinct from the perceptions themselves. “We have therefore no idea of substance,” he argues, “distinct from that of a collection of particular qualities, nor have we any other meaning when we either talk or reason concerning it” (THN, 16). Accordingly, the mind is neither an entity over and beyond the collection of perceptions that it is, nor it is an underlying substance in which they inhere. Alternatively, it is the kind of various operations of perceptions and their causal connexions that will enable us to identify mental states and events as those of belief, memory, passions, etc. in terms of the functionalist theory of mind. Indeed, it is the causal relations among our perceptions that produce these mental phenomena. As Pitson has summed it up, “[t]here is a causal flow among impressions and ideas, originating with impressions of sensation, with particular impressions of reflection, volition and the passions, giving rise not only to ideas but also to the actions we perform. On this account, then, the mind is *organized* in a certain way, with volition, belief, passion, etc., each performing certain functions and, by their interaction, producing bodily behaviour” (31). These indicate that Hume understands the mind as a system of causal connection.

To summarize, I have shown that Hume identifies the mind as a system or collection of different perceptions standing in certain causal relations to each other. I have shown that there is textual support for Hume’s adherence to teleofunctionalism. Hume’s explicit identification of the mind as a teleofunctional system follows Locke’s characterization of object identity in terms of an underlying functional continuity (internal organization) striving towards a *common end* or *purpose*. He argues that we ascribe identity to variable and changing objects, like vegetables, animals, ships, and houses, even though that any change of considerable part in a mass of matter, strictly speaking, destroys its identity. Hume claims that there is another artifice, which induces the imagination to produce a reference of the parts to some *common end* or purpose. This common end, in which all the parts conspire, continues to be the same under all the variations. An oak, grows from a small plant to a large tree, but it is still the same tree; an infant becomes a man, sometime fat, sometimes lean, without any change in his

identity; a church that is built from brick, fell to ruin, and is rebuilt from free-stone, but it is still the same; and a river that is totally altered in less than twenty four hours, continue to be the same during several ages. These all clearly show that there is an underlying functional continuity in these objects, and only this kind of continuity constitutes their identity. Therefore, *identity of function* makes these objects identity alone.

More fundamentally, Hume claims that in order to explain the nature of personal identity, the same method of reasoning, which so successfully explained the identity of plants, animals, ships, and houses, must be used. “The identity, which we ascribe to the mind of man,” Hume concludes, “is only a fictitious one, and of a like kind with that which we ascribe to vegetables, and animal bodies” (THN, 259). For Hume, personal identity, like object identity, must be constituted by the facts about the purpose, or function, of such being. If the identity of an oak tree, for instance, is constituted by the participation of its parts in order to maintain the same continued life over time, the identity of the human mind is also constituted by the participation of its members (impressions and ideas) to maintain its overall psychological continuity, or a sense of persisting self, over time. In the same way that continuity of function provides the identity condition for organic objects and artifacts, a similar functional continuity accounts for the identity condition of persons over time. Hume calls both of his accounts of object identity and personal identity “fictitious.” Nonetheless, he does not want to suggest that strictly there is no such thing as an identical self. In fact, he wants to say only that a “perfect identity” is not part of its essential nature. Hume firmly believes that we assume a continuity, or constancy, in the self, as in a plant or animal, which persists throughout the most radical transformations, but it falls short of providing perfect and absolute identity. He indicates that all that correspond to this assumed continuity or constancy is identity of function alone.

We have seen that Hume identifies the mind with a causal system of perceptions. It is by virtue of these causal relations that we arrive at the fictional view of the self as something to which our perceptions belong. We also saw that there is a sense according to which the self persists over time by virtue of some functional continuity. At this point, the question that naturally arises is how, in Hume’s view, mind and body are related to each other. For it can be argued that it is impossible to give an adequate account of the

diachronic identity of the self independently of its relation to body (32). Hume deals with this topic primarily in the section of “Of the Immateriality of the Soul.” In the next section, I turn to his discussion about the nature of the mind in this section.

I.4: Hume on the Nature of the Mind

Descartes and his followers described the mind as an unextended simple substance that interacts with the body as an extended substance. The problem that Descartes’ substance dualism encountered was how an unextended substance could causally interact with an extended substance. Hume strongly rejects the Cartesian idea of the mind, or self, as something *simple* and *identical* underlying our perceptions (THN, 253). Alternatively, the materialists argued for a conception of mind that reduces it into material substance altogether. Hume also denies that the nature of mental states could be somehow reduced to the physical *simpliciter*. On his account, the materialist and the immaterialist arguments for the immateriality of the soul are both equally inconclusive. On the one hand, he argues that the immaterialists are mistaken since some perceptions derived from sight and touch are extended and divisible, yet joined to indivisible perceptions. On the other hand, he holds that since whatever is extended consists of parts, and whatever consists of parts is divisible, it follows that thoughts or perceptions that are indivisible cannot be conjoined to anything divisible. Therefore, thought and extension are incompatible. Hume concludes that since some perceptions are unextended and indivisible while others, sight and touch, are extended and divisible, it follows that both materialists and immaterialists are mistaken. But what does Hume himself believe? Hume defends a type of ‘property dualism.’ I discuss this now.

Hume’s Dualism

It is known to us now that in 1734 Hume went to La Flèche, in Anjou, the Jesuit seminary attended by Descartes in the seventeenth century, where he must have spent much of his time writing the penultimate draft of Books I & II of the *Treatise*. Although this single incident may not have any philosophical significance, it can be said, at least symbolically, that Hume chose La Flèche to write his masterpiece in a Cartesian spirit. Of course, it is inaccurate to claim that Hume is a Cartesian in the sense that, for instance, Father Malebranche was a Cartesian philosopher. There is no doubt that the differences between them are substantial and numerous. At least, they belong to two different traditions of

philosophy. Nevertheless, some of Hume's fundamental philosophical principles are essentially Cartesian. For instance, Hume adopts his Theory of Ideas without any criticism from his predecessors, which can be traced back to Descartes (33). Moreover, it is recently argued that Hume is unshakably committed to the "assumptions of Cartesian Representationalism" (34). More importantly, Hume is committed to a form of mind-body dualism. David. M. Armstrong argues that Hume has used "a version of the Properties argument, which we have seen Descartes using" (35). Daniel E. Flage claimed that "Hume adhered to a rather sophisticated form of dualism, a dualism of both entities and systems" (36).

Hume's property dualism is best captured in the following passage in the *Treatise*:
Whatever is extended consists of parts; and whatever consists of parts is divisible, if not in reality, at least in the imagination. But 'tis impossible any thing divisible can be *conjoin'd* to a thought or perception, which is a being altogether inseparable and indivisible. For supposing such a conjunction, wou'd the indivisible thought exist on the left or on the right hand of this extended divisible body? On the surface or in the middle? On the back- or fore-side of it? If it be conjoin'd with the extension, it must exist somewhere within its dimensions. If it exist within its dimensions, it must either exist in one particular part; and then that particular part is indivisible, and the perception is conjoin'd only with it, not with the extension: Or if the thought exists in every part, it must also be extended, and separable, and divisible, as well as the body; which is utterly absurd and contradictory. For can any one conceive a passion of a yard in length, a foot in breadth, and an inch in thickness? Thought, therefore, and extension are qualities wholly incompatible, and never can incorporate together into one subject (THN, 234).

This 'remarkable' argument attempts to establish the basis of Hume's principle that mental objects, thoughts or perceptions, must be ascribed to an immaterial mind. This is because mental objects lack extension and spatiality and, thus, only belong to the category of indivisible and simple whilst matter is divisible. He goes on to write:

This maxim is *that an object may exist, and yet to be no where*. And I assert, that this is not only possible, but that the greatest part of beings do and must exist after this manner. An object may be said to be no where, when its parts are not so situated with respect to each other, as to form any figure or quantity; nor the whole with respect to other bodies so as to answer to our notions of contiguity or distance... A moral reflection cannot be plac'd on the right or on the left hand of a passion, nor can a smell or sound be either of a circular or a square figure. These objects and perceptions, so far from requiring any particular place, are absolutely incompatible with it, and even imagination cannot attribute it to them (THN, 235).

He concludes:

For we have only this choice left, either to suppose that some beings exist without any place; or that they are figur'd and extended; or that when they are incorporated with extended objects, the whole is in the whole, and the whole in every part. The absurdity of the last two suppositions proves sufficiently the veracity of the first. Nor is there any fourth opinion. For as to the supposition of their existence in the manner of mathematical points, it resolves itself into the second opinion, and supposes, that several passions may be plac'd in a circular figure, and that a certain number of smells, conjoin'd with a certain number of sound, may make a body of twelve cubic inches; which appears ridiculous upon the bare mentioning of it (THN, 239).

Accordingly, a moral reflection or a smell or sound cannot be placed or said to be circular or square; it would be absurd to ascribe thickness, length, and breadth to our desires. They are incompatible with extension and local conjunction with matter and body. Here Hume rejects the materialist notion of the mind, which was acknowledged by Hobbes and others, and endorses a form of 'dualism.'

Hume's commitment to property dualism is unwavering. He makes a categorical distinction among immaterial and material properties. Hume shows that both materialists and immaterialists are mistaken. The immaterialist philosophers argue that whatever is material has parts and is divisible, but thought is indivisible and thus it cannot be joined to the divisible. In reply, Hume argues that the immaterialists are mistaken since some perceptions derived from sight and touch are extended and divisible yet joined to indivisible perceptions. Likewise, he argues that the materialists are also mistaken because some perceptions are unextended and indivisible. These perceptions "*may exist, and yet be no where*" (THN, 235). For an object of thought to be no where means that "its parts are not so situated with respect to each other, as to form any figure or quantity; nor the whole with respect to other bodies so as to answer to our notions of contiguity or distance" (THN, 235). This includes all our perceptions except those of sight and touch. Hence, a moral reflection or a smell or sound cannot be placed or said to be circular or square. They are incompatible with extension and conjunction with matter and body.

The maxim that an object may exist without any place calls for further comment. In Hume's view, when we ascribe a local conjunction with body to an object that has no extension we are inclined to a kind of illusion or fiction. To illustrate, Hume tells us that we are used to suppose that the taste of, say, a fig, lies in the object itself. But reflection

makes us to ask ourselves if the taste is in one part of the object or it is so extended in every part of it. According to Hume, the imagination determines us to associate the taste with the extended object. But when we ascribe the taste of the fig to every part of its body we come to believe that the taste itself has shape and size. Reason assures us, however, that this is impossible because taste cannot have shape, size, and location. As a result, a conflict arises: On the one hand, we are influenced by the “inclination of our fancy by which we are determin’d to incorporate the taste with the extended object” (THN, 238), and on the other hand, reason shows us the impossibility of such a union. To resolve this contradiction, “[w]e suppose, that the taste exists within the circumference of the body, but in such a manner, that it fills the whole without extension, and exists entire in every part without separation” (THN, 238). This is typical of using the scholastic principle of *totum in toto et totum in qualibet parte* (the whole in the whole and the whole in every part). Hume concludes that this brings us to the absurdity of saying “that a thing is in a certain place, and yet is not there” (THN, 238). In order to avoid this absurdity we must recognize that certain things exist but have no place. “A moral reflection cannot be plac’d on the right or on the left hand of a passion,” he argues, “nor can a smell or sound be either of a circular or a square figure. These objects and perceptions, so far from requiring any particular place, are absolutely incompatible with it” (THN, 236). Thus, this gives way to the maxim *that an object may exist, and yet be no where*. The problem of materialism which supposes that thought or perceptions could be conjoined with matter and extension is that it ascribes spatial location to something that is entirely incapable of it. In summary, Hume thinks that since odors, sounds, and tastes—all non-visual and non-tactual perceptions—are incompatible with location and, thus, have no shape, size and extension, it suffices to conclude that total materialism is unwarranted.

On the other hand, Hume claims that some perceptions, viz., perceptions of sight and touch, are extended and, thus, conjoined to matter. That is, these perceptions—visual and tactual—have location, extension, size, and shape. The reason that he holds that these perceptions have these qualities is that they *represent* external objects, and that representation requires a resemblance between that which represents and that which is represented. For Hume, ideas are faint images or copies of impressions. On numerous occasions, he remarks that simple ideas are images of simple impressions. In the opening

paragraph of the *Treatise*, he asserts that ideas are ‘faint images’ of impressions (THN, 1). Later on, he tells us that ideas are copies of impressions (the Copy Principle): “all our ideas or more feeble perceptions are copies of our impressions” (THN, 72). Or generally, he says that ideas are *derived* from impressions (THN, 4, 7, 19, 33, 37, 161, 243). Ideas, as images, resemble their impressions: “impressions can give rise to no ideas, but to such as resemble them” (THN, 63). In another occasion, he says that “every idea is deriv’d from some impression, which is exactly similar to it” (THN, 33). In regard to the resemblance between simple ideas and simple impressions he claims “that all our simple ideas in their first appearance are deriv’d from simple impressions, which are correspondent to them, and which they exactly represent” (THN, 4). Therefore, Hume concludes that “ideas always represent the objects of impressions, from which they are deriv’d” (THN, 37). In fact, he explicitly states that our ideas that are copied from our impressions “represent them in all their parts” (THN, 96). That is to say, representation of all parts of external objects by perceptions consists in the exact resemblance between the qualities of the external objects and the qualities of their copies in the mind. Hume’s central idea here is that the idea of an extended object is itself extended. Thus he writes:

That table, which just now appears to me, is only a perception, and all its qualities are qualities of a perception. Now the most obvious of all its qualities is extension. The perception consists of parts. These parts are so situated, as to afford us the notion of distance and contiguity; of length, breadth, and thickness. The termination of these three dimensions is what we call figure. This figure is moveable, separable, and divisible. Mobility, and separability are the distinguishing properties of extended objects. And to cut short all disputes, the very idea of extension is copy’d from nothing but an impression, and consequently must perfectly agree to it. To say the idea of extension agrees to any thing, is to say it is extended. (THN, 239)

Robert Anderson in a seminal paper convincingly argues that “Hume intends that just as an idea, being copied from an impression, must “agree to” or resemble that impression, so also that impression, being copied from and representing an object, must “agree to” or resemble that object. Thus the impression of an extended object, such as a table, must itself be extended, and the idea derived from that impression must also be extended” (37). Hence, it is safe to say that, for Hume, the impression or idea of the extended object is literally an image of that object, and resembles it in that sense. And since Hume believes that what is extended must have a particular figure (THN, 235), it follows that these

perceptions also must have some shapes—that is, the same shapes as the objects that they represent.

Moreover, Hume seems to be suggesting that ideas of vision and touch that represent objects resemble them in having size. He claims:

But our ideas are adequate representations of the most minute parts of extension; and thro' whatever divisions and subdivisions we may suppose these parts to be arriv'd at, they can never become inferior to some ideas, which we form. (THN, 29)

In fact, the size of the idea may be small, but still it resembles it perfectly:

This however is certain, that we can form ideas, which shall be no greater than the smallest atom of the animal spirits of an insect a thousand times less than a mite. (THN, 28)

Finally, it seems plausible to say that “[i]f adequate representation requires the perception of an extended object to be itself extended, and if adequate representation requires the perception of an object having size itself to have size, then adequate representation would seem to require the perception of an object having location itself to have location” (38). These arguments and passages concerning the extension, figure, size, and location of some perceptions show that, for Hume, there is an exact resemblance between the perception and the object that it represents. There is no doubt that visual and tactual perceptions perfectly resemble the objects they represent in terms of extension, shape, size, and location, in spite of the fact that these perceptions cannot have the same location and size as the objects they represent—they are within the spatial limits of the brain (39).

Hume’s theory of representation, as explained here, is essentially an *iconic* one. An iconic mode of representation is a pictorial or imagistic kind of representation in which all the parts of the object that is represented appear in the representation. That is to say, the representing perception has a similar shape and size to the represented object, as well as location and extension, though in a smaller size and, of course, in a different location. Many contemporary philosophers hold that Hume’s theory of mental representation (Theory of Ideas) is imagistic. Jerry Fodor has argued that Hume’s theory of mental representational is a kind of non-conceptual, or more precisely, an *iconic* mode of representation (40). Don Garrett has also said that Hume “treats all mental representations as fundamentally *imagistic*—that is, as qualitatively *resembling* the contents of sensory and reflective experience” (41).

What we can gather from these arguments is that Hume holds that both

materialists and immaterialists are mistaken about the question of the substance of the mind because it appears that some perceptions are unextended and indivisible, and thus, immaterial, while others (sight and touch) are extended and divisible, and thus, material. Therefore, we must reject both materialism and immaterialism and treat the question concerning the substance of the mind as unintelligible. In fact, Hume's strategy in treating the question of the immateriality or the materiality of the mind as unintelligible is to maintain a neutral position about the *immortality* of the soul (THN, 250). As a result, he leaves open the possibility that the soul is immortal, as much as he does the possibility that it is mortal. The question that still remains to be answered is that what is Hume's own position on the mind/body relation? As we saw so far, he has rejected substance dualism, and even Spinoza's non-dualist parallelism, according to which there is only one kind of substance that is neither material nor immaterial, though has both material and immaterial attributes. Hume remains committed to a form of property dualism that distinguishes between objects that belong to two fundamentally different ontological existences; some perceptions have extension, size and figure, and are located in the brain while others lack all these qualities and are "no where." But what is notable about Hume's own theory is that in spite of the fact that it rejects both materialism and immaterialism, it simultaneously incorporates both material and immaterial elements into his theory of mind. The outcome of this synthesis is the rise of a new alternative theory of mind. Of course, this has given way to diverse interpretations of Hume's theory. Some commentators, like John Bricke, argue that Hume endorses the type identity theory of mind (42). Whereas others, most notably Daniel Flage, have argued that Hume adheres to a rather sophisticated form of dualism, a dualism of both entities and systems (43). Pitson also ascribes a two-fold dualism to Hume that I should briefly explain now.

Pitson has suggested that Hume's inclusion of both materialism and immaterialism into his theory of mind results in "a kind of twofold dualism: between the spatial [visual and tactual] and non-spatial [non-visual and non-tactual] perceptions which make up the mind, and between perceptions themselves and the sensible qualities which constitute the body" (44). To elaborate, he argues that "[i]mplicitly, at least, Hume appears to distinguish both extended as well as unextended perceptions from the body on whose existence and functioning they depend" (45). More precisely, on the one hand, there

is a dualism between the extended and unextended perceptions, and on the other, there is a dualism between the unextended perceptions and the body on whose existence and functioning they depend upon. Phillip D. Cummins and Flage have attributed a version of mental/physical dualism to Hume on the ground of his distinction between extended and unextended perceptions ⁽⁴⁶⁾. Pitson rejects this interpretation, and I believe he is correct to do so. As I will argue shortly, Hume seems to be committed to the idea that all thoughts and perceptions, extended and unextended, have material causes. Pitson's solution to this problem is imaginative. However, it suffers problems of its own.

Pitson suggests that we can avoid the problem related with Hume's twofold dualism, if we consider Hume's claim that some of our perceptions are extended in a non-material sense. Basically, we should consider the possibility that 'extension' in this sense could mean that "some perceptions might have *phenomenal* extension, to the extent that they occupy a certain portion of the corresponding sensory field" ⁽⁴⁷⁾. He supports his interpretation by appealing to contemporary theories about our experiences of colour in which typically the colour-appearances of things take up various expanses of our visual field ⁽⁴⁸⁾. Strictly speaking, the outcome is this: no perceptions are literally "classified as something which exists with any place" ⁽⁴⁹⁾ having size, shape and location. Thus, "perceptions in general belong to that category of things which, according to Hume, may exist and yet be nowhere" ⁽⁵⁰⁾. In this way, Pitson reduces Hume's extended category of perceptions into his unextended class of perceptions. In a nutshell, he "suggests that the difficulties associated with Hume's twofold dualism might be avoided by applying his principle that something may exist without any place to perceptions in general, while recognizing that all sensible qualities are space-occupying even if there are some which lack any determinate location (or 'particular place' as Hume puts it)" ⁽⁵¹⁾.

There are a few problems with this interpretation. First, as Pitson himself confesses, this would deprive Hume of his main objection against immaterialism (i.e., if all perceptions are nowhere it means that materialism is mistaken) ⁽⁵²⁾. Secondly, this interpretation is unsympathetic to the text. Hume never wavers in his view regarding the extension of visual and tactual perceptions laid out in the section of "Of the Immateriality of the Soul" in the *Treatise* or elsewhere. Finally, and more importantly, Pitson's interpretation is only able to resolve the first kind of dualism that he ascribes to

Hume—that is, the dualism between the spatial and non-spatial perceptions. It just reduces the first kind of dualism into the second one. It leaves the dualism between the unextended perceptions and the body that they depend upon intact. At best, on Pitson's account, Hume is a 'property dualist.' Nonetheless, it seems to me that we do not need the kind of reductionism that Pitson appeals to rescue Hume. Anderson has convincingly shown us, I think, that there is an irresolvable inconsistency in Hume's account of the location of perceptions (53). As Anderson has said, if Hume had understood the notions of 'extension,' 'shape,' 'size,' and 'location' of perceptions *equivocally*, as Pitson understands, he would be making no point at all against either materialists or immaterialists (54). We can conclude, therefore, that Hume employs 'extension' and other terms *univocally* throughout. If this is true, then, Pitson's interpretation is incorrect.

So far I have attempted to explain Hume's position in relation to materialism and immaterialism. It is shown that he rejects both materialism and immaterialism, though he incorporates elements of both into his own theory of mind. Furthermore, he seems to be committed to a form of property dualism. From what I have argued so far, however, it is not clear what Hume's own position is on the difficult question of the relation of mind and body. Specifically, there are two main issues that need be settled. First, the difficult ontological question concerning the nature of mind remains to be resolved, and second, I shall explore Hume's views on the question regarding the metaphysics of mind and its relation to the ontology of the mind. The thesis that I will be trying to defend is this: Hume answers the metaphysical question without answering the ontological question. In contemporary philosophy of mind this is fundamentally a functionalist project (55). This seems to be Hume's programme as well.

It is my view that Hume is not only aware of the inconsistency that Anderson attributes to him, but also, I suggest, he knows that the inconsistency may not be resolved at all. As pointed out above, Hume seems to think that we are not in a position to resolve the ontological question concerning the nature of mind. Of course, he allows that we may well be in a position to demonstrate the error, or fallacy, involved in both materialism and immaterialism, but we cannot decisively make any claim about the ontology of the mind on this ground. Before dealing with this error, I briefly explain the distinction between the ontology and the metaphysics of the mind.

Functionalism distinguishes between the ontology and the metaphysics of mental states, and rejects the ontological claim without rejecting the metaphysical one. Accordingly, theories of the mind prior to functionalism have been concerned with both the ontology—what there *is*—and the metaphysics—what gives each type of mental state its own identity, for example what pains have in common in virtue of which they are pains—of mental events. Substance dualism claims that there are both mental and physical substances, while behaviourism and physicalism are monistic, claiming that there are only physical substances ⁽⁵⁶⁾. In fact, to speak about ontology here is to speak about either the existence of two distinct substances, namely, mental and physical, or the existence of one single substance, i.e., the physical. These are the ontological claims. On the contrary, concerning the metaphysical claims, we may say that, “behaviourism tells us that what pains (for example) have in common in virtue of which they are pains is something behavioural; dualism gave a non-physical answer to this question, and physicalism gives a physical answer to this question” ⁽⁵⁷⁾. What happens in theories of mind prior to functionalism is that they answer both metaphysical and ontological questions together. Basically, answering the ontological question depends upon answering the metaphysical question and vice versa. For example, Descartes could not maintain that ‘pain’ is a non-physical state if he had not already established that there are two fundamentally distinct ontological substances. Similarly, physicalism would not be able to claim that, for instance, pain = C-fibre, if it had not already reduced everything into the physical. To put it differently, to be able to say that pain = C-fibre requires that we already establish that there is only one single physical substance. Nonetheless, functionalism answers the metaphysical question without answering the ontological one. It says that, for example, what pains have in common is their *function*. But it does not specify if the subjects that have pains are non-physical or otherwise ⁽⁵⁸⁾. So what is in common to human pain and Martian pain is that pain could be realized either in our body or in a Martian regardless of their physical descriptions. In fact, the functional state is said to be neutral to the physical description of the body, i.e., to what it is made of; “it does not tell us how the machine works or what it is made of, and in particular it does not rule out a machine which is operated by an immaterial soul, so long as the soul is willing to operate in the deterministic manner specified in the table” ⁽⁵⁹⁾. Similarly, it allows that

a functional state may be caused by a physical human brain or silicon-based computer regardless of the fact that how they work or what they are made of. Hence, functionalism is impartial/neutral about the ontological question.

To return to Hume, the reason that I think he rejects the ontological claims of both the immaterialists and the materialists is that they both claim that they know the substance of the mind—that is, in their views the mind is either two fundamentally distinct ontological substances or a single substance. Hume, however, rejects the idea that we can know the substance of the mind; all that we can know, or are determined to know, is the causes of our thoughts and perceptions. For “the essence of the mind [is]...unknown to us” (THN, xvii); we only know the causes of its perceptions. This implies that we are not in a position to make any claim about the ontology of the mind. Epistemologically speaking, our lack of knowledge about the substance of things means that we are not in a position to say if there are either two distinct types of existences, namely, the mental and the physical, or just one kind, i.e., the physical, in the world. And not being able to talk about the substance of the mind is, indeed, the same as not being able to talk about its ontology. If Hume is unable to determine the substance of the mind—whether it is a physical or non-physical entity—he is not, by the same token, in a position to determine the ontological status of the mind. For Hume, to be completely shut off from the substance of things entails that there is no solid ground for distinguishing between the material and the immaterial realities. Hence, he holds that we can neither afford to maintain that there are two separate and distinct physical and non-physical substances, as the immaterialist does, nor are we able to reduce all mental states and processes into crude matter, as the materialist does. Therefore, Hume sees himself in a conundrum and, as a result, dismisses the intelligibility of the question concerning the substance of the mind. Consequently, he dismisses the ontological question concerning the mind.

Nevertheless, as I will show shortly, Hume unequivocally argues that we could know the cause of our thoughts or perceptions. For him, to know what has caused our thoughts and perceptions is, indeed, to know the nature of mind. Hume’s central thesis is that the existence of causal relations among perceptions is the key for understanding the nature of mind. More precisely, it is a question about knowing the origins of our ideas and the way that perceptions in the mind cause new ideas. In this sense, we are able to

determine the nature of mental events and differentiate them from the physical on the basis of our empirically warranted perceptual experiences of them. We may not know the essence of things, or the things-in-themselves (to use Kantian terminology), but we can certainly know the mind or self as far as we have empirically warranted perceptual experience of it, and we can distinguish it from the physical on that ground. The same is true about the physical or the external objects insofar as we have empirically warranted perceptual experiences of them. Hence, in light of our perceptual experience of reality, we may readily allow, as Hume does, that there are two sorts of different ‘properties’—the mental and the physical—in the world. This constitutes Hume’s so-called ‘property dualism.’ One may argue that this implies a kind of ontology insofar as it says that there are two sorts of properties in the world, and thus, Hume is committed to an ontological position here. Hume makes a weak ontological claim when he asserts that everything in the world, including the human mind, is ultimately *physical*, or depends on the physical. This view is most vividly presented in his essay entitled “Of the Immortality of the Soul.” This is also evident in his argument that matter and motion can cause thought and perception. Hume does believe that everything in the world is made up of entirely minute particles that compose matter. In this sense, Hume is committed to materialism.

To return to Hume’s metaphysics of mind, he distinguishes between three kinds of questions about our perceptions in general. The first is about *substance*, the second is about *local conjunction*, and the last is about the *cause of our perceptions* (THN, 246). Hume’s verdict on the question of the substance of the mind is that it should be condemned, as absolutely unintelligible. “All our perceptions are not susceptible of a local union,” he argues, “either with what is extended or unextended; there being some of them of the one kind, and some of the other” (THN, 250). If we were able somehow to render all our perceptions as either material or immaterial, then, we would have been able to determine the substance of the mind. But since, on Hume’s account, that is something impossible to arrive at, it follows that we should dismiss the question of the substance of the mind altogether. The essence of the mind is unknown to us (THN, xvii). Provided that one is unable to determine the substance of a thing, it can hardly be the case that one knows what there *is* or what there *is* not. And because we are unable to determine the substance of the mind, then we have no choice but to dismiss the ontological question

concerning the nature of mind altogether.

Hume claims, however, that there is a philosophically valid question concerning the compatibility or incompatibility of thought and perceptions with extension—that is, their local conjunction with matter (THN, 235). Here, as I explained above, he shows that some perceptions, namely, sight and feeling, are capable of local conjunction and those other perceptions, which “*may exist, and yet be no where,*” are “incapable of any conjunction in place with matter and body, which is extended and divisible” (THN, 236). Hume treats the question of local conjunction as a metaphysical problem. “It may be better worth our while to remark,” he claims, “that this question of local conjunction of objects does not only occur in *metaphysical disputes* concerning the nature of the soul, but that even in common life we have every moment occasion to examine it” (THN, 236; emphasis added). But more importantly, the reason that may account for Hume’s metaphysical concerns here is his psychological explanation of the idea of ‘local conjunction.’ According to Hume, the same natural tendency of the mind, which is responsible for generating the ideas of necessary connection, object identity, and personal identity, feigns the idea of local conjunction from the relations of *causation* and *contiguity in time* between two objects (THN, 236-240). According to Hume, an extended object is incapable of local conjunction with an unextended object that exists without any place or is ‘no where.’ In other words, there is no *contiguity in place* between the two objects. Nevertheless, there is the relation of causation and contiguity in time between the extended object and the unextended perceptions (THN, 237). The *inclination* of our fancy makes us to confuse these relations of causation and contiguity in time with that of contiguity in place; they “must have such an effect on the mind, that upon the appearance of one it will immediately turn its thought to the conception of the other” (THN, 237). Consequently, “we not only turn our thought from one to the other upon account of their relation, but likewise endeavour to give them a new relation, *viz.* that of a *conjunction in place*, that we may render the transition more easy and natural” (THN, 237). In fact, Hume argues that the idea of local conjunction, like the idea of the self, is a fictional one because we confuse, or mistake, the relations of causation and contiguity in time with the relation of contiguity in place.

The psychological mechanism producing the idea of local conjunction parallel’s

with the one that produces the ideas of the external object and the self (THN, 237). As I briefly pointed out earlier, Hume argues that the bitter taste of an olive or the sweet taste of a fig, for instance, cannot be contained in the circumference of its body. We cannot say that these qualities are only in a particular part because experience tells us that every part has the same taste. Similarly, we cannot say that they are in every part because we, then, have to regard them as figured and extended, which is absurd and incomprehensible. As a result, we either must suppose that some beings exist without any place; or that they are figured and extended; or that when they are incorporated with extended objects, the whole is in the whole, and the whole in every part ⁽⁶⁰⁾. Hume concludes that the latter supposition is absurd, and it sufficiently proves the veracity of the former (THN, 239). Hume argues that although extended objects are “incapable of a conjunction in place with another [perceptions], that exists without any place or extension, yet are they susceptible of many other relations” (THN, 237). For example, Hume claims that “the taste and smell of any fruit are inseparable from its other qualities of colour and tangibility [sight and touch]; and which-ever of them be the cause or effect, ’tis certain they are always co-existent” (THN, 237). He goes on to argue that “[n]or are they only co-existent in general, but also co-temporary in their appearance in the mind; and ’tis upon the application of the extended body to our senses we perceive its particular taste and smell” (THN, 237). According to Hume, the mere co-existence and co-temporary appearance of the smell and taste of these fruits with the qualities of colour and tangibility indicate that there are relations of *causation* and *contiguity in time* between the extended objects and the unextended perceptions, i.e., smell and taste, that exist without any place. However, Hume thinks that the mere co-existence and co-temporary appearance of these qualities and perceptions in the mind do not amount to *contiguity in place* between them. Hume claims, then, that causation and contiguity in time “must have such an effect on the mind, that upon the appearance of one of it will immediately turn its thought to the conception of the other” (THN, 237). In consequence, from the relations of causation and temporal contiguity, “we feign likewise that of a conjunction in place” (THN, 238) in order to render the transition more easy and natural. Of course, Hume’s claim that we confuse the relations of causation and temporal contiguity with that of spatial contiguity, might be considered problematic in its own right. But, as he indicates, the mechanism responsible

for producing the fictional idea of local conjunction is the same psychological mechanism that generates the fictional ideas of bodies and selves. Hume conceives this psychological belief-forming mechanism as a genuine metaphysical inquiry (THN, 236).

For Hume, the question of local conjunction of perceptions does not address ontological concerns, even though he arrives at a conclusion pertaining to the ontological status of perceptions claiming that some perceptions are locally conjoined and some others are incapable of contiguity in place. Nevertheless, Hume's main objective is to show that the ontological question concerning the materiality and the immateriality of the mind leads us to a standoff between materialism and substance dualism. Hume wants to avoid this standoff even at the cost of the ontological question itself. Furthermore, Hume's interest for the question of local conjunction of perceptions is of particular concern about the relation between thought and extension. The question of local conjunction of perceptions is directed at replacing the question concerning the substance of the mind with the question of how mental states may be conjoined with matter. For Hume, that question is meant to explain different types of mental states in terms of the properties in virtue of which they are mental states. To put it differently, Hume is interested to account for the nature of mental states despite their being locally conjoined (visual and tactual perceptions) or their lack of local conjunction. In order to do so Hume tries to show how *experience*, rather than reason, provides an explanation for the nature of our perceptions in general. According to him, experience shows a constant conjunction and temporal contiguity between the extended and the unextended objects, and that sufficiently proves the existence of causal relation between mind and body. Hume clearly states that all mental states, including those perceptions without any particular place, have physical/material causes. This brings me to Hume's last question, namely, the question concerning the cause of our perceptions.

Having distinguished between these three questions concerning our perceptions in general, Hume states that the last one, i.e., the cause of our perceptions, is the most crucial one. In spite of the fact that some of our perceptions, or mental states, are incapable of conjunction in place, Hume argues that the relation of temporal contiguity and causation function as the material cause of our perceptions. Hume seems to be committed to the idea that thoughts and perceptions are caused by material motion (61).

The key principle for understanding Hume's view here is the distinction that he makes between the relation of causation and temporal contiguity, on the one hand, and the relation of spatial contiguity on the other. Here he argues against the Scholastics, who concluded that it is impossible that thought and perception can ever be caused by matter and motion, "that we must separate the question concerning the substance of the mind from that concerning the cause of its thought; and that confining ourselves to the latter question we find by the comparing their ideas, that thought and motion are different from each other, and by experience, that they are constantly united; which being all the circumstances, that enter into the idea of cause and effect, when apply'd to the operations matter, we may certainly conclude, that motion may be, and actually is, the cause of thought and perception" (THN, 248). Hume proceeds to remark that since we are never sensible of any real connexion between causes and effects except by that of our experience of their constant conjunction it must follow that our knowledge of cause and effect in the current situation depends entirely on experience of constant conjunction between thought and perception and matter and motion. And since all objects, which are not contrary, are susceptible of a constant conjunction, there is no reason why any object may or may not be the cause of any other. Consequently, "to consider the matter *a priori*, any thing may produce any thing, and that we shall never discover a reason, why any object may or may not be the cause of any other, however great, or however little the resemblance may be betwixt them" (THN, 247). The conclusion that Hume draws here is that there is basically nothing that precludes the possibility of a causal relation between thought and matter. Furthermore, he argues that the real question in this matter is that whether we are having, in fact, the experience of a relation of constant conjunction among thought and motion. "Nay 'tis not only possible we may have such an experience," Hume says, "but 'tis certain we have it; since every one may perceive, that the different disposition of his body change his thoughts and sentiments. And shou'd it be said, that this depends on the union of soul and body" (THN, 248). In a nutshell, "having argued that the absence of any intelligible connection between thought and motion does not prevent them from being causally related, Hume goes on to claim that experience reveals that motion 'actually is, the cause of thought and perception'" (62).

Hume discusses this aspect of his theory elsewhere as well. In an essay entitled,

“Of the Immortality of the Soul” he rejects *a priori* metaphysical and moral arguments for the immortality of the soul and concludes that the only philosophical valid argument concerning this matter is, indeed, the physical arguments from the analogy of nature, as *a posteriori* argument, which conclusively proves the mortality of the soul. There he argues that “[w]here any two objects are so closely connected, that all alterations, which we have ever seen in the one, are attended with proportionable alterations in the other; we ought to conclude, by all rules of analogy, that, when there are still greater alterations produced in the former, and it is totally dissolved, there follows a total dissolution of the latter” (STE, 165). He claims:

Sleep, a very small effect on the body, is attended with a temporary extinction, at least a great confusion in the soul. The weakness of the body and that of the mind in infancy are exactly proportioned, their vigour in manhood, their sympathetic disorder in sickness; their common gradual decay in old age. The step further seems unavoidable; their common dissolution in death. The last symptoms which the mind discovers are disorder, weakness, insensibility, and stupidity, the fore-runners of its annihilation. The farther progress of the same causes increasing, the same effects totally extinguish it. Judging by the usual analogy of nature, no form can continue when transferred to a condition of life very different from the original one, in which it was placed. Trees perish in the water, fishes in the air, animals in the earth. Even so small a difference as that of climate is often fatal. What reason then to imagine, that an immense alteration, such as is made on the soul by the dissolution of its body and all its organs of thought and sensation, can be effected without the dissolution of the whole? Every thing is in common betwixt soul and body. The organs of the one are all of them the organs of the other. The existence therefore of the one must be dependent on that of the other. (STE, 166)

Hume concludes:

Nothing in this world is perpetual, every thing however seemingly firm is in continual flux and change, the world itself gives symptoms of frailty and dissolution. How contrary to analogy, therefore, to imagine that one single from, seemingly the frailest of any, and subject to the greatest disorders, is immortal and indissoluble? What daring theory is that! how lightly, not to say how rashly entertained! (STE, 166)

Here, Hume makes his main metaphysical claim about the nature of mind according to which the mind is in continual flux and change and subject to great disorders and, thus, dissoluble and mortal. He claims that everything in the world is *physical*.

Moreover, Hume seems to be suggesting that mental states have physical—or, more specifically, neuro-physiological causes. He has expressed this view on different

occasions. First, in his examination of the causes of both impressions of sensation and ideas, Hume claims that they depend upon certain traces in the brain. Here he discusses the physiological functioning of the animal spirits and the way that they excite ideas in the brain (THN, 61). Second, Hume implicitly, on certain occasions, indicates that mental phenomena can generally depend upon neural events (THN, 269, 270). Besides, he maintains that certain mental states can generate physiological changes in the body. This is best shown in his treatment of voluntary actions and their connection with the motion of the body (THN, 632). This establishes that, for Hume, there is a two-way causal interaction between the mental and the physical. Hume's notion of mental causation is best shown in his rejection of the occasionalist view of Malebranche who claimed that "it is not any energy in the will that produces local motion in our members: It is God himself, who is pleased to second our will, in itself impotent, and to command that motion which we erroneously attribute to our own power and efficacy" (EHU, 71). In reply, Hume tells us that if there were not any energy or power in ourselves producing the motion of the body, then we are not in a position to ascribe that faculty to the supreme mind itself (EHU, 64-5). Hume argues that volition can give rise to action through various physical or physiological events which do not themselves form the immediate object of volition (EHU, 66). This reflects Hume's commitment to the idea of mental and physical causation.

To sum up, Hume's reason for thinking that perceptions and thoughts have material causes stems from three main sources: First, he thinks that the extended perceptions have size, figure and location. Second, he holds that causation and temporal contiguity between the unextended perceptions function as the basis of material cause of thoughts and perceptions. Finally, he argues that experience shows that bodily changes affect our thoughts and sentiments. According to him, all perceptions—extended and unextended—are caused by motion in the body. This line of thought has important ramifications on his theory of mind, which I will explain shortly. Before doing this, I want to consider an objection.

So far, on the one hand, I have attributed a form of 'property dualism' to Hume, and on the other, I have argued that he is committed to a kind of physical-to-mental causation. This gives rise to a natural objection. It is often believed that dualism comes

into conflict with the possibility of mental and physical causation, and because I have been ascribing both positions to Hume at the same time it means that my reading is facing difficulty. There is a response to this objection. Hume argues that the notion of causation is based on the idea of ‘constant conjunction’ between cause and effect, rather than the existence of some real connection among them. Pitson, for instance, notes that “[t]he Humean response to this, however, would surely be that the idea of identifying something which would enable us to understand how certain events are causally related itself reflects a mistaken ‘rationalistic’ conception of the relation of cause and effect. All that matters here is whether the events in question are constantly conjoined—there does not have to be, and in fact there never is, any discoverable connection between them. Thus, ‘...tho’ there appear no manner of connexion betwixt motion or thought, the case is the same with all other causes and effects’ (T, 1.4.5.30)”⁽⁶³⁾. Moreover, the worry about the incompatibility of mental causation and dualism arises from an assumption about the nature of causation in the physical world—that is, causation is a matter of flow of energy or force from one object or event to another. But we know that, on Hume’s account, “we have really no idea of a power or efficacy in any object, or of any real connexion betwixt causes and effects” (THN, 168). All that we have is the idea of necessary connection, and this idea of necessary connection consists in the constant conjunction of certain objects as well as the ‘determination of the mind.’ These two elements constitute physical necessity. In fact, power or efficacy, as certain qualities or the determination of the mind, is transferred to external objects (THN, 168). Therefore, the causal relation that exists between matter and thought does not need to be based on real connection between events. Rather, the mere existence of constant conjunction between cause and effect along with the ‘determination of the mind’ is sufficient to causally relate thought and matter. Hume is essentially concerned to show how the standard understanding of the notion of causation, which assumes that causal relations must be rationally intelligible, is mistaken. According to Hume, it is experience rather than reason that determines when a causal relation obtains, and experience, at most, tells us about constant conjunctions between different kinds of objects and events.

To return to my discussion, Hume establishes that perceptions in general are caused by physical properties one way or the other. Some of them are caused by physical

properties while locally/spatially conjoined, and some others are caused by physical properties while lacking local conjunction. However, this does not mean that all mental states (perceptions) are identical with some physical properties in the brain because we are told that most of our perceptions lack local conjunction. Furthermore, since the causal relation, on Hume's account, requires only the existence of constant conjunction between cause and effect, which is strictly supplied by experience, it implies that the existence of the relation of spatial contiguity is not necessary to connect cause and effect. Perhaps, one could attribute a kind of partial type identity theory to Hume insofar as he allows that visual and tactual perceptions are extended and have size, shape, and location, but it seems that this kind of interpretation is inconsistent with Hume's theory of mind in general. On the contrary, I suggest that Hume's intent for claiming that extended perceptions are spatially located has to do with his strategy to refute the immaterialist claim concerning the substance of the mind. Most likely, Hume thinks about the immaterialists as his main contenders, rather than the materialists who were right, at least, in thinking that everything in the world is ultimately caused by some type of physical property. And if all perceptions are caused by some physical properties, why do we need to distinguish between the extended and unextended perceptions? Why not just materialism?

Hume may have intended to show that a causal, or functional, theory of mind is unbiased or neutral to the physical/non-physical basis, or the make-up of the functional system as whole. Contemporary literature on functionalism acknowledges that functional description, whether machine systems or otherwise, is impartial to how the system works or what it is made of. Accordingly, even an immaterial soul can be a functional system as long as it is willing to operate in a deterministic manner specified by the functional input-output description (64). A pain is a pain in virtue of the functional/causal description that it has regardless of the fact that whether it is generated by the human brain or by a silicon-based computer. Hume's view that some perceptions are spatially located and some others are not is quite symptomatic of the fact that his causal, or functional, theory of the mind is impartial to the make-up of the functional system. The fact that Hume's notion of physical-to-mental causation does not require spatial contiguity as the necessary condition connecting cause and effect, strongly suggests that, for Hume, the make-up or

the physical/non-physical basis of our mental states is entirely irrelevant to the functionality of the system as a whole. Consequently, we may suggest that, in Hume's view, all our mental states are functional states whether they are physically/spatially located in the brain or they lack local conjunction as is the case of unextended perceptions. In short, the distinction that Hume makes between the extended and unextended perceptions not only fits squarely within the boundaries of functionalism, but it may also indicate how Hume ingeniously sorted out perceptions in order to show that a functional/causal theory of mind could be produced without falling victim to the mistaken views of both materialism and immaterialism.

I would like to end this section by quoting an extremely remarkable passage by Pitson. He argues:

Hume thus arrives at a version of dualism which arguably reflects much of our common sense thinking about mental states and their relation to the body: in particular, that what are involved are different kinds of thing—but not different kinds of *substance*, in the philosophical sense of this notion—which nevertheless stand in causal relations to each other. At the same time, the difficulties which are often thought to attend this kind of position are met by certain crucial principles which explain why, for Hume, there is really no such a thing as the mind/body *problem* but only questions about the relation between the different kinds of state involved which are to be settled by reference to experience. (65)

This statement, I think, captures the spirit of Hume's motivation for introducing a new functionalist theory of mind, which is supposed to resolve the mind/body problem.

To sum this section up, we have seen that Hume finds the question as to whether perceptions inhere in a material or immaterial substance unintelligible. This claim of course should not come as a surprise to us insofar as Hume firmly rejects the intelligibility of the notion of substance. However, he considers the question concerning the local conjunction and the cause of our perceptions very important. In regard to the question of local conjunction, he establishes that some perceptions may exist without any place. What concerns Hume here is the issue of the relation of thought and extension. He holds that our perceptions in general lack extension. However, this does not cause any problem concerning our perceptions relation to extension insofar as experience shows that causation and temporal contiguity can supplement that of a conjunction in place. Overall, Hume intends to show that the materialist view, which holds that there is a

genuine possibility of our perceptions being spatially located, encounters the difficulty of explaining how perceptions have form, size, and quantity. In respect to the cause of our perceptions, Hume identifies matter and motion as the cause of our thoughts and perceptions. He claims that experience, rather than reason, determines the causal relation between body and mind. This has an important implication for Hume's view on the mind/body problem. It implies that there is more to the question of mind/body problem than that of the material cause of our mental states. One thing that, for instance, it says is that there is no obstacle to recognize a causal relation between mind and body, given that the idea of necessary connection and that of the constant conjunction between different objects or events is revealed to us by experience. And insofar as experience adequately accounts for the genuine possibility of our mental states lacking local conjunction, it follows that *experience* alone must determine the cause of our perceptions.

Moreover, I argued that Hume's distinction between these questions provides a context for my claim that Hume dismisses the ontological question concerning the nature of mental states. In fact, Hume's rejection of both materialism and immaterialism, and as a result, his dismissal of the question concerning the substance of the mind, indicates that he is convinced that the ontological question is irresolvable. However, since he holds that experience establishes the existence of causal relation between mind and body, it will allow us to answer the metaphysical question without answering the ontological one. In a nutshell, my main contention is that Hume's distinction between these questions and the particular emphasis that he places on the question of the cause of our perceptions could be easily translated into contemporary functionalist view which says knowing the nature of mind does not require our knowledge of what it is made of.

Conclusion

We are now in a position to review some of the ideas that presented in this chapter. First, I wanted to show that Hume, like Locke, was writing in the tradition in which the problem of personal identity was seen as that of giving an account of what *constitutes* personal identity. Second, I outlined Hume's characterization of his theory of the self as a fiction. The main idea behind his fictional theory of personal identity was that the idea of the self is produced by the psychological mechanism of the imagination. As explained, Hume's theory of the self has a negative phase and a positive phase. In the negative

phase, Hume wants to establish that personal identity is neither constituted by the identity of a simple material or immaterial substance nor it is derived from the identity of a simple impression. In the positive phase, he shows that personal identity is constituted by the *identity of function*. Then, I argued that this identity of function adequately accounts for the specification of the conditions of personal identity over time. Finally, I discussed Hume's views on the question of the nature of mind. I ascribed a version of 'property dualism' to him. I suggested that Hume distinguishes the ontological aspect of his theory from its metaphysical claim, and that he rejects the ontological claims about the nature of the mind altogether. Instead, he shifts the focus onto the metaphysics of mental states. As a result, he answers the metaphysical question without answering the ontological question concerning the nature of mind. To substantiate this claim, I have shown that Hume dismisses the question of the substance of the mind because he thinks that it is unintelligible; rather, he argues that there are significant philosophical advantages in reflecting upon the question concerning local conjunction of our perceptions and the question about the cause of our perceptions. The main implication of Hume's dismissal of the ontological question was that his theory maintains a neutral/unbiased position about the physical/non-physical make-up of the functional system. As a consequence, to know the nature of mental states, for Hume, is to know their causal/functional roles rather than knowing their spatial location or their material make-up. We may know something about the location, size, figure, and extension of some perceptions, but this cannot sufficiently answer our question about the nature of the mind. On the contrary, it shows how helpless both the materialists and the immaterialists are in the face of the question. In particular, Hume thinks that their philosophical endeavours to settle the ontological question is simply unproductive.

What we can gather from these is that Hume's philosophy of mind seems to succeed in setting out the groundwork of a causal/functional theory of mind. He characterizes the mind as a system of causal connection in which each perception is identified in terms of the role that it plays in the perceptual system that it is a member of. He also establishes that some sort of identity of function adequately accounts for our continuity over time. In the next chapter, I turn to Hume's computational theory of mind.

II

Hume's Computational Theory of Mind

That the experimental reasoning, which we possess in common with beasts, and on which the whole conduct of life depends, is nothing but a species of instinct or mechanical power that acts in us unknown to ourselves; and in its chief operations, is not directed by any such relations or comparisons of ideas, as are the proper objects of our intellectual faculties. (EHU, 108)

The central claim of this work is that Hume's theory of the self can be interpreted, or reconstructed, in terms of the functionalist theory of mind. In a broad sense, functionalism defines mental properties or states like a desire or a pain solely in terms of its function, or the role that it plays in the cognitive system that it is a member of. Functionalism identifies mental states in terms of what they do rather than by what they are made of. Functionalism remains neutral about the internal constitution of mental properties. In this respect, functionalism takes the nature or identity of a mental state to be determined by its causal relations to sensory inputs, other mental states, and behavioural outputs. The claim that Hume's theory of the self can be interpreted, or reconstructed, on causal and/or functional grounds is a thesis about Hume's characterization of mental particulars (impressions or ideas) in terms of the kind of roles that each plays in our cognitive system. Hume's characterization of mental states and processes as defined by their causal/functional roles in a cognitive system finds its natural place in his causal theory of mind. Hume claims that "the true idea of the human mind, is to consider it as a system of different perceptions or different existences, which are link'd together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other" (THN, 261). In fact, Hume here argues that the best empirical theory about the nature of mind is the causal theory that is the result of a complex of mental states and processes, produced, modified, influenced, and individuated in terms of the roles that each play in the cognitive system that it is part of. He also recognizes that the mind as defined by functionalism consists of a complex causal system of mental states,

events, processes, etc.

As indicated earlier, I distinguished between two different senses of functionalism in Hume's theory of mind: teleofunctionalism and psycho-functionalism. The former is rooted in his teleological characterization of the mind according to which the bundle of perceptions persists over time by maintaining functional continuity, similar to the kind of functional continuity which constitutes the identity of organic objects and artifacts. I have discussed Hume's teleological functionalism in the previous chapter. The main source of Hume's psycho-functionalism lies in his Theory of Ideas or his Representational Theory of Mind (RTM). One of the most important features of psycho-functionalism is that it identifies mental states and processes as defined by their roles in a cognitive psychological theory. Indeed, what is distinctive about this feature of psycho-functionalism is its claim that mental states and processes are postulated by the best scientific explanation of human behaviour. And that is why psycho-functionalism is primarily derived from the methodology of cognitive psychological theories. There are substantial evidences which show that Hume adopts the same methodology of cognitive psychology in his characterization of mental states and processes as defined by their causal role in the mind. I imagine that Hume's psychological analysis of the human mind and his introduction of the experimental method of reasoning into moral subjects show that his commitment to the methodology of cognitive psychology is robust. In addition, Hume's articulation of a new kind of philosophy, an experienced-based *science of human nature*, indicates that, in his view, the best explanation of the human mind, and its behaviour, is the one that his science of human nature provides us with. In this respect, Hume's science of man not only has a lot in common with contemporary 'cognitive science,' but is also in some respect anticipation of, and superior to, the current theories in the tradition of cognitive science.

In this chapter, I want to examine Hume's psycho-functionalist theory of the mind. RTM is a major theme in contemporary psycho-functionalist theory of mind, and in what follows, I will show that it is a major theme in Hume's causal/functional theory of the self as well. Of course, as the title of the current chapter indicates, I have attributed a Computational Theory of Mind (CTM) to Hume. CTM, which is a leading version of RTM, claims that the brain is a kind of digital computer and mental states and processes

are simply computation, or formal symbol manipulation. According to CTM, all psychological states and processes are constituted by computational relations—the syntactic manipulation of symbols—to mental representations. Hume never thought that the brain is a kind of digital computer, or that mental states and processes are constituted by their purely formal syntactic manipulation of symbols. However, since CTM combines RTM with a computational account of reasoning in which it purports to show how reasoning could be a sort of causal process, this shows a link between Hume and CTM. Moreover, while CTM differs from Hume’s version of representationalism which likened perceptions to pictures rather than symbols, it is philosophically important to see how Hume anticipates an early version of CTM. In Hume’s version of CTM, instead of ‘computation’ or ‘formal symbolic manipulation’ we are dealing with an iconic mode of representation. Also, Hume’s version of CTM includes topics about the mechanics of the mind and the psychological mechanism of imagination. This chapter is an investigation of these issues.

In II.1, I discuss Hume’s mechanics of the mind. I want to show that, for Hume, mental states and processes are essentially automatic, unreflective, and mechanical, occurrences that cannot be justified on the basis of reason or experience; rather, they are brute and unavoidable facts of the mind. In II.2, I will show that the idea of causal inference that is fundamental to Hume’s philosophy is produced by the psychological mechanism of imagination. In II.3, I will examine Hume’s views on the architecture of the mind. I want to show that Hume is committed to the sub-doxastic and modular view of the architecture of the mind. In II.4, I will defend Hume’s version of RTM.

II. 1: Hume’s Mechanics of the Mind

Hume is a philosopher of human nature ⁽¹⁾. He asserts that human nature is the only science of man that provides a solid foundation for the other sciences (THN, xvi). He characterizes his science of man as examining the mind to discover its most secret springs and principles (THN, xviii). Although the truth of these springs and principles “must lie very deep and abstruse” (THN, xiv), we may hope to arrive at it by “anatomizing human nature in a regular manner” (Abs., 2). As “an Anatomist”, Hume is supposed to explain “all the operations of the mind” according to “The Experimental Method of Reasoning.” This method requires “careful and exact experiments” in order to “render all our

principles as universal as possible.” The experimental method of reasoning requires “tracing up our experiments to the utmost, and explaining all effects from the simplest and fewest causes” (THN, xvii). This new experimental science of man may “discover, at least in some degree, the secret springs and principles, by which the human mind is actuated in its operations” (EHU, 14). In short, anatomizing about human nature is to discover the laws governing the operations of the mind. This is the positive phase of Hume’s teaching—his naturalism.

However, these governing laws and “ultimate springs and principles are totally shut up from human curiosity and enquiry” (EHU, 30). Hume claims that “we can give no reason for our most general and most refin’d principles beside our experience of their reality” (THN, xviii). Hume explicates “all the operations of the mind” by means of the three universal principles of association, i.e., resemblance, contiguity, and causation. He asserts that we must recognise their existence because their “effects are everywhere conspicuous; but as to its causes, they are mostly unknown, and must be resolv’d into *original* qualities of human nature” (THN, 13). This is the negative phase of Hume’s teaching—his scepticism. Nevertheless, his scepticism about the possibility of metaphysical explanation of these laws and principles must not make us sceptical about the possibility of a science of mind (2). Through observation and experience, a careful analysis and description of the origins of the ideas, as well as of the operations of these principles, we can overcome the temptation to scepticism. Specifically, Hume’s scepticism is directed at speculative metaphysics, which claims deeper and more certain metaphysical knowledge of reality. Similarly, he rejects dogmatic scepticism because it denies the possibility of scientific knowledge. Contrastingly, Hume shifts the focus “from a vain attempt to give a philosophical justification of our fundamental beliefs to a scientific account of their origin in the operations of our minds” (3). He calls this a “Sceptical Solution” to a sceptical challenge (EHU, 5). For Hume, this sceptical solution is itself a scientific enquiry; it explains the kinds of mechanisms or underlying principles governing the mental world. Accordingly, Hume’s scepticism concerning the ultimate laws and principles of the science of the mind does not involve the impossibility of knowledge, as dogmatic scepticism claims; rather, it asserts that these laws and principles are ultimately reduced to certain brute mental facts, or instincts, or *mechanical powers*,

that are philosophically inexplicable. We must note, however, that “this recognition of our unreflective, instinctive, and unavoidable acceptance of certain basic beliefs must not be confused with claiming to have a philosophical justification of those beliefs” (4).

Generally speaking, for Hume, mental phenomena such as the processes, states, and events in the mind are essentially unreflective, mechanical, and unavoidable occurrences that cannot be justified on the basis of reason or experience. Hume’s intention for reducing the laws and principles of the science of the mind to brute mental facts has to do with this idea that the nature of mental states is best explained as certain automatic and unreflective phenomena occurring in the mind. And the fact that mental phenomena are automatic, mechanical and unreflective occurrences in the mind, may suggest that there is no other explanation over and above the very mechanical explanation itself. The ultimate laws and principles of the science of the mind remain inexplicable because they are ultimately reduced to mechanical and automatic occurrences, which cannot be explained differently. In other words, the ultimate laws and principles of the science of the mind consist of brute mental facts—the automatic operations and functionings of the mind. Therefore, Hume’s new experimental science of man does not claim that the ultimate laws and principles of the mind are *totally* inexplicable, but that mental phenomena are best explained as brute mental facts, or automatic and mechanical occurrences in the mind. That is why Hume asserts that the new experimental science of man may “discover, at least in some degree, the secret springs and principles, by which the human mind is actuated in its operations” (EHU, 14). Supposedly, for Hume, this kind of mechanical explanation may well be the best available explanation that philosophy could offer. In fact, Hume’s rejection of the question of the substance of the mind and its substitution with the questions of local conjunction and causation, as explained in the previous chapter, indicates that, for him, any concerns about the ultimate reality of the mind, or its ontological status, is a vain question. That is why he warns us that the ultimate springs and principles of the science of the mind are totally shut up from human curiosity and enquiry. In contrast, he holds that the only legitimate question concerning the mind is a kind of scientific enquiry about the causes of our perceptions. The kind of scientific enquiry about the causes of our perceptions is, indeed, the very causal/functional explanation of these ultimate springs and principles of the science of the

mind itself. In a nutshell, it is true that these ultimate springs and principles of the science of the mind are inexplicable, but that is so because these springs and principles are ultimately reduced into certain brute mental facts, or automatic and mechanical powers in the mind that cannot be explained otherwise. Therefore, in this sense they are inexplicable, but the causal/functional account of mental phenomena is in itself a kind of philosophical explanation. With this background in place, I turn now to the analysis of Hume's mechanics of the mind.

All perceptions of the human mind resolve themselves into two distinct kinds: *impressions* and *ideas*. Ideas are the exact, but faint images, or copies, of impressions. Hume's "first principle," called the Copy Principle, clearly expresses this critical point. The copy principle claims that "*all our simple ideas in their first appearance are deriv'd from simple impressions, which are correspondent to them, and which they exactly represent*" (THN, 4). The difference between impressions and ideas is that impressions have a greater degree of force, liveliness, and vivacity than the latter groups of perceptions. Impressions themselves could be of two distinct kinds: those of *sensation* and those of *reflection*. Impressions of sensation arise "in the soul originally, from unknown causes" (THN, 7). Hume does not have a theory about the causes of sense impressions, which explains why he claims that "[t]he examination of our sensations belongs more to anatomists and natural philosophers" (THN, 8). Hume's main preoccupation in the *Treatise* and in the *First Enquiry* is to examine impressions of reflection, viz. passions, desires, and emotions, which are derived in a great measure from our ideas consisting of the raw materials or contents upon which many of our faculties of the mind operate.

There are two key principles which play significant roles in Hume's mechanics of the mind which need mention here: the principle of association (relations) and the principle of imagination. These two principles are of much more importance than impressions and ideas, as the raw materials or the contents upon which these two principles operate. Particularly, the key roles played by the association of ideas are the most distinguishing aspects of Hume's science of the mind. In the *Abstract* he proudly proclaims that "if anything can entitle the author to so glorious a name as that of an *inventor*, 'tis the use he makes of the principle of the association of ideas, which enters

into most of his philosophy” (Abs, 35). There are three principles of association: resemblance, contiguity, and cause and effect. Hume calls these associating principles *natural relations* because in each such relation one idea naturally introduces another. “[T]here is a secret tie or union among particular ideas,” he says, “which causes the mind to conjoin them more frequently together, and makes the one, upon its appearance, introduce the other” (Abs, 35). Hume concludes that “’twill be easy to conceive of what vast consequence these principles must be in the science of human nature, if we consider, that so far as regards the mind, these are the only link that bind the parts of the universe together...they are really *to us* the cement of the universe, and all the operations of the mind must, in a great measure, depend on them” (Abs, 35). The key function of these associating principles is to explain uniformities among mental phenomena. Most important of all, they are supposed to render the imagination as uniform as possible. The imagination has great authority over ideas; it is at great liberty to transpose and change its ideas (THN, 10). “As all simple ideas may be separated by the imagination, and may be united again in what form it pleases,” Hume asserts, “nothing wou’d be more unaccountable than the operations of that faculty, were it not guided by some universal principles, which render it, in some measure, uniform with itself in all times and places” (THN, 10). And if there were no principles regulating the empire of the imagination, it would have freely separated, and joined, and composed our ideas into all varieties of fiction (Abs, 35). There is a certain kind of uniformity or, I would prefer to say, a kind of automatic and mechanical feature in the way that these associating principles operate.

The second key principle that plays a vital role in Hume’s mechanics of the mind is the imagination. Hume’s notion of the imagination is a complex and multifaceted one. He states in a footnote to the *Treatise* that he uses the term ‘imagination’ in two different senses:

When I oppose the imagination to the memory, I mean the faculty, by which we form our fainter ideas. When I oppose it to reason, I mean the same faculty, excluding only our demonstrative and probable reasoning. When I oppose it to neither, ’tis indifferent whether it be taken in the larger or more limited sense, or at least the context will sufficiently explain the meaning. (THN, 117n22).

The first sense of ‘imagination’ is concerned with the *feigning* of ideas, whereas the second one is concerned with a belief-forming faculty of the mind. The imagination as opposing memory is the faculty of our fainter ideas, and in the latter sense, i.e., the anti-

reason sense, it is more directly concerned with signifying the vivacity, force, and liveliness of ideas (5). Some philosophers argue that Hume uses the notion of the imagination in three distinct senses: (a) as opposed to memory, (b) as opposed to reason, and (c) as opposed to the understanding (6).

We may distinguish between three different faculties or functions of the imagination in Hume's cognitive psychology. The first kind of function is the *metaphysical* faculty of imagination. Accordingly, Hume regards the metaphysical systems which he wants to demarcate from science as mere fiction of the imagination (THN, 108), or as merely the offspring of the imagination (THN, 108), or as trivial suggestions of the fancy (THN, 267) (7). A kind of blind imagination has led the metaphysicians to hypothesize about other reality over and above our experience on a purely conceptual basis. The second function that he enumerates is the *artistic* function of the imagination. Here, the faculty of imagination, as Hume says, is at liberty to transpose and change its ideas (THN, 10) the way it wishes; it imposes new order on sense impressions. The faculty of imagination forms new complex ideas by uniting and joining our simple impressions and ideas. The imagination in this sense is that of our very creative power to invent fairy tales (8). Finally, we may distinguish the *scientific* function or faculty of imagination. Here, the function of the imagination is to bridge the gap between the observed and the unobserved. It presents the idea of a continuously existing object by producing resembling perceptions despite the fragmentary and interrupted nature of the series of impressions. In particular, the faculty of imagination in this sense is supposed to generate the ideas of necessary connection, the existence of external objects and personal identity (9). In a sense, Hume's major philosophical contribution takes place when he explains how the faculty of imagination naturally generates these fundamental philosophical ideas, and that how the metaphysicians have speciously attributed the rise of these ideas to other unjustified sources.

The artistic function of imagination is central to Hume's cognitive psychology. Its principal function is to generate new complex ideas, which have no corresponding complex impressions. Hume holds that impressions and ideas can be either *simple* or *complex*. While the simple ideas have their exact corresponding simple impressions, many of our complex ideas do not have such corresponding impressions. So, where do

the complex ideas come from? In reply, Hume says that the mind can combine simple impressions in such novel ways into newborn complex ideas. The imagination, as a non-reflective faculty, does this fundamental job. Stated otherwise, the imagination reorganizes our simple impressions and simple ideas by placing them into new relationships, and thus gives rise to an entirely new cognitive order. To compare the imagination with memory, memory is supposed to enable us to recall our past experiences, but the imagination enables us to impose new order on our past experiences, which results in producing new ideas. In this way, the imagination generates numerous kinds of complex ideas that have no corresponding complex impressions. For example, one can imagine “such a city as the *New Jerusalem*, whose pavement is gold and walls are rubies, tho’ [she] never saw any such” (THN, 3). Presumably, the most important kinds of complex ideas that arise in this manner are the ideas of space and time. Hume argues that we have no simple impressions of space and time, and yet we have the ideas of spatial extension and temporal duration. He says that, however, we do have compound impressions of an array of coloured or tangible mathematical points (THN, 32), and this gives us the idea of space. This does not cause any problem for him because he had concluded earlier that “all our simple ideas proceed, either mediately or immediately, from their correspondent impressions” (THN, 7). All he needs to do is to show how the ideas of space and time are derived from sense impressions since he is only concerned about the origins of our ideas of space and time.

For Hume, the ideas of space and time cannot be derived from simple impressions. Hume repudiates the Lockean *additive* theory according to which we combine, or add, simple impression to simple impression and then, produce the idea of length, and eventually, of extension in general ⁽¹⁰⁾. Instead, Hume thinks that we have experimental evidence indicating that our simple impressions of physical objects are unextended and indivisible. Historically speaking, Hume’s treatment of the problem of space and time is interpreted as an alternative to the established *infinite divisibility* theory of space and time according to which the ideas of space and time consist of “an infinite number of parts or inferior ideas” (THN, 39). In contrast, Hume argues that these ideas must consist of finite numbers that are “simple and indivisible” (THN, 39). These minimal ideas cannot be divided. He concludes, “’tis certain, that the imagination reaches

a *minimum*, and may raise up to itself an idea, of which it cannot conceive any sub-division, and which cannot be diminish'd without a total annihilation" (THN, 27). Similarly, we have minimal impressions that we receive through sensory or introspective experiences (11). The idea of space and time is derived from the impressions of sets of these minimal indivisible and unextended but visible mathematical points (THN, 34). Hume argues that the simple, indivisible, and unextended parts (the visible mathematical points), which make up the ideas of space and time, would be inconceivable unless they turn out to be coloured and tangible. He claims that "*that the idea of space or extension is nothing but the idea of visible or tangible points distributed in a certain order*" (THN, 53). Since these ideas are tangible and coloured we can perceive an array of them in such a way as to be given the idea of space or of the "manner or order, in which objects exist" (THN, 40). The "idea of extension," Hume says, "is nothing but a copy of these colour'd points, and of the manner of their appearance" (THN, 34). The arrangement of these coloured points is a compound impression. The psychological mechanism that generates such ideas is deep-seated within the faculty of imagination. It is said that the imagination can transform the simple impressions into a "compound impression, which represents extension" or the abstract idea of space itself. This is the most fundamental function of the faculty of imagination in its artistic sense. Hume accounts for the idea of time on similar grounds.

There is another situation in which the artistic function of imagination is well observed. It is argued that, on Hume's account, the belief in the uniformity of events is instinctive and not subject to our will (12). In the *First Enquiry*, Hume explicitly claims that our belief in the uniformity of nature cannot be explained by the principles of association (EHU, 55). This belief in the uniformity of nature that engages all animals is not founded on "any process of argument or reasoning, by which he concludes, that like events must follow like objects, and that the course of nature will always be regular in its operations. For if there be in reality any arguments of this nature, they surely lie too abstruse for the observation of such imperfect understandings" (EHU, 106). Thus, reason does not give us the belief in the uniformity of nature; it is all given by an instinctive habit that we call imagination. Hume writes:

Animals, therefore are not guided in these inferences by reasoning; neither are children; neither are

the generality of mankind, in their ordinary actions and conclusions: neither are philosophers themselves, who, in all the active parts of life, are, in the main the same with the vulgar, and are governed by of the same maxims. Nature must have provided some other principle, more ready, and more general use and application; nor can an operation of such immense consequence in life, as that of inferring effects from causes, be trusted to the uncertain process of reasoning and argumentation. (EHU, 106)

He concludes:

It is custom alone, which engages animals, from every object, that strikes their senses, to infer its usual attendant, and carries their imagination, from the appearance of the one, to conceive the other, in that particular manner, which we denominate *belief*. No other explication can be given of this operation, in all the higher, as well as lower classes of sensitive beings, which fall under our notice and observation. (EHU, 106-107)

Accordingly, nature does not give us the liberty to decide whether or not to believe in its regularity. Instead, it is the nature that has entrusted upon us “an operation of such immense consequence in life” (EHU, 106). The key point is that the belief in the constancy of nature is not a rational one; rather, it is an unjustified and unreasonable belief. The belief in the uniformity of nature is also an irrational one. In some places, Hume seems to distinguish the irrational belief from the natural or instinctive belief, but here he seems to suggest that both kinds of unjustified beliefs, whether natural (instinctive) or irrational beliefs, are equally natural and instinctive. This distinction is most clearly expressed in the *First Enquiry*.

Hume explicitly reduces both kinds of natural and irrational ⁽¹³⁾ beliefs into a single category of belief in the *First Enquiry*. There, he argues that all kinds of beliefs arise from the experimental reasoning or from our natural attitudes that are merely instinctive and mechanical. He claims:

But our wonder will, perhaps, cease or diminish, when we consider, that the experimental reasoning itself, which we possess in common with beasts, and on which the whole conduct of life depends, is nothing but a species of instinct or mechanical power, that acts in us unknown to ourselves; and in its chief operations, is not directed by any such relations or comparisons of ideas, as are the proper objects of our intellectual faculties. Though the instinct be different, yet still it is an instinct, which teaches a man to avoid the fire; as much as that, which teaches a bird, with such exactness, the art of incubation, and the whole economy and order of its nursery. (EHU, 108).

Hume not only identifies those natural beliefs arising from the artistic function of imagination as instinctive and mechanical, but also argues that even the experimental

reasoning itself is an instinctive and mechanical power acting on us unknown to ourselves. Perhaps the idea that even experimental or causal reasoning is a species of natural or instinctive operation of the mind is the most crucial move that Hume makes in his cognitive psychology. Gerhard Streminger, however, has wrongly argued that this development occurred in Hume's later philosophy, i.e., mainly in the *First Enquiry*. On the contrary, Hume unambiguously expresses the same idea in the *Treatise*. In the concluding paragraph of the section of "Of the Reason of Animals," he writes:

Reason is nothing but a wonderful and unintelligible instinct in our souls, which carries us along a certain train of ideas, and endows them with particular qualities, according to their particular situations and relations. This instinct, 'tis true, arises from past observation and experience; but can any one give the ultimate reason, why past experience and observation produces such an effect, any more than why nature alone shou'd produce it? Nature may certainly produce whatever can arise from habit: Nay, habit is nothing but one of the principles of nature, and derives all its force from that origin. (THN, 179).

Hume definitely has a consistent view of causal reasoning as a species of the non-reflective or natural faculty of the mind. Hume firmly believes that 'causal inference,' which is "a true species of reasoning, [and] the strongest of all others" (THN, 97n20), is essentially automatic and non-reflective. In Hume's view, 'causal inference' or 'experimental reasoning,' which we share with infants and brute beast, is a non-theoretical process that is organized and run by the ever-active and non-rational faculty of the imagination. Hume argues that the inference from the observed to the unobserved is not determined by reason, but by the imagination. He furthermore tries to demonstrate that certain natural and primitive dispositions of the mind, namely, the principles of associations and relations of ideas, are responsible for the making of that inference. The observation of the constant conjunction between objects will have the inevitable effect of creating a 'union in the imagination' among their ideas. Whenever the idea of one appears in the mind, its 'usual attendant' immediately follows, without any intervening reflection or reasoning. Such a transition is so automatic and non-reflective that it cannot be prevented from occurring. As mentioned briefly earlier, in Hume's view, natural belief, which arises by 'custom,' as a result of repetitions in our experience, is an inevitable process that stands free and independent from our conscious decision-making faculty. We cannot decide what to believe or what not to believe. "Nature, by an absolute

and uncontrollable necessity has determin'd us to judge as well as to breathe and feel," Hume asserts, "nor can we any more forbear viewing certain objects in a stronger and fuller light, upon account of their customary connection with a present impression, than we can hinder ourselves from thinking as long as we are awake, or seeing the surrounding bodies, when we turn our eyes towards them in broad sunshine" (THN, 183). Roughly, the artistic function of imagination consists in this sort of explanation.

The automatic and unreflective character of the imagination is best observed in its so-called scientific function ⁽¹⁴⁾. Although the distinction between the artistic and the scientific functions sometimes overlaps, it is still an effective and useful analytic tool to distinguish different functions of the imagination. The scientific function of imagination is understood in contrast to reason. Indeed, a fundamental aspect of Hume's new science of the mind is the distinction that he makes between the imagination and reason. Reason, or the understanding, is the reflective faculty of the mind that compares two different ideas concerning their similarities and differences. The relations that are involved here are merely "philosophical relations" ⁽¹⁵⁾. The imagination, on the other hand, is the kind of non-reflective faculty that operates on the basis of three "natural relations," i.e., contiguity, resemblance, and causation. Here the imagination takes us from one idea to another idea associated to it by means of the three natural relations. As a result, the imagination produces a pattern of behaviour called "habit" or "custom," which is responsible for the mind's propensity to continue any process or motion once it has begun it. Habit is the most fundamental feature of the imagination. In particular, habit explains how the imagination fills the gappy and interrupted series of perceptions. And by doing so it generates the feeling of continuity and uninterruptedness in our perceptions, which this in turn gives rise to the fictional idea of the self as well as to the idea of the continued and distinct existence of external objects. In other words, the imagination is responsible for producing the qualities or characteristics of *constancy* and *coherence*, which these in turn generate the idea of continuity of objects over time. By and large, this function of the imagination deals with the psychological features of our perceptual system in producing the ideas of object identity and personal identity. This is an essential part of the current project, and to be discussed exclusively in the next chapter. But the most important occasion during which the scientific function of imagination is observed is the role of the

imagination in producing the notion of causal inference. I discuss this in more detail in the following section.

II. 2: The Imagination and Causal Inference

Causal inference is the most important relation for ordering our perceptual experiences, and it is by means of such ordering that we are able to establish a science of human nature. Hume argues that causal relation is not determined by reason; rather, the faculty of imagination is responsible for making the transition from cause to effect. The distinction between reason and the imagination is fundamental to Hume's new science of the mind. He distinguishes reason from the imagination by making a distinction between seven philosophical relations. The first four of these relations, "*resemblance, proportion in quantity or number, degrees in any quality, and contrariety*," Hume asserts, "depend entirely on the ideas, which we compare together" (THN, 69), whereas the other three relations, "*identity, relations of time and place, and causation...* may be chang'd without any change in the ideas" (THN, 69). For him, the first category of philosophical relations is the source of demonstrative reasoning, and the latter category is supposed to produce probable reasoning. This distinction says that "the difference between the two kinds of relation is the difference between relations that are *independent* of the manner or order in which perceptions are presented or considered, and relations that *depend* on this manner" (16). For instance, if two unchanging ideas, say, *X* and *Y*, are related, and they come into mind whether simultaneously, or *X* precedes *Y*, or vice versa, it does not have any effect on the relations of resemblance, quantity, quality, and contrariety. Therefore, these relations are not affected by any change in the manner or order in which perceptions are presented to the mind.

On the other hand, the relations of identity, causation, and contiguity in time and place, are affected and determined by the manner and order in which perceptions are presented to the mind. It makes a fundamental difference, if the idea of *X* as the cause of the idea of *Y* precedes *Y* at all times. If somehow the order is not maintained between these ideas as cause and effect, we will hardly have any idea of causation. The same is true of the ideas of identity and contiguity in time and place. Let's say a perceptual subject experiences the idea of *T* in an uninterrupted manner. In this case, we are inclined to ascribe a notion of identity to the idea of *T*. If, however, we have an interrupted series

of resembling ideas, say, T , T' , T'' , and so on, we are inclined to ascribe diversity to these ideas. Consequently, the manner or order in which perceptions are presented to the mind makes a fundamental difference with respect to the relations of identity, causation, and relations of time and place. On Hume's account, the first four relations provide the ground for 'demonstrative reasoning' and thus for certainty and knowledge, while the remaining three gives us nothing more than 'probable reasoning.'

Hume's objective for differentiating between these two categories of relations is to establish that the inference from the observed (cause) to the unobserved (effect) is not a move that reason determines us to make. "If reason determin'd us, it wou'd proceed upon that principle," Hume argues, "*that instances, of which we have had no experience, must resemble those, of which we have had experience, and that the course of nature continues always uniformly the same*" (THN, 89). This is called the 'uniformity principle' (17). According to this principle, inferences from the observed to the unobserved proceed upon the supposition that it is universally true. Hume, however, argues that demonstrative reasoning never establishes the truth of the uniformity principle because:

We can at least conceive a change in the course of nature; which sufficiently proves, that such a change is not absolutely impossible. To form a clear idea of any thing, is an undeniable argument for its possibility, and is alone a refutation of any pretended demonstration against it. (THN, 89)

In a nutshell, Hume's conception of reason, which assumes that all reasoning is deductive, cannot determine us to make the transition from the observed to the unobserved. All we have is the constant conjunction between objects, and "*that even after the observation of the frequent or constant conjunction of objects, we have no reason to draw any inference concerning any object beyond those of which we have had experience*" (THN, 139). As a consequence, Hume suggests that, if reason does not make the transition, we must look for the source of it elsewhere because there is no doubt that we make such a transition, or inference, from the observed to the unobserved. He then argues that the 'imagination' as a natural propensity or disposition of the mind is responsible for the transition that we make from the observed to the unobserved.

Hume's core idea is that the experience of constant conjunction between the observed and the unobserved creates a "union in the imagination." On this account, "we always find that from the constant conjunction the objects acquire an union in the

imagination” (THN, 93). When the idea of the observed becomes present in the mind, we immediately form the idea of its “usual attendant” without having any further reflection or reasoning (THN, 93). Whenever we have the idea of an A in the mind, the idea of a B immediately occurs. This is an automatic process, and cannot be prevented. Of course, the role of past experiences is quite vital to making this transition. The mind’s tendency operates on the basis of our past experiences “in such an insensible manner as never to be taken notice of, and may even in some measure be unknown to us” (THN, 103). And “the imagination can draw inferences from past experiences, without reflecting on it; much more without forming any principle concerning it, or reasoning upon that principle” (THN, 104). Hume offers the example of a person who stops at a river before him. He foresees the consequences of his action, which is conveyed to him by his knowledge of past experiences, but he cannot think that “on this occasion he reflects on any past experience, and calls to remembrance instances, that he has seen or heard of, in order to discover the effects of water on animal bodies” (THN, 104). Instead, this tendency of the mind, viz., custom, “operates before we have time for reflection” (THN, 104). In this situation, “the idea of sinking is so closely connected with that of water, and the idea of suffocating with that of sinking, that the mind makes the transition without the assistance of memory” (THN, 104). Here, “[t]he objects seem so inseparable, that we interpose not a moment’s delay in passing from the one to the other” (THN, 104). Hume concludes that “as this transition proceeds from experience, and not from any primary connexion betwixt the ideas, we must necessarily acknowledge, that experience may produce a belief and a judgment of causes and effects by a *secret operation*, and without being once thought of” (THN, 104: emphasis added). It is this secret operation, or functioning of the imagination, that convinces the mind that “*that instances of which we have no experience, must necessarily resemble those, of which we have*” (THN, 104).

As we have seen, the imagination can draw inferences from past experience, but without reflecting on it, or without forming any principle concerning it. We are never conscious of any such operation in the mind (THN, 102). It is important to note that what Hume means by not being conscious in this context is “to see that ... we have no direct, introspective, access to the processes in question. In making causal inferences, for example, we obviously do not consciously recall the previous instances of constant

conjunctions on which the inference is based” (18). In fact, “CUSTOM, which proceeds from a past repetition, without any new reasoning or conclusion” (THN, 102), guides the mind to make the customary transition from the observed to the unobserved. This secret operation that guides the mind to make that transition is an automatic, mechanical, and non-reflective tendency that the imagination brings about. The imagination when “set into any train of thinking, is apt to continue, even when its object fails it, and like a galley put in motion by the oars, carries on its course without any new impulse” (THN, 198). Nothing is “more usual, than for the mind to proceed after this manner with any action, even after the reason has ceas’d, which first determin’d it to begin” (THN, 48).

Let me use some formality to explain this. Adam, for instance, observes a moving ball striking a resting ball, and then sees that the resting ball begins to move. He then observes this experience repeatedly. Indeed, he observes that:

- (1) A_1 strikes B_1 , then B_1 moves;
- (2) A_2 strikes B_2 , then B_2 moves;
- (3) A_3 strikes B_3 , then B_3 moves;
- (4) A_4 strikes B_4 , then B_4 moves;

Now, when:

- (5) A_n strikes B_n , Adam concludes that, B_n is going to move.

The observation of the constant conjunction of As and Bs creates a mental habit in Adam’s mind, forcing him to *infer* that something unobserved will follow something observed. Eventually, he comes up with the rule (law) that:

- (6) Whenever As happens, then, Bs occurs.

The move that Adam’s mind makes gradually from episode (1) to episode (6) is due to the smooth passage of the vivacity and liveliness of the impressions resulting from the observation of the constant conjunction of the previous episodes, which are made available by memory. When Adam’s mind moves from episode (1) to episode (2) and then all the way to episode (5), it collects memories of different but resembling episodes. These collected memories are essentially the ideas of different impressions stored in the mind. These ideas generate the internal impression as the origin of the idea of necessary connection. In fact, in episode (5), when Adam’s mind observes the constant conjunction between the observed and the unobserved, it only experiences another episode of such

constant conjunction without any idea of necessary connection. But, as I said, Adam's mind has stored a collection of resembling ideas from previous episodes. When his mind moves from (5) to (6), and thus gives rise to the idea of necessary connection, the mind bridges the gap or interruption between these episodes through generating a resembling internal impression that has originally got into Adam's mind through his first and subsequent observations of constant conjunction between cause and effect. Therefore, Adam's mind slides smoothly across the gap and creates the feeling of necessity between the observed and the unobserved.

In making the transition from the observed to the unobserved, we are fundamentally involved in coming to *believe* that when A happens, then, *necessarily* B occurs. Here we are concerned about Hume's theory of belief rather than his theory of ideas. There is a strong sense of necessity involved in making the transition, and Hume is concerned to explain the psychological mechanism that generates the kind of natural belief that we come to have. On his account, what distinguishes the simple idea, A, from a belief about it, is related to what distinguishes an impression from an idea. Hume holds that impression differs from idea in respect to its degree of force, vivacity, and liveliness. He claims that a belief is "a lively idea related to or associated with a present impression" (THN, 96). Or a belief is "a more vivid and intense conception of an idea, proceeding from its relation to a present impression" (THN, 103). After establishing that belief is a lively idea related to some present impression Hume explains why belief arises in the mind. He writes that "*when any impression becomes present to us, it not only transports the mind to such ideas as are related to it, but likewise communicates to them a share of its force and vivacity*" (THN, 98). If impression did not have this feature of communicating its force and vivacity to ideas, we would never be able to produce any beliefs. The source or origin of belief, therefore, is traced back to the manner or order in which impressions first appear in the mind.

What really explains, for Hume, the real difference between a belief and a simple idea is how we *feel* them. Accordingly, there is no power, necessity, and connexion between an idea of A and an idea of B, but only when there is an experience of A, then there a *felt* expectation of B arises. Feeling or sentiment, which is "a superior *force*, or *vivacity*, or *solidity*, or *firmness*, or *steadiness*, ...[is] that act of the mind, which renders

realities more present to us than fictions, causes them to weigh more in the thought, and gives them a superior influence on the passions and imaginations” (THN, 629). Belief is this feeling or sentiment that is determined by the mind. And it is in this sense that Hume asserts that “all probable reasoning is nothing but a species of sensation” (THN, 103). In fact, he elsewhere expresses this idea more forcefully: “*that all our reasonings concerning cause and effects are deriv’d from nothing but custom; and that belief is more properly an act of the sensitive, than of the cogitative part of our natures*” (THN, 183). We can argue that the reason that Hume searches for the origin of belief in our feeling and sentiment is that he wants to emphasize that our belief in the unobserved arises completely naturally, like any other phenomenon in nature (19). He thinks that it arises by ‘custom’ or ‘habit,’ as a “result of repetitions in our experience. We do not decide to believe what we do; we are not free not to believe those things that are most fundamental for us” (20). “Nature,” Hume argues, “by an absolute and uncontrollable necessity has determin’d us to judge as well as to breath and feel” (THN, 183). Thus, the belief that arises in this constant conjunction of two different ideas is accompanied by a feeling of determination, and this process in turn is produced by a primitive and natural disposition of the mind, namely, the imagination.

Hume argues that the mere repetition of ideas in the mind produces something new in those instances that are not observed. He claims this while he is committed to the idea that each instance or occurrence is independent of all the other instances and unable to produce anything new in the relation between objects. In each instance of the relation of a cause and effect we only observe one thing following the other; we never get the impression of any necessary connection. It is only after constant conjunction, or repeated observation of the observed and the unobserved that we come to have the idea of necessary connection. So if we have no impression that generates the idea of necessary connection, the question that Hume asks is this: where does the idea of necessary connection come from? Hume holds that the imagination produces an impression from which the idea of necessary connection is derived. That impression is “an internal impression of the mind,” or a reflective impression, which is produced by the immediate feeling determined by the mind (THN, 165, 156). “The idea of necessity arises from some impression,” Hume asserts, “there is no impression convey’d by our senses, which

can give rise to that idea. It must, therefore, be deriv'd from some internal impression, or impression of reflection" (THN, 165). He goes on to say that "there is no internal impression, which has any relation to the present business, but that propensity, which custom produces, to pass from an object to the idea of its usual attendant" (THN, 165). Accordingly, Hume identifies the internal impression as the origin of the idea of necessity, which is produced by the repetition of constant conjunction in the imagination.

My main objective in this section was to show that causal inference is produced by the imagination. The imagination, as explicated in the previous section, comprises the basic structure of Hume's conception of the mind as an automatic, non-reflective, and mechanical cognitive system. We have seen that the mechanism of imagination generates an internal impression in the mind, and this in turn gives rise to the idea of necessary connection. In this respect, the imagination generates an internal impression from which the idea of necessary connection is derived. Such an internal impression has no original or direct origins in our sense-impressions. It is an invention of the imagination. But more importantly, I wanted to show that the idea of causal inference, which is so fundamental to Hume's philosophy, is produced by the psychological mechanism of imagination. For Hume, causal inference as such is essentially a mechanical and non-reflective process.

The claim that Hume's theory of the mind is computational calls for further investigations into the nature his views on the architecture of the mind. In fact, in order to explain how the mental states and processes are natural phenomena, or to explain them in causal or functional terms, requires an account that explains the connection between functional architecture of the mind and its mental contents. Contemporary philosophers of mind think that a modular view of the structure of the mind is essential to explain this connection. In the following section, I want to discuss Hume's view on the architecture of the mind.

II. 3: Hume on the Architecture of the Mind

Fodor in *The Modularity of Mind* has ascribed an *associationist* mental structure to Hume's theory of mind. He argues that the Associationists, including Hume, view the association of ideas "as a mechanical relation *among* mental contents, not as a computational relation defined over them" (21). According to Fodor, "Hume speaks of associations between Ideas on the model of gravitational attraction between physical

objects” (22). However, he changed his view on Hume’s philosophy of mind in his later work (23). Recently, Fodor has forcefully argued for a functionalistic reading of Hume’s theory of mind in his *Hume Variations*. There, he argues that “Hume’s *Treatise* is the foundational document of cognitive science: it made explicit, for the first time, the project of constructing an empirical psychology on the basis of a representational theory of mind” (24). I will come back to this in a moment. But, first, I must discuss Hume’s views on the architecture of the mind and why I think his theory of mind includes the kind of psychological architecture that functionalism requires.

I claim that there is a common recognition by both Hume and modern cognitive science that our cognitive capacities and performances are, in principle, *automatic*, *modular*, and *sub-doxastic* (25). It is central to the Modularity Thesis that the input systems do not get help from the processes of belief fixation (central system) because the modules are domain specific. Also, as Fodor has said, the input systems are ‘informationally encapsulated’ and have ‘limited central accessibility.’ Informational encapsulation refers to the idea that information available to a given module for processing is limited to that which is contained in its proprietary database; it cannot have access to information stored in other systems (modules). In other words, the flow of information *into* the input system is constrained, whereas inaccessibility refers to the flow of information out of the input system. Therefore, the input systems have neither access to external information and resources, for example, belief fixation (central monitoring), nor receive information from other modules. The main idea here is that a system is inaccessible in the sense that its internal processing is unavailable and inaccessible to consciousness and, thus, to a certain extent, opaque to introspection. Overall, the idea that the internal processing is inaccessible to consciousness amounts to the idea that operations and processes involved are found to be task-specific and function in isolation from each other and the higher-level cognitive processes and functions. This refers to the sub-doxastic character of our mental phenomena as taking place below the threshold of our consciousness, and thus, below the level of belief (26). They are automatic simply because the perceiver has no control over these operations and processes in the mind. This is, in principle, analogous to Hume’s belief that what we perceive is largely independent of volition as well as belief fixation. Hume has famously said that we cannot

decide what to believe or what not to believe. Hence, Hume's conception of cognitive processes is essentially computational because he thinks of the architecture of the mind as automatic, modular, and sub-doxastic. In a broader sense, Hume's new science of the mind is an anticipation of contemporary 'cognitive science.'

John Biro, for instance, has argued that Hume's conception of cognitive processes (or states) is well observed in the following specific matters: The Problem of Qualia and the case of the Missing Shade of Blue, the problem of Exempt Agent or Undischarged Homunculi. I briefly explain these cases now. To begin with the latter, Daniel Dennett has levelled an objection against Hume, known as "Hume's Problem." Dennett claims that "Hume wisely shunned the notion of an inner self that would intelligently manipulate the ideas and impressions." He then states that "this left [Hume] with the necessity of getting the ideas to "think for themselves. His associationistic couplings of ideas and impressions, his pseudo-chemical bonding of each idea to its predecessor and successor, is a notorious non-solution to the problem" (27). In fact, Dennett claims that Hume's problem of the 'exempt agent' is the problem of all kinds of representational theory of mind. He argues that representationalism, including Hume's, presupposes an undischarged homunculus, whose intelligence is not explained. In other words, any system of mental representation supposes, or requires, a user or interpreter "external" to that system. The reason is that a representation is only *for* or *to* someone and since functionalism and Hume's bundle theory suppose representationalism—that is, it represents something to or for someone—it follows that they somehow must require an exempt agent, a homunculus (28). He concludes that "on the one hand, how could *any* theory of psychology make sense of representations that *understand themselves*, and on the other, how could *any* theory of psychology avoid regress or circularity if it posits at least one representation-understander in addition to the representation?" (29). Dennett thinks that both Hume and representationalism fail because they cannot avoid positing the idea of ideas understanding themselves.

Obviously, this is not an easy objection; it seriously challenges some underlying assumptions of RTM, including Hume's. I think that Dennett's objection is a subversive one. It states that representation is *for* or *to* someone (or a self or mind) and that, seemingly, RTM presupposes the existence of a mind that representation is represented *to*

or *for*. It is widely acknowledged that intentionality is something relational, as in between the object represented and someone that the object is represented to or for. The existence of the subject of intentionality is essential to the relation as a whole, but the object itself does not need to exist in reality at all because it can involve e.g., fictitious entities (30). Although this is one serious objection, it can be avoided. Biro has pointed out, for example, that “part of the difficulty is that Hume does make *perceptions*—not bundle of them—the subjects of certain quasi-intentional verbs” (31). Perceptions are said to do things such as ‘producing,’ ‘attracting,’ ‘destroying,’ and ‘influencing’ each other. Since Hume uses these verbs as if they are the subject of our intentional activities, some commentators assume that there is an “active self,” as explained in Chapter One. Nevertheless, according to Biro, Dennett is mistaken because he fails to distinguish between these quasi-intentional verbs, on the one hand, and the mind, or self, as the genuine subject of our intentional predications, on the other. It seems that, for Hume, person or mind, as opposed to perceptions, is the real subject of intentionality. No doubt, Hume frequently writes in a way that makes perceptions the subject of intentionality; that would be misleading, however, especially if we note that perceptions may, for example, ‘produce,’ or ‘attract,’ each other, but Hume never suggests that perceptions themselves can ‘think,’ ‘understand,’ ‘desire,’ and ‘will’ (32). Therefore, in Hume’s view, we must differentiate between personal and sub-personal levels of cognition; he wants us to recognise that to be able to explicate intentional behaviour at personal level requires adverting to processes at sub-personal levels, and that the sub-personal processes are always describable in non-intentional terms (33).

One may argue that Dennett raises a more fundamental problem here. It can be argued that his chief claim is that RTM presupposes a concept of the self, or of the mind, that it is supposed to explain in the first place. Dennett’s idea is based on an intuition telling us that in order to explain mental representations as intelligent phenomena playing certain roles, or functions, requires another intelligent being to do the supervising for those representations. In other words, the intuition says how could an occurrence of mental representation be both a representation of an object in the mind as well as an agent supervising its own activities? This is the kind of intuition that has made many of Hume’s scholars to draw the same conclusion as Dennett did. Dennett’s criticism poses a

historical problem for the functionalists as well. From a historical viewpoint, the extensive philosophical works done within the tradition of functionalism, or RTM, have paid almost no attention to the question of the self, and its relation to functionalism, or to RTM for that matter. As a matter of fact, it looks though the question of personal identity does not arise for them at all. And that's why Dennett's criticism of RTM should not come to functionalists as a surprise insofar as they have not offered an adequate treatment of the question of personal identity. Sydney Shoemaker, however, defends a psychological-functional continuity account of personal identity over time (Chapter IV). It is a main thesis of this project that Hume likewise offers a functional continuity account of personal identity.

Dennett's strategy to avoid Hume's problem—the problem of homunculus—is to adopt an *intentional stance*, a theory in which all homunculi have been discharged. He calls those objects that behave rationally and exhibit elements of beliefs, desires, and other mental phenomena described as intentional, *intentional systems*—systems whose behaviours we can predict. We can adopt three distinct stances towards an intentional system: the physical stance, the design stance, and the intentional stance. If predicting the behaviour of a system is determined by its physical constitution (that is, its microphysical constituents) and our knowledge of the laws of physics or chemistry, there involves a physical stance. If we predict the behaviours of a system on the basis of the assumption that it has certain designs or functions, then we have adopted a design stance. Finally, we take an intentional stance when the other two stances fail to predict the behaviours of an intentional system. It consists of a few elements: first, we treat the object as a rational agent; then, we figure out what beliefs that agent ought to have, given its purpose and place in the world. Similarly, we figure out what desires it ought to have, and that we assume, or predict, that a rational agent acts in a way to *optimize* its goals⁽³⁴⁾. However, the belief and desire of an intentional system have no *real* content; rather, these putative states can be relegated “to the role of idealized fictions in an action-predicting, action-explaining calculus”⁽³⁵⁾. In Dennett's view, these intentional states are not what Reichenbach calls “illata—posited theoretical entities.” Instead, they are “abstracta—calculation bound entities or logical constructs”⁽³⁶⁾. Accordingly, when we consider an object—an organism or an artifice—as an intentional system neither do we

make any commitments about its internal states as physically characterized, nor do we say anything about its design or function of the system. Rather, we ascribe belief, desires, and other intentions to an intentional system as conceptual constructs and predict and explain its behaviours on the basis of ascribing belief and desire to it (37).

Some people have found this unsatisfactory. Fodor and Ernest Lepore have pointed out, for example, that Dennett's so-called intentional stance, is unintelligible or question-begging because it depends on either evolutionary or a transcendental argument (38). I will put aside, for the moment, the more fundamental objection that Fodor and others have raised against Dennett's so-called instrumentalism: the claim that mental events have subjective features, namely, qualia, and that mental contents and representations are ontological existences in the mind (39). But I shall say that the way that Dennett approaches the problem of homunculus leaves almost no hope for resolving that problem at all. To put it differently, the same intuition that he uses to refute RTM could work against his intentional stance approach as well. In fact, the problem of homunculus is more vivid in the intentional stance approach rather than in RTM or Hume's bundle theory of the self. According to the intentional stance approach, one has to presuppose a highly abstracted notion of the self cognizing the behaviours of the other self; there has to be *someone* to take a stance on the behaviours of the other self. In fact, one can ask: on what does Dennett's intentional stance stand? I think that this is a challenging question for Dennett, and some people have raised it as well (40). This is certainly an ontological question concerning the status of the intentional stance. Crudely stated, it asks how the intentional stance and those entities posited by it, e.g., beliefs, desires, etc., stand in relation to the physical stance and the design stance. In other words, is the intentional stance embedded in the "furniture of the physical world" or is it something over and beyond it? Dennett has defended a position—he calls it a small "r" realist position—in which he reduces Hume's homunculus into what has come to be known as "homuncular functionalism." According to homuncular functionalism,

The AI programmer ...breaks the computer down into subsystems, each of which is given intentionally characterized tasks; he composes a flow chart of evaluators, rememberers, discriminators, overseers and the like. These are homunculi with a vengeance. . . . Each homunculus in turn is analyzed into smaller homunculi, but, more important, into less clever homunculi. When the level is reached where the homunculi are no more than adders and

subtractors, by the time they need only the intelligence to pick the larger of two numbers when directed to, they have been reduced to functionaries "who can be replaced by a machine." (41)

Dennett replaces the little man, homunculus, with a group of homunculi. I do not see how Dennett is going to avoid this regress. Fodor has pointed out that Dennett faces a regress and has to drop the homuncular anthropomorphising. Apparently, things will boil down to the fact that the intentional stance that one adopts is either an intelligent one or a dumb one. If it is an intelligent observing standpoint, then one may say that it is a homunculus of a sort, or it eventually comes down to one such homunculus. If it is a dumb state of affairs, then we may call it the functional role, and thus it may turn out that functionalists are right to claim that their RTM approach does not presuppose a homunculus. Therefore, we can see that the same intuition is at work against Dennett's intentional stance as well. And Biro is right to say that "the problem in question is everyone else's problem, not Hume's and that he in fact comes closest to having anything like a solution to it" (42).

I think that the solution that Hume puts forward is just the one that Dennett has proposed to replace it, namely, homuncular functionalism. As I have tried to show above, Hume distinguishes between two distinct levels of intentionality, i.e., personal level and sub-personal level. On the one hand, Hume uses certain intentional verbs to describe the sub-personal level as the subject of intentionality, and that how each perception is a representation of an object having mental content in the mind, and on the other hand, it seems that, for Hume, it is ultimately the personal level that is the real subject of intentionality. It seems that this has made Dennett ascribe the so-called "Hume's Problem" to him. Dennett does not undermine Hume's role in the history of functionalism, and in fact, credits him with formulating the central problem of cognitive science. But he thinks that Hume faces the homunculus problem simply because he lacks the notion of an automatic computing machine, something like the theory of Turing machine functionalism (43). I disagree with Dennett that the bundle theory could not have worked because it needed the explanatory mechanism, which later on Artificial Intelligence (AI) provided us with, namely, the mechanism in which organised and automated dumb homunculi can collectively give rise to mental properties. Of course, I am not interpreting Hume as a late twentieth century functionalist because that would be simplistic and unsympathetic to the spirit of Hume's philosophy. It seems, however, that,

for Hume, each perception as the subject of intentionality in the sub-personal level functions as if it *is*, or *similar to*, an organized and automated dumb homunculus playing its role in the functioning of the mental system as a whole. As explained above, although Hume uses certain intentional verbs to describe the sub-personal level, it is ultimately the personal level that is the real subject of intentionality. Hume's perceptions operate just like Dennett's notion of "robotic neurons" (44), or members of a committee sitting in a meeting and participating in the process of decision-making. For Hume, perceptions are simply members of a collective entity, namely, the bundle of perceptions, each playing a causal/functional role within the system as a whole. Therefore, to a certain extent Hume's bundle theory of the self is a precursor to what Dennett calls homuncular functionalism.

It is important to note that Hume's strategy is to refute not only the substantial theory of the self, as its main contender, but also denies that the idea of the self is derived from a sort of *introspecting* experience. What he finds in introspection is, indeed, only perceptions, or a collection of perceptions, giving rise to the idea of the "bundle self." Hume's notion of the introspecting self is similar to Dennett's notion of the homunculus self. And since he rejects the idea of the introspecting self, I must conclude that he rejects the notion of the homunculus self as well. If what I am suggesting is correct, then, Hume could be interpreted as someone who, long before Dennett, offered a solution to avoid the homunculi problem. Essentially, Hume's bundle theory is a psychological mechanism that explains how a collection of perceptions by means of certain dumb and mechanical processes generate a notion of a self that is the subject of thinking, having beliefs, desires and other cognitive states and dispositions.

According to Biro, another situation in which Hume's conception of cognitive science is observed is the Problem of the Qualia and the case of the Missing Shade of Blue. Basically, the idea is that Hume's dismissal of the case of the missing shade of blue is in line with the cognitivist view that many of our mental processes are not introspectable and that those processes involved in the production of any idea may be unavailable to consciousness. This reading is influenced by two main sources: the notion of absent qualia and the pictorial interpretation of Hume's copy principle. Ned Block, for example, has argued that it is possible for a functional entity to satisfy all the functional descriptions, and yet lack the qualitative content or subjective feature of our experience,

i.e., qualia, associated with the processes specified by those descriptions ⁽⁴⁵⁾. Kenneth Bower has also put forward an interpretation of Hume's "copy principle", according to which mental representation need not be in terms of exact imagery or pictorial representation; the exact representation may be conceived in terms of information-theoretic terms, which only requires that all the information in an impression be present in the corresponding idea ⁽⁴⁶⁾. Biro thinks that representation understood in information-theoretic terms, that is, non-pictorial representation, does not depend on qualitative content, or qualia. For Biro, the argument applies the other way around too. He argues that "there seems to be no reason why one could not come up with an idea in pictorial form, even though the information embodied in that idea is not present in the same pictorial form in any particular prior impression" ⁽⁴⁷⁾. The implication of this for the case of the missing shade of blue is that "the information for "deriving" the idea of the previously unencountered shade of blue is there in the impressions one has had of the *other* shades one has encountered and in the relations among these, and there is no reason why it could not be re-constituted—usually quite subconsciously, at the sub-personal level where most of our perceptual processing takes place—in the form of a novel image" ⁽⁴⁸⁾. The upshot of this argument is that Hume's dismissal of the case of the missing shade of blue shows that his theory of mind leaves room for the sort of cognitivism that contemporary cognitive science calls for.

To conclude, these situations show that many of our mental processes occur while largely unavailable to our consciousness. Arguably, we call those mental processes and events as automatic and sub-doxastic states because they are typically distinct from the states of beliefs that we are in. In the state of belief, a thinker holds a belief about something, which can be characterized as a conscious accessibility to the mental content of that belief, whereas in the sub-doxastic state one has no conscious accessibility to the content of his mind. As it is shown above, Hume's theory of mind could be interpreted as a theory that makes room for such an interpretation.

So far we have seen that Hume's new science of human nature share, and to some extent anticipates, the basic principles of contemporary cognitive science. One of the main ideas of cognitive science is that the best way to understand human behaviours is in terms of internal states and processes that underlie such behaviours. In other words, one

of the most striking features of our cognitive capacities and operations is that those internal states, processes, mechanisms, and performances must be thought of as subdoxastic, modular, and automatic. I have shown some parallels between Hume's concerns and some of those in contemporary cognitive science. Next, I turn to Hume's RTM.

II. 4: Hume's Representational Theory of Ideas

Hume's naturalism is twofold. On the one hand, he tries to show that none of our actions and beliefs 'arise from reason' or 'have a foundation in reason.' This *sceptical* phase of Hume's programme is intended to deflate the pretensions of reason. According to the sceptical phase, the traditional conception of man as a detached rational agent is inadequate and thus must be rejected altogether. On the other hand, he attempts to show the origin of our most fundamental ideas and beliefs in human nature. This *naturalistic* phase explains our ideas and beliefs "by appeal to the nature of our experience and to certain fundamental properties and dispositions of the human mind" (49). Hume proceeds with the positive causal explanation of the origins of our natural beliefs after he has established his negative sceptical deflationary arguments about the foundation of our beliefs and ideas in reason. Hume's methodological naturalism is the ground upon which he develops an empirical theory of the mind. This phase of Hume's programme shall be understood as an "early attempt to construct a naturalistic theory of the mind within the assumptions of Cartesian Representationalism" (50). In Fodor's language, "Hume is remarkably perceptive and remarkably prescient about the architecture of such theories; in particular, he's exceptionally good on *what else* you have to do if you want to run Cartesian Representationalism as an empirical option in cognitive psychology" (51).

Hume's theory of the mind is Cartesian in certain ways, even though he rejects the Cartesian Substance Dualism. I have shown, for example, that Hume is committed to a version of dualism—that is, 'bundle' or 'property' dualism'. But more importantly, Hume's Theory of Ideas is essentially a Cartesian one. We "cannot separate Hume's commitment to an empirically warranted cognitive psychology from his commitment to the Theory of Ideas because, as it turns out, the Theory of Ideas *is* the cognitive psychology that is warranted empirically" (52). Indeed, Hume tries to show how a Cartesian account of the mind and especially the "Theory of Ideas" "might be developed

into a naturalistic and empirically plausible psychology of cognition” (53). I continue to quote Fodor:

Hume saw that accepting...the “Theory of Ideas” is central to constructing an empirically adequate account of cognition; indeed, that it is primarily the commitment to the Theory of Ideas that determines what form an empirically adequate cognitive psychology must take. For Hume, as for our contemporary cognitive science, the mind is preeminently the locus of mental representation and mental causation. In this respect, Hume’s cognitive science is a footnote to Descartes’s, and ours is a footnote to his (54).

I quote him again:

“Hume’s *Treatise* is the foundational document of cognitive science; it made explicit, for the first time, the project of constructing an empirical psychology on the basis of a representational theory of mind; in effect, on the basis of the Theory of Ideas” (55).

Hume’s commitment to the Theory of Ideas is unshakable. Historically speaking, he adopts his theory of ideas from his predecessors, most notably Descartes, without much exposition or improvement. As a representationalist, Hume holds that perceptions (or concepts) are mental particulars and that they represent objects in thought. According to Hume, “representing things in thought *is* the defining function of concepts and the like” (56). The functionalist account of mind consists of a Representational Theory of Mind (RTM) that postulates a system of symbols, which function as mental representation. In other words, functionalism claims that cognitive processes are computational processes and as such presupposes a medium of computation—that is, a representational system.

The proponents of RTM argue that it can be defined as having certain loose, but connected, theses (57). Among these theses, three are relevant to my discussion. I explain each of them briefly:

First, “mental representations are the primitive bearers of intentional content” (58).

RTM takes our common sense mental states like thinking, believing, desiring, wishing, and fearing, known as *propositional attitudes*, to be *about* or *refer to* things. These states of referring to or being about have ‘intentional properties,’ a kind of *relation* between the object that is represented and the mental representation, more broadly called the mental object. Moreover, RTM claims that propositional attitudes are ontologically *real*.

Intentional Realism about propositional attitudes is a fundamental component of RTM.

Intentional realism claims that when we adopt a realist view about propositional attitudes we stand in a certain relation to an internal representation. Furthermore, intentional

realism entails the possibility of laws about intentional states. Thus, “the laws that psychological explanations invoke typically express causal relations among *mental states that are specified under intentional description*, viz. among mental states that are picked out by reference to their contents. Laws about causal relations among beliefs, desires, and actions are the paradigms” (59). So, intentional explanation is typically nomic.

Furthermore, RTM claims that propositional attitudes are understood as mental particulars with semantic properties. In other words, to say that propositional attitudes have intentional content is to say that they have certain semantic properties. And to say that propositions have semantic properties is to say that propositions are objects with truth conditions that must be satisfied or, in short, semantically evaluated. For example, the proposition *P* that ‘it’s raining’, can only be satisfied if, in fact, ‘it’s raining.’ To flesh this out a bit further, according to RTM, token mental representations are symbols and these tokens of symbols are physical objects with semantic properties. To put it differently, mental representations belong to a symbolic or representational system in which representations of the system have both combinatorial syntax and semantics. It further distinguishes between structurally complex (molecular) representations and structurally simple (atomic) representations in which the former is systematically built up of the latter constituents. The “semantic content of a (molecular) representation is a function of the semantic contents of its syntactic parts, together with its constituent structure” (60). This feature of RTM holds that mental representation is committed to “symbolic structures” and, as a result, RTM postulates a ‘language of thought.’

In brief, the first thesis claims that for having a propositional attitude with the content *P*, there is a corresponding event that relates the agent to a token of mental representation that has the content *P* (61). The first thesis furthermore claims that the existing relation *R*, which holds between the mental representations in one’s so-called ‘belief box’ and the world outside the mind, is intentional. The final claim is that such intentional relation is semantically evaluable.

Second, RTM claims that thinking is computational (62).

This thesis concerns the nature of mental *processes* rather than mental *states*. It simply states that mental processes consist of causal sequences of tokenings of mental representation (63). In other words, it states that mental processes are causal relations

among mental representations. It furthermore argues for a causal relation between, for example, a tokening of a mental representation that “it’s going to rain” and a tokening of another mental representation that “one must take the umbrella.” Fundamentally, mental processes, which operate upon any mental representation that satisfies a given structural description, are causally sensitive to the syntactic or formal structure of such representation. It is here that it has been said that RTM is a species of the so-called Computational Theory of Mind (CTM). According to this underlying idea, “it is possible to devise machines whose function is the transformation of symbols, and whose operations are sensitive to the syntactical structure of the symbols that they operate upon” (64). To quote Fodor: “token mental representations are symbols. Tokens of symbols are physical objects with semantic properties... [and] computations are those causal relations among symbols which reliably respect semantic properties of the relata” (65). In a nutshell, every causal process is a kind of computation, which consists of causal relations among symbols that preserve semantic properties or contents. Another significant aspect of such a causal relation among symbols is that it entails law-likedness among states.

Third, RTM states that “meaning is information” (66).

This thesis simply asserts that the content of a mental representation is bestowed upon it by its informational semantics. According to Fodor’s informational theory of meaning, concepts, or ideas, to use Humean terminology, are mental states which have representational properties carrying information about our environment. To put it differently, having a concept in the mind is for that mind to be in some kind of relation with the world; concept individuation (tokening) takes place in a nomic mind-world relationship. Here, I do not intend to pursue this subject further. Suffice it to say, Fodor’s informational atomism targets both the Inferential Role Semantics theory of meaning as well as Hume’s pictorial theory of meaning (67). Now I shall turn to Hume.

Hume on Mental Representation

Many philosophers now agree that Hume has a theory of mental representation. Don Garrett, for example, has ascribed a “naturalistic theory of representation” to Hume. Stroud, following Thomas Reid, has said that Hume subscribes to the common theory of ideas that originated with Descartes. Fodor has extracted a full-blown representational theory of mind out of Hume’s philosophy. Hume’s representational theory of mind (or

theory of ideas) requires a theory about the structure and content of perceptions or concepts. Hume holds that there are two types of mental particulars: *impressions* (roughly sensations) and *ideas* (roughly concepts) and that all mental particulars have semantic properties or contents. Furthermore, he maintains that the content of concepts comes from (copied from) the content of impressions. The copy principle is the theory which explains the structure and the content of perceptions for Hume. In particular, it explains mental causation and intentional contents of the concepts. It states “*that all our simple ideas in their first appearance are deriv’d from simple impressions, which are correspondent to them, and which they exactly represent*” (THN, 4). Fundamental to this principle is that there is no mental activity or thought unless there are impressions of sensation prior to the idea. This principle involves both a Resemblance Thesis (every simple idea has a resembling simple impression) and a Causal Thesis (every simple idea is partly caused by a simple impression). The copy principle implies that the structures of ideas are exactly copied from the structures of impressions and, thus, the former do not contain any more structural information than the corresponding impressions. Although Fodor has rejected this aspect of Hume’s copy principle in favour of some form of nativism of the architecture of the mind, he rightly points out that, even if we take away Hume’s empiricism and his copy principle, what “is left is a perfectly standard Representational Theory of the Mind, one that’s compatible with as much (or as little) nativism as the facts turn out to require” (68). To put it differently, whether ideas have the same structures as impressions or have some extra informational ingredients, “perceptual theory needs impressions to be of things in order to explain how perceptual judgments can be true or false” (69). Therefore, Hume’s cognitive psychology requires that perception starts with impression and ends with conceptualization. Even if it turns out that Fodor is right about the copy principle and the fact that the mind does add something to impressions in the course of moving from impressions to ideas, it is safe to say that Fodor’s nativism is not going to pose a serious problem for Hume’s representational theory of mind insofar as his commitment to the copy principle is unshakable.

According to this reading, the mind is the locus of mental representation. Hume seems to be committed to the view that all ideas, whether of memory or imagination, represent (70). According to him, “ideas always represent their objects or impressions; and

vice versa, there are some objects necessary to give rise to every idea” (THN, 157). Or in another occasion he claims that “[i]deas always represent the objects or impressions, from which they are deriv’d, and can never without a fiction represent or be apply’d to any other” (THN, 37). Nevertheless, Hume admits that some impressions, namely, the passions do not represent:

A passion is an original existence, or, if you will, modification of existence, and contains not any representative quality, which renders it a copy of any other existence or modification, when I am angry, I am actually possest with the passion, and in that emotion have no more a reference to any other object, than when I am thirsty, or sick, or more than five foot high. (THN, 415)

Annette Baier calls this passage “unfortunate” and a “very silly paragraph” (71). Garrett calls it “paradoxical” because “Humean passions are typically, if not universally, *of, at, or for* someone or something” (72). But “Hume does not deny that representation is involved in having passions; he claims only that the passion itself is a non-representational impression of reflection, while the representation content involved in the passion’s being *of, at, or for* someone or something is provided by one or more associated ideas” (73). Moreover, he seems to hold that moral and aesthetic sentiments, which are impressions of reflection, are associated with ideas having representational content. There is even the possibility that moral and aesthetic sentiments themselves represent certain qualities such as virtue and vice, beauty and deformity (74). Hume also claims that certain features of the mind’s own operations such as the feelings of determination or ‘necessity’ and facility or ‘ease’ (in making a mental transition), which are impressions of reflection, have representational contents. Hence, we may conclude that for Hume some impressions of reflection represent, while others do not. But what about Hume’s impressions of sensations?

Hume also recognizes that impressions of sensation can represent bodies and their qualities, though in an imperfect way. He claims that “the only defect of our senses is, that they give us disproportion’d images, and *represent* as minute and uncompounded what is really great and compos’d of a vast number of parts” (THN, 28). In his discussion of the idea of space he argues that “compound impression, which represents extension, consists of several lesser impressions, that are indivisible to the eye or feeling” (THN, 38). Finally, in the *First Enquiry* Hume claims that “no man, who reflects, ever doubted, that the existences, which we consider, when we say, this house and that tree, are nothing

but perceptions in the mind, and fleeting copies or representations of other existences, which remain uniform and independent” (EHU, 12).

However, some commentators have objected that impressions generally do not represent because “ideas are derived from impressions, not the other way around. [and] so ideas can represent, but impressions [including impressions of sensation] cannot”⁽⁷⁵⁾. There are some textual supports for this reading in the *Treatise*. On one occasion, Hume argues that the impressions of touch “neither represent solidity, nor any real object” (THN, 231). Elsewhere, he suggests that since nothing is ever present to the mind but perceptions, and that ideas are derived from the impressions, it must follow that we are incapable of forming any idea different from our ideas and impressions (THN, 67). So they concluded that because “[w]e cannot even conceive what it would be like for a perception of the mind to resemble something that is not a perception of the mind”, it must follow that we are not in a position to make any claim about the possible relation between an impression and the related external object ⁽⁷⁶⁾.

To respond: first, the case of the impressions of touch is a unique case and does not seem to undermine Hume’s general theory of representation. In respect to the conceivability of external objects as distinct from perceptions, it must be noted that “Hume’s remark does not at all entail that bodies distinct from perceptions of the mind are *inconceivable*; it entails only that we cannot conceive of them as being specifically different from ideas and impressions”⁽⁷⁷⁾. As Hume immediately says, we do not “suppose them specifically different; but only attribute to them different relations, connexions, and durations” (THN, 67). In fact, this passage suggests that we cannot comprehend the specific intrinsic qualities of the bodies, but only understand certain specified relation to objects. In other words, we are only equipped to conceive bodies as entities that are conceived through the medium of perceptions in the mind. It seems that impressions of sensation are caused by and resemble, and hence represent bodies.

To summarize, Hume holds that impressions, whether of reflection or sensation, can represent. The impressions of sensation can represent external objects and their qualities. It is only some impressions of reflection, like the passions, which are not representational, whereas other impressions of reflection, like moral and aesthetic qualities, represent their objects. Ideas, however, represent a wider range of objects than

impressions. Ideas not only represent their objects or impressions and bodies, but they can also represent other ideas. Garrett even suggests that ideas can represent third order types, i.e., ideas *of* ideas of ideas (THN, 106). Moreover, ideas are representational in a wide range of ways: some ideas represent objects as individual, while others represent as part of a compound (THN, 8, 10). Ideas can also represent the objects either in the past or in the present (THN, 9), as related or unrelated (THN, 14-15), as substances or modes (THN, 16-17), and as particulars or generals (THN, 18). Ideas are representational through and through. If what is said so far is true, then we may suggest that Hume is committed to a version of RTM. But what about its intentionality?

Hume's explanation of intentionality specifically occurs in the context of our belief in the existence of an external world. David Raynor has argued that Hume, in response to Malebranche and Berkeley, attempted to "show *how* mind-dependent perceptions can represent mind-independent objects." He argues that "Hume had to show how the mind, without design, creates its ideas of external things, and does so by attributing to its perceptions "different relations, connexions and durations" (78). Generally speaking, it seems quite fair to say that Hume's conception of mental *states* is that they are relations to mental representations and as such they are intentional. It is a fact that Hume holds that perceptions, whether impressions or ideas, can at least be *about*, *of*, or *refer*, to something or someone. I assume that this constitutes Hume's notion of intentionality. I conclude that Hume has a representational theory of mind that has the following features: "First, there exist "Ideas," which are a species of mental symbol. Second, having a belief involves entertaining an Idea. Third, mental processes are causal associations of Ideas. Fourth, Ideas are like pictures. And fifth, Ideas have their semantic properties by virtue of what they resemble: the Idea of John is about John because it looks like him" (79). Enough has been said about Hume's notion of intentionality. I now turn to the question of computation.

Hume on Computation

Hume holds that mental processes are *causal* relations among mental representations. This constitutes Hume's account of mental *processes* and the nature of thinking as computation. Hume denies that ideas are *intrinsically* representational. He explicitly states that "the reference of the idea to an object [is] an *extraneous*

denomination, of which in itself it bears no mark or character” (THN, 20; emphasis added). Descartes held that ideas contain, by nature, both “objective reality” as well as “formal reality,” and that ideas are intrinsically representational because they have objective reality ⁽⁸⁰⁾. On the contrary, Hume seems to be only committed to the formal reality of ideas and, thus, rejects the claim that ideas are intrinsically representational. The reason that Hume rejects the intrinsic representational character of ideas is that he cannot reconcile it with the claim that passions are not representational. Moreover, it is incompatible with his thesis that there are only seven philosophical relations ⁽⁸¹⁾.

If ideas are not intrinsically representational, then they are to be dependent upon some “extraneous denomination.” According to Hume, representation depends on extrinsic denomination like the relations of *resemblance* and *causal derivation*. Resemblance is the necessary condition for representation; the ideas one forms are exact representations of the impressions one feels (THN, 3). Elsewhere, Hume argues that “how can an impression resemble a substance, otherwise than by resembling it” (THN, 233). However, resemblance might be important for representation but not sufficient for it ⁽⁸²⁾. It is possible that objects can resemble other objects without representing them. For instance, impressions can resemble ideas but do not represent them. But what about causal derivation? Can it provide the sufficient condition for intrinsic representation? Garrett, for instance, remarks that even resemblance and causal derivation combined together, which makes the Copy Principle, are insufficient for intrinsic mental representation ⁽⁸³⁾. There are indications in Hume’s texts that even suggest that resemblance and causal derivation do not appear to be necessary for inherent representation either. In order to flesh this out, I shall distinguish between mental and non-mental cases of representations in Hume’s cognitive theory.

Among the most obvious cases of non-mental representations are *words*, which may represent objects and facts (THN, 113); theatre, which may represent the “determinate portion of space” where the real action of a play has occurred (EHU, 3 editions E-N only); children, which can represent their parents’ families and the pride involved (THN, 309); money, which represents beautiful and agreeable objects by the power it affords of obtaining them (THN, 359); one action represents another, for instance, “the giving of stone and earth” can represent the delivery of a manor in property

law, or the conveying of the key to a granary can represent the delivery of corn contained in it, tapers, habits, or grimaces can represent religious mysteries (THN, 515); and the lower house of Parliament represents the entire commons (FPG, 35). In the cases of the lower house of Parliament and children, we can trace both some resemblances and causal derivations, while in the cases of theatre and religious symbolism, there are no original causal derivations but there may be some resemblances. In the case of words, there might be causal derivation but no resemblance. Lastly, in the case of money, it seems that there is neither resemblance nor causal derivation ⁽⁸⁴⁾. To sum up so far, it has been said that ideas are not inherently representational and that resemblance and causal derivation, which combined constitute the copy principle, are not sufficient, or even necessary, for intrinsic representation. Thus, the question that was raised above—on what “extraneous denomination” representation depends upon?— is still pressing for an answer. Assuming that resemblance and causal derivation are insufficient for intrinsic representation, then in what do mental representations consist? The answer to this question brings us to the core of Hume’s computational theory of mind.

That “extraneous denomination” upon which representations depend is the causal and/or functional role of the represented objects or events. “If we examine more closely what all [Hume’s] examples of non-mental representations have in common,” Garrett argues, “we find that, in each case, the representation plays at least a significant part of the causal and/or functional role of what is represented; and it does so specifically in virtue of generating (a) mental effects such as associated ideas, beliefs, sentiments, passions, and volitions; and/or (b) dispositions to further mental effects, such as associated ideas, beliefs, sentiments, passions, and volitions” ⁽⁸⁵⁾. It is not hard to see that in the cases of non-mental representation Hume discusses them in order to show that representation stands causally or functionally in relation to the things represented by producing mental effects and disposition-inducing effects. As a result, representation gives rise to certain effects or dispositions such as dramatic spectacles in the case of theatre, passions in the case of children’s family pride, religious faith in the case of religious mysteries, and political action in the case of the Parliament.

One may object that these kinds of non-mental representations have no bearings on Hume’s mental representation. However, Hume conceives the mind as an integral part

of human nature, and thus it is not possible that the cases of non-mental representations receive different treatments in Hume's philosophy of mind ⁽⁸⁶⁾. More importantly, for Hume, the mind is the locus of representation, whether mental or non-mental. As I have argued earlier, the mind, in Hume's view, is an entity that unites its own experiences—impressions and ideas—providing us a unified picture of the world. According to Hume, our perceptual experiences are fragmentary, piecemeal, interrupted and discontinuous, and the mind is ordained to unify these fragmentary experiences and generate a coherent conception of the mind. The mind consists of perceptions, each with its own ontological independence. The world, similarly, is made of atoms, continuing in flux and motion. There is neither perfect identity in the world, nor is there perfect identity in the self. But the mind comes up with the ideas of object and personal identity. This suggests that, for Hume, external reality is only conceivable through the medium of perceptual experiences. The outer reality may exist outside the mind, but it cannot be conceived without the mind. In a similar fashion, representation, whether mental or non-mental, without the mind is inconceivable.

To resume my discussion, effects may function in two distinct ways. First, in some cases, the representation plays the role of a *causal indicator*, which functions “as a causal intermediary by which that cause produces its typical effects in the mind.” For instance, a word can indicate the presence of the thing named and thus functions as a causal intermediary; or the lower house of Parliament can be a causal indicator of the political desires that caused the election of its members and thus functions as a causal intermediary. Second, in other cases, “the representation plays the *functional* role of the represented object by producing mental effects and dispositions that *model* or *replicate* parallel effects and disposition typically produced by the represented object itself in similar or parallel circumstances” ⁽⁸⁷⁾. For example, the stage set of a play replicates and produces parallel effects and dispositions in the minds of the audience members. Furthermore, the mental representations have the same structure as the non-mental ones. In a similar way, a sense impression can indicate a physical quality such as heat and hence function as a causal intermediary. An idea of an object, like a dark cloud, can stand in parallel relation to the qualities of the cloud itself, which may produce the ideas of the cloud's usual effects, such as rain. Hence, it is fair to say that both mental and non-mental

representations play, by means of producing mental effects and mental dispositions, the causal and/or functional role of the object it represents. “*Mental* representations—i.e., perceptions that represent—constitute, on this Humean view,” Garrett asserts that “a special case of representation: they not only play the causal and/or functional role of what they represent by evoking mental effects and dispositions, but they are also *themselves* mental entities. Ideas, in turn, constitute yet a further special case: for they are a class of mental representations that are copied from other mental entities and always have other mental entities as their proximate causes” (88). One can hardly deny this conclusion.

Three points are in order here: First, as it is known, some ideas, in Hume’s view, involve beliefs, (i.e., the idea of the self, the ideas of necessary connection, and the idea of the continued and distinct existence of objects) and other simple ideas do not involve belief. I should emphasize that the same causal and/or functional roles that simple impressions and ideas play in representing their objects is at work in the cases of the ideas that involve belief. Second, an idea could represent another idea or even a group of, or a bundle, of ideas. Most notably, the idea of the self represents the entire bundle of perceptions constituting the mind (self). The idea of the self clearly plays the causal and/or functional role of the entire bundle of perceptions by means of the mental effects that the bundle produces. Third, the causal and/or functional roles of ideas within the mind go beyond the non-mental representations that are explained thus far. Perceptions in general and ideas in particular, can represent bodies, minds, persons, individuals and compounds, modes and substances, and particulars and generals. The main idea is that the representational content of an idea on a particular occasion is still determined by its causal and/or functional capacities. A unique case that the representational content of ideas is determined by their causal/functional capacities occurs in the case of particulars and general kinds. Hume holds that an idea can represent an entire kind when it becomes an abstract or general idea. In his view, qualitatively identical ideas can represent a particular, i.e., a triangle, on one occasion, and an entire kind on some other occasion. An idea can become an abstract idea by being associated with or annexed to, via custom, a general term, “which gives them a more extensive signification, and makes them recall upon occasion other individuals, which are similar to them” (THN, 17). Therefore, the general idea thereby plays the functional roles of all the members of the kinds combined

together. Therefore, for Hume, the idea of a particular equilateral triangle can represent on different occasions either a particular equilateral triangle, or all equilateral triangles, or a figure, or a rectilineal figure, or a regular figure, or all triangles (THN, 21). So a particular idea can represent a general kind by being annex'd to a general term; "*some ideas are particular in their nature, but general in their representation*" (THN, 22).

In the same manner, a complex idea that is composed of simple ideas, connected in the mind by the relations of resemblance, causation, and contiguity, can represent a *substance* or a *mode*. Such a representation, however, does not depend upon the identity of the simple ideas; rather, it derives from the causal/functional roles of the complex idea. Thus Hume writes:

The idea of a substance as well as that of a mode, is nothing but a collection of simple ideas, that are united by the imagination, and have a particular name assigned them, by which we are able to recall, either to ourselves or others, that collection. But the difference betwixt these ideas consists in this, that the particular qualities, which form a substance, are commonly referred to an unknown something, in which they are supposed to inhere; or granting this fiction should not take place, are at least supposed to be closely and inseparably connected by the relations of contiguity and causation. The effect of this is, that whatever new simple quality we discover to have the same connexion with the rest, we immediately comprehend it among them, even though it did not enter into the first conception of the substance. (THN, 16)

In regard to *modes* Hume writes:

The simple ideas of which modes are form'd, either represent qualities, which are not united by contiguity and causation, but are dispers'd in different subjects; or if they be all united together, the uniting principle is not regarded as the foundation of the complex idea. The idea of a dance is an instance of the first kind of modes; that of beauty of the second. The reason is obvious, why such complex ideas cannot receive any new idea, without changing the name, which distinguishes the mode. (THN, 17)

Accordingly, a complex idea can take on the causal/functional roles of all the members of the collection, namely, the simple ideas combined together. Here, I should also add that ideas can represent things as parts of a compound, and that the ideas as a compound play almost the same causal and/or functional roles of the members of the compound itself (THN, 10).

Furthermore, ideas can represent things standing in *relations*. It seems that Hume thinks that ideas can represent spatial, temporal, and other kinds of parallel relations. This

could be seen in his remarks about the relation of cause and effect. He claims that “we must not here be content with saying, that the idea of cause and effect arises from objects constantly united; but must affirm, that it is the very same with the idea of these objects” (THN, 405). Moreover, ideas can represent bodies. Most notably, for the “vulgar”, that is all of us most of the time, the ideas represent objects as they have been conceived. Even from the philosopher’s point of view, ideas can represent objects as well. Hume has famously said that the hearing of a sudden noise as of a door turning upon its hinges, following with the sight of the porter and then receiving a letter from his friend thousands of miles away, generate the lively ideas of a turning door, a lively idea of stairs, a lively idea of a human body (the porter), the idea of the ocean between him and his friend, the idea of his friend existing far away from him (THN, 196). These lively ideas represent the door, the stairs, the porter’s body, the ocean, and the friend as continuing to exist all along, and independently from Hume’s mind. Thus, ideas can represent bodies and objects. Finally, it seems fair to suggest that an idea may represent a mind or a self (person) as well. I have already argued that the idea of the self represents the entire bundle of perceptions. The very basic point here is that the idea of the self “presumably represents the entire bundle by functioning as an appropriate abstract idea—that is, as an idea of some particular perception or perceptions together with a disposition to call up as necessary ideas of other perceptions suitably related by causation and resemblance” (89). There is, furthermore, the possibility that the idea of the self may represent a person in the holistic sense of both mind and body as causally related.

In sum, we have seen that impressions can represent objects, minds or persons, with their qualities, whereas ideas represent impressions, and other ideas as well as objects, minds or persons. It is also said that ideas can represent things in different modes. It can represent individuals, or parts of a compound, or as a compound; modes or substances; objects in relation or out of relation; as particulars and general kinds. Nevertheless, I have argued that perceptions are not intrinsically representational, “the “extraneous denomination” on which the thing’s “representational capacities depend are nevertheless closely and essentially related to *mentality*—for it is only in the networks of causally related perceptions that constitute *minds*” (90). As a result, Hume rejects the old Cartesian theory of mental states according to which they possess ‘objective reality’ and

makes them intrinsically representational, in favour of the view that their representational feature is an “extraneous denomination” based on causal capacities limited by their qualitative characters. On this account, perceptions represent because of the associative, inferential, sentimental, emotional, and conative causal networks in which they participate. In a nutshell, perceptions are representational in virtue of the causal and functional roles they play. The representational content of ideas is determined by their causal/functional capacities. I conclude that these arguments allows for a computational construal of Hume’s representational theory of mind. Now I should discuss Hume’s representational theory of meaning.

Hume on Meaning

Contemporary philosophers of mind do not accept the details of Hume's theory of mind, though they endorse it in general. Many cognitive psychologists and philosophers reject the idea that the pictorial theory of meaning is the main medium of mental representation. Fodor, for example, emphatically rejects Hume’s pictorial theory of meaning. In his view, what Hume’s representational theory of mind, particularly his functionalism, needs is an informational/semantic theory of meaning. However, Hume’s lack of an informational/semantic theory of meaning does not undermine his representationalism considerably. It seems that Hume’s empiricism, not his cognitive psychology, creates problems for his representationalism (91). Fodor claims that if we take away Hume’s empiricism, his motivation for the copy principle goes as well, and what is left is a perfectly standard Representational Theory of the Mind (92).

The main reason for rejecting Hume’s pictorial theory of meaning is that “ideas can’t copy their constituent structure from impressions because impressions don’t have constituents; all they have is parts” (93). Hume holds that the structure and content of concepts (ideas) comes from the structure and content of intact impressions, which explains why he is committed to the pictorial theory of meaning. But these “ideas contain more structural information, [constituents], than the corresponding impressions do” (94). Basically, Hume believes that the contents of complex ideas come from the contents of complex impressions. Fodor denies this claim because the content and structure of complex ideas is crucially different from the content and structure of complex impressions and, therefore, the former cannot be copied from the latter. Fodor states that

each complex idea, for instance, a BROWN COW, has a *canonical* decomposition into simple ideas, namely, BROWN and COW. Conversely, the complex impressions do not have canonical decompositions into simpler impressions; rather, a complex impression is the impression of parts of the thing. The impression of a BROWN COW can be sliced in any number of ways. Hence, Fodor, unlike Hume, who thinks that the structure of ideas are copied exactly—as images—from the structure of impressions, concludes that the imagistic or iconic mental representation fails because it cannot account for the fact that ideas have canonical decompositions. Therefore, the structures, or contents, of ideas are substantially different from the structure and content of impressions.

Basically, the content of impressions, including complex impressions, is pre-conceptual, or non-conceptual. Fodor relies on the experimental psychology of vision similar to that which is presented in the works of Bella Julesz's and his so-called *random dot stereograms* experimentation. Without going into the details of Julesz's experimentation, it has to be pointed out that the purpose of such experimentation was to show that two-dimensional images seem three-dimensional (95). The implication of this for Fodor's argument here is that sensory representations carry unconceptualized information. This tells us that the contents of our sensory perceptions are not entirely available to us. In other words, the unconceptualized information contained in our sensory perceptions is typically sub-doxastic, that is, impressions carry unconceptualized contents that are not consciously accessible to us (96). So, in principle, there is a fundamental difference between the contents of impressions and the content involved in the process of concept-formation. What these leave us with is this: concepts or ideas have more structural information than their corresponding impressions, and thus the contents of ideas are fundamentally different from the contents of its impressions. This means that ideas or concepts are not mere copies of impressions, and therefore, Hume's theory of concept formation, or his theory of meaning, does not depend entirely on his copy theory. This leads to the conclusion that Hume's pictorial theory of meaning is untenable.

Now the pressing question is this: Does the failure of Hume's theory of meaning entail the failure of his representational theory of mind? I suggest that Hume's pictorial theory of meaning does not stand in the way of his representational theory of mind. Hume's epistemology does not stand in the way of his cognitivism. The main reason is

that even though the content of impressions is non-conceptual and the structure, or the content, of the concepts is different from the structure, or content, of impressions, this does not mean that the content of impressions is unable to warrant concept-formation. It is so because concept-formation consists of a causal sequence of mental representations between an impression and a concept being formed. As long as such a causal sequence of mental representation is maintained between impressions and ideas, even though some extra ingredient is added to the structure and the content of ideas in the process of conceptualization, a representational theory of mind can still be secured. More specifically, since Hume's representational theory of mind contains elements of compositionality, mental causation, and intentionality, it can fairly be called representational. First, Hume's representational theory of mind can be defined as compositional because, for Hume, compositionality, though a kind of complex representation, still constitutes the relation between complex representations (complex ideas) and simpler representations (simple impressions). Hume, no doubt, holds that the mind consists of representational states, which are semantically evaluable and also that such states play certain causal and/or functional roles in cognitive processes in order to produce behaviours. The important point is that Hume distinguishes between the structure of simple and the structure of complex concepts and that the semantics of the former determines the semantics of the latter. More importantly, semantically speaking, all conceptual contents, whether simple or complex, are somehow reduced to simple impressions and to the contents that we receive through our sense experiences. This process satisfies the compositionality condition (thesis) in the sense that it requires that complex representations inherit their meaning, or content, from simple ones. Hence, Hume's pictorial theory of meaning may be false, or he should perhaps do away with his empiricism, but his associationism can produce a representational theory of meaning. Second, no matter how large is the difference between the structure of simple impressions and the structure of simple concepts it is a fact that Hume recognizes a robust theory of mental causation between impressions and ideas. Furthermore, the same thing could be said about Hume's notion of intentionality. I have discussed these issues above and shall not return to them again here.

Conclusion

In this chapter, I have discussed Hume's computational approach to the study of the mind. Here, I have tried to show that Hume's commitment to functionalism is more than the mere teleological functionalism; I have claimed that Hume's theory of mind is, in principle, computational. In particular, I have argued that his computationalism consists of a representational theory of mind. I have also discussed Hume's mechanics and the architecture of mind and their relations to his computationalism. Hence, I conclude that not only Hume's teleological functionalism is explicit in the text, but also, and more importantly, the text vindicates Hume's commitment to a naturalistic theory of representation, which is fundamental to his psycho-functionalism. Furthermore, I have discussed Hume's views on the architecture of the mind, and claimed that he is committed to the view of the mind as modular, sub-doxastic, and automatic.

In the next chapter, I will discuss Hume's psychology of identity ascriptions. I want to establish two important things there. First, I will explain how the ideas of object identity and personal identity, which persists over time by some form of functional continuity (identity), arise in the mind. Second, I interpret the underlying psychological mechanism that explains the ascription of identity to these successive but related objects as an essential case study for Hume's account of the sub-doxastic nature of mind.

Hume's Psychology of Identity Ascriptions

When we have been accustom'd to observe a constancy in certain impressions, and have found, that the perception of the sun or ocean, for instance, returns upon us after an absence or annihilation with like parts and in a like order, as at its first appearance, we are not apt to regard these interrupted perceptions as different, (which they really are) but on the contrary consider them as individually the same, upon account of their resemblance. But as this interruption of their existence is contrary to their perfect identity, and makes us regard the first impression as annihilated, and the second as newly created, we find ourselves somewhat at a loss, and are involv'd in a kind of contradiction. In order to free ourselves from this difficulty, we disguise, as much as possible, the interruption, or rather remove it entirely, by supposing that these interrupted perceptions are connected by a *real existence*, of which we are *insensible*. This supposition, or idea of continu'd existence, acquires a force and vivacity from the memory of these broken impressions, and from the propensity, which they give us, to suppose them the same; and according to the precedent reasoning, the very essence of belief consists in the force and vivacity of the conception. (THN, 199)

That action of the imagination, by which we consider the uninterrupted and invariable object, and that by which we reflect on the succession of related objects, are almost the same to the feeling, nor is there much more effort of thought requir'd in the latter case than in the former. The relation facilitates the transition of the mind from one object to another, and renders its passage as smooth as if it contemplated one continu'd object. This resemblance is the cause of the confusion and mistake, and makes us substitute the notion of identity, instead of that of related objects. However at one instant we may consider the related succession as variable or interrupted, we are sure the next to ascribe to it a perfect identity, and regard it as enviable and uninterrupted. Our propensity to this mistake is so great from the resemblance above-mention'd, that we fall into it before we are aware; and tho' we incessantly correct ourselves by reflection, and return to a more accurate method of thinking, yet we cannot long sustain our philosophy, or take off this bias from the imagination. Our last resource is to yield to it, and boldly assert that these different related objects are in effect the same, however interrupted and variable. In order to justify to ourselves this absurdity, we often feign some *new and unintelligible principle*, that connects the objects together, and prevents their interruption or variation. Thus we feign the continu'd existence of the perceptions of our senses, to remove the interruption: and run into the notion of a *soul*, and *self*, and *substance*, to disguise the variation. (THN, 254)

As described in the previous chapters, Hume's fictional self persists over time by functional continuity. Hume holds that the idea of the self, which has no origin in our

sense-impressions, is produced by a complex psychological belief-forming mechanism. In this chapter, I aim to, first, spell out the details of such psychological mechanism that produces the ideas of personal identity and object identity. Second, I aim to interpret Hume's psychology of identity ascriptions as a case study of his views on the *automatic* and *sub-doxastic* nature of mind. I should say a little more about these objectives in this introduction.

In the section of "Of Scepticism with Regard to the Senses," Hume discusses a "propensity," ascribing identity to a succession of related perceptions in order to explain the belief in the continued and distinct existence of external objects. Accordingly, perfect identity—identity in strict sense—requires an object to "remain uninterrupted and invariable thro' a suppos'd variation of time" (THN, 200-201). Yet, by philosophical reflection, we come to realize that the interruption of the perceptions seems contrary to the identity, "and naturally leads us to regard these resembling perceptions as different from each other" (THN, 205). As a result, we find ourselves involved in a contradiction (THN, 199, 205, 208). In order to resolve the contradiction, the smooth "passage of the imagination along the ideas of these resembling perceptions makes us suppose that such perfectly resembling perceptions have a continued and distinct existence (THN, 205). The propensity operates to unite these broken perceptions to produce the fiction of continuity of objects (THN, 205).

In the immediately following section of "Of the Ancient Philosophy," Hume attempts to explain how the same propensity of the mind generates the belief in a *substance*, or *original first matter*, *substantial forms*, *accidents*, and *occult qualities* (THN, 220). When we observe an object that undergoes gradual and successive changes, for example, an oak tree growing over the years, we, indeed, experience a succession of related perceptions. Like before, we are here presented with the idea of diversity. Again, perfect identity requires invariability and uninterruptedness of the object. As a result, and again like before, we are involved in a contradiction (THN, 220), and in order to reconcile this contradiction, we suppose, or "feign something unknown and invisible [object], which it supposes to continue the same under all these variations" (THN, 220). The ancients used to call that fictional object a substance or a material substratum.

In Hume's view, the same propensity is at work to produce other related ideas. In

the section of “Of Personal Identity,” Hume explains how the belief in an immaterial soul, or *self*, gets generated by the same propensity of the mind. Once again, perfect identity requires an object that is invariable and uninterrupted (THN, 253), but philosophical reflection tells us that the mind is, indeed, the collection of the succession of distinct but related perceptions. Similarly, we are involved in a contradiction, and as a result of it, the above mentioned propensity of the mind makes us ascribe identity to the succession of related objects (THN, 254). Thus, we suppose “some new and unintelligible principle, that connects the objects [perceptions] together and prevents their interruption or variation” (THN, 254). We feign “the notion of a *soul*, and *self*, and *substance*” (THN, 254).

These explanations conform to a general schema. In the case of the belief in external objects, we have successive sensible objects that are invariable, though interrupted; in the case of the belief in substances or substantial forms we experience successive sensible qualities that are uninterrupted, though variable; and in the case of the belief in the self or souls, we have successive perceptions that are both interrupted and variable. In all these situations, we have a succession of related perceptions that are interrupted or variable and that conform to a general psychological belief-forming mechanism responsible for the generation of the ideas of object and personal identity. This psychological mechanism consists of a multistage process: (1) due to a “propensity” or “inclination” we tend to ascribe identity to the succession of related perceptions (THN, 199, 203, 220, 253-4); (2) philosophical reflection makes us aware of the palpable interruption in the successive related perceptions, and thus, induces us to consider constant but interrupted perceptions as diverse and different (THN, 199, 205, 220, 254); (3) we find ourselves involved in a contradiction (because we take constant but interrupted perceptions both as identical and diverse simultaneously), and in order to reconcile this contradiction, we feign or postulate an ‘unperceived perception’ and thus, remove entirely the interruption by *supposing* that these interrupted perceptions are connected by a real existence, of which we are insensible (THN, 199, 205, 254). These psychological processes function uniformly and consistently in all the three cases discussed above. In the case of the belief in external objects, we suppose that perceptions have a continued and distinct existence during the interruptions when they are not

perceived; in the case of the belief in a substantial form we postulate an unknown and invisible object; and in the case of the belief in a soul, we suppose a self or substance that is uninterrupted and invariable. Hence, Hume has a parallel account for the cases of the belief in the self, the belief in external objects, and the belief in material substrata, and that the same sort of psychological belief-forming mechanism accounts for the generation of their ideas (1).

Hume sets out his full explanation of this psychological belief-determining mechanism in the section of “Of Scepticism with Regard to the Senses” and discusses it partly in the section of “Of the Ancient Philosophy.” When he arrives at the section of “Of Personal Identity” he does not provide a detailed account of the psychological mechanism involved there; rather, he briefly, in one paragraph (THN, 253-54), describes how the imagination produces the fictional idea of personal identity. Consequently, one can argue that since Hume does not fully explain the psychological belief-determining mechanism generating the idea of personal identity, it means that the same mechanism operating in the generation of the idea of object identity cannot produce any result in this case. Therefore, Hume does not have a parallel account of both object and personal identity. It seems to me that this sort of objection is, at best, a mistake. Hume does not explain the psychological belief-determining mechanism operative in the generation of the idea of the self simply because he has already, and fully, explained such a mechanism. In fact, in the section of “Of Scepticism with Regard to the Senses,” Hume explicitly says that “we shall afterwards see many instances of this tendency of relation to make us ascribe an *identity* to *different* objects” (THN, 204).

Hume’s major contribution to the philosophical discourse on the question of object identity and personal identity is that he explains the underlying psychological mechanism that generates our beliefs in an enduring self as well as the continued and distinct existence of objects. He insists that there is no substantial difference between the belief-ascription mechanism involving object identity and the one involving personal identity. For Hume, all that exists and philosophically matters is our perceptual experiences and the belief-determining mechanism about the surrounding world and ourselves. Of course, Hume does not deny the existence of mind-independent reality. He seems to be agreeing, at least, with the ‘vulgar,’ about a form of ‘naive realism.’ He holds

that it would be a fundamental mistake to think that the mind experiences itself, or knows itself fundamentally different from the way it knows the external objects. The mind knows itself no differently from how it knows any other objects. Hume “puts forward the need for a conception of mind that places it alongside the conception of objects and within the realm of the experienceable. Within this realm fundamental gaps between mental and non-mental entities do not exist: we seem to be able to experience our own mind as much as we seem to be able to experience the objects of our natural surroundings: we experience them in the same perceptual guises” (2). This means two things: first, all perceptions could be interpreted as external persisting objects. According to Hume, there are two kinds of objects: (a) mind-independent objects (external world), and (b) mind-dependent objects (perceptions). However, (a) objects are not fundamentally different from (b) objects because the mind-independent object is made accessible through the medium of mind-dependent objects. In other words, perceptions are the medium through which the surrounding world represents itself to us. In a sense, talking about the external world without talking about perceptions is an empty enterprise. Second, it follows that all that really matters in our psychological belief-determining mechanism is the perceptual system itself, which is responsible for the identity ascriptions. Hence, Hume can consistently say that the identity of persons and objects strictly consist in the psychological mechanism operating on us, or the natural propensity, forming the beliefs about ourselves and the objects in the physical world. Simply put, perceptions constitute the identity of persons and objects alone. Hume, unlike Locke who specified different criteria of identity for different objects, believes that there is one single psychological criterion of identity (Chapter IV). This gives him the ground to parallel his accounts of personal identity and object identity. The two accounts conform to a general schema. In part, in this chapter I want to show how Hume develops his parallel accounts of personal identity and object identity, and the psychological belief-determining mechanism that produces the belief in their continuity over time.

In respect to the second objective, it is important to note that Hume places high value on sustaining stable doxastic states. Hume has repeatedly talked about a feeling of uneasiness that arises from the opposition of two contrary principles (THN, 206), or a contradiction that gives rise to a “sensible uneasiness” (THN, 205). But “the mind...will

naturally seek relief from the uneasiness” (THN, 206; 215; 224) in order to reach a settled or stable state. In some cases, the uneasiness or contradiction cannot be resolved. “Since the uneasiness arises from the opposition of two contrary principles,” Hume says, “it [the mind] must look for relief by sacrificing the one to the other” (THN, 206). Or sometimes, since neither of the contrary belief gives way, the mind yields to “reluctance” (THN, 206) or it runs into a principle that is “obstinate” (THN, 215). In other cases, the contradiction leads into “confusion” (THN, 238), or into an “unintelligible” hypothesis (THN, 220, 224, 254), and that serves as to “conceal” (THN, 219) or “disguise” (THN, 254) the contradiction. In fact, Hume thinks that some of our most fundamental metaphysical beliefs—the belief in external object, the belief in material substrata, and the belief in immaterial soul or the self— which arise from the propensity of the mind to ascribe identity to related objects, are produced by way of concealing and disguise the contradiction. Accordingly, when the mind experiences a contradiction between, for instance, our propensity to ascribe identity to a succession of related objects and the palpable interruption or changes in the related objects, which induces the mind to consider them as different or diverse, it tries to reconcile or remove the contradiction by way of supposing the existence of an object that is strictly uninterrupted and invariable (THN, 199, 205, 208, 219, 220, 253-254). Contradiction is an unstable doxastic condition and feels uneasy, discomforting, and irritating. In order for the mind to remove the contradiction and sets itself at ease, or to relive itself from the disturbing and uneasy feeling, it has to sustain a stable and steady doxastic state. Such a doxastic state is a calm and tranquil condition.

Louis Loeb has argued that a stability-based theory of justification establishes that, for Hume, the states that result from the psychological belief-determining mechanism is *automatically* justified, simply in virtue of being derived from such a belief-determining mechanism. In this sense, a stable doxastic state is the natural function of belief; nature had endowed us with belief in order to sustain stability. The beliefs in external objects, material substrata, and the self all function to generate stable doxastic states. These beliefs are justified because they sustain stable and steady doxastic state. I argue that Loeb’s interpretation provides us with an essential case study that shows Hume’s firm commitment to the mechanical, automatic, and sub-doxastic nature of the

mind. Partly, my concern is to show how all the processes involved in the psychology of identity ascription remain under the threshold of reflective philosophical thinking. That is, Hume's psychology of identity ascription and especially the mechanisms of constancy and coherence operate without being accessible to our reflective consciousness. I suggest that Hume's psychology of identity ascription, including the mechanisms of constancy and coherence, is such an extraordinary occasion in which we can unequivocally observe his commitment to the *sub-doxastic* and *modular* view of mind.

Two more introductory points are in order: First, Hume talks about the continued and distinct existence of external objects rather than their identity. He explicitly distinguishes between the question of "why we attribute a CONTINU'D existence to objects, even when they are not present to the senses" and the question of "why we suppose them to have an existence DISTINCT from the mind and perceptions?" (THN, 188). Hume specifies that the distinctness of objects refers to "their situation as well as relations, their *external* position as well as the *independence* of their existence and operation" (THN, 188). He argues that these two questions are "intimately connected together" simply because "if the objects of our senses continue to exist, even when they are not perceiv'd their existence is of course independent of and distinct from the perception; and *vice versa*, if their existence be independent of the perception and distinct from it, they must continue to exist, even tho' they be not perceiv'd" (THN, 188). However, Hume, in fact, "devotes almost all his attention to the first question, treating the second as subordinate to it" (3). That is why Hume exclusively focuses on the question of the continuity (identity) of objects rather than on their distinct and independent existences from the perceiving mind. It is not clear why Hume neglects the question of objects' distinctness from the perceiving mind. Two possible answers come to mind: The first answer has to do with his claim that the two questions are intimately connected together. Apparently, Hume thinks that his treatment of the question of continuity is equally applicable to the question of distinctness.

Second, according to Hume, there are two different kinds of object: (1) the objects whose spatial characteristics, viz., shape, size, pattern, etc., change (variable objects); (2) the objects whose spatial features do not vary (invariable objects). The first category of objects include all the organic bodies and cases like fire (Hume's example), and the

second category of objects include mainly the inorganic bodies. Hume does not provide two different types of explanations for these different kinds of objects; rather, he thinks that the series of perceptions involved in these different categories of objects affect the mind differently. In the first category, the effect shows itself in the form of an interrupted but variable series of perceptions exhibiting coherence, and in the second category the effect shows itself in the form of an interrupted and invariable series of perceptions exhibiting constancy.

In III.1, I will sketch out a four-stage psychological mechanism which produces the idea of the continued and distinct existence of external objects. In III.2, I will distinguish between the principle of constancy and the principle of coherence. In III.3, I explain Hume's principle of constancy. In III.4, I examine the principle of coherence.

III.1: The Continued and Distinct Existence of External Objects

There is general agreement that the aim of Hume's philosophy is to explain the origins of certain ideas, i.e., the idea of necessary connection, the idea of the continued and distinct existence of external objects, and the idea of the self. In fact, Hume's central preoccupation is to account for the *genesis* of the psychological causes of the universal but mistaken *beliefs* in causal necessity, the continued and distinct existence of bodies, and personal identity. Hume uses the same underlying mental mechanism to explain the origins of our beliefs in these ideas. He uses the same general structure to explain the origins of all these ideas. According to him, we must get these ideas and beliefs either from 'reason,' or from the 'senses,' or from the 'imagination.' He, however, argues that reason and the senses cannot determine us to believe in the existence of these ideas and beliefs. In consequence, he holds that the imagination, a natural and primitive disposition of the mind, is the only mental belief-determining mechanism that produces our ideas and beliefs in causal necessity, object and personal identity. In fact, the basic idea is that in the same way that we come to believe that the external bodies continue to exist over time without being perceived by us, likewise, we believe that we have an idea of the self that continues to be the same over time in spite of the fact that we know that there is no 'perfect' identity involved there.

Hume begins the section of "Of Scepticism with Regard to the Senses," which deals with the identity of objects, by asserting that "'tis in vain to ask, *Whether there be*

body or not? That is a point, which we must take for granted in all our reasonings” (THN, 187). He claims that we must assent to the principle concerning the existence of external objects. “Nature has not left this to his choice,” he argues, “and has doubtless esteem’d it an affair of too great importance to be trusted to our uncertain reasonings and speculations” (THN, 187). Putting aside this ‘naive realism’ of the ‘vulgar,’ Hume states that “[w]e may well ask, *What causes induce us to believe in the existence of body?*” (THN, 187). This explicitly indicates that Hume’s primary concern is to explain the cognitive system that causes one’s belief in the existence of external objects. This becomes clearer when Hume immediately points out that the above question raises the two following but related questions: First, “why we attribute a CONTINU’D existence to objects, even when they are not present to the sense;” and secondly, “why we suppose them to have an existence DISTINCT from the mind and perception” (THN, 188). The first question asks why do we suppose that objects endure or continue to exist when we have no perception of them; and the second question asks why do we suppose that objects are external to us? Hume’s chief point is that it is of no use to ask whether objects exist or not, because reason cannot establish that they exist; our human nature compels us to believe that they exist anyway. Instead, philosophy must aim at investigating those principles of the human mind that compel us to believe that objects continue to exist when they are not perceived and are independent from us. Accordingly, Hume’s programme is essentially a psychological one rather than a metaphysical one. In Hume’s view, we cannot seriously talk about the truth of our beliefs in the existence of objects and persons, though we take them for granted.

Hume remarks that we believe that objects continue to exist while they are not being perceived; and that they exist distinctly and independently of their being perceived. Since we have such a belief it follows that we must have an idea of an enduring and independent world. And because we have such an idea there must be an intelligible way to explain both the idea and our belief in a continued and distinct world. As pointed out earlier, Hume maintains that we can get the idea and the belief from three sources: the senses, reason, and the imagination. He shows that the first two cannot give rise to the idea and the belief in the enduring objects. As a result, he holds that the identity of external objects is something that we merely ascribe to them as a result of various

functions of the imagination.

The reason that the senses are incapable of giving rise to the idea of *continued* existence of object is that it requires the object to be continually perceived when, indeed, it is not being perceived as such. To put it another way, the senses generate ‘impressions,’ but impressions cannot give rise to our belief in objects. Objects are supposed to exist for a long period and continuously and independent from the mind, but impressions of sensation are relatively short-lived. They last for some short instances, or moments, and if they last longer, they will be discontinuous. Therefore, it seems that the senses are incapable of generating our belief in the continuous existence of objects. They only give rise to single perceptions and single perceptions cannot produce anything beyond themselves. At most, what the senses can provide us with is the idea of *distinct* objects. Moreover, Hume argues that impressions are parts of the mental world and as such “never give us the least intimation of anything beyond” themselves (THN, 189). That is, we are only aware of our impressions, and to say that we perceive something distinct from our impressions is already to ascribe external existence to those impressions.

Furthermore, Hume shows that ‘reason’ also cannot give rise to the idea of and the belief in the continued and distinct existence of objects. He argues that although philosophers may offer arguments in support of our belief in the existence of external objects and “may fancy they can produce to establish the belief of objects independent of the mind, ’tis obvious these arguments are known but to very few, and that ’tis not by them, that children, peasants, and the greatest part of mankind are induc’d to attribute objects to some impressions, and deny them to others” (THN, 193). On the one hand, the philosopher makes clear the distinction between impressions and the external objects, and “informs us, that everything, which appears to the mind, is nothing but a perception, and is interrupted, and depend on the mind,” and on the other, “the vulgar confound perceptions and objects, and attribute a distinct continu’d existence to the very things they feel or see” (THN, 193). The upshot of this is that the philosopher’s view, which tells us that whatever appears to the mind is nothing but interrupted perceptions, is incompatible with the vulgar view, which attributes a distinct and continued existence to many of his impressions. Since almost all of us believe in the distinct and continued existence of objects it must follow that it is not the understanding or reason that gives us

an assurance of the continued and distinct existence of body, but an entirely unreasonable sentiment, i.e., *imagination*. Something that is keen to the vulgar view.

The central theme of Hume's argument about our belief in the continued and distinct existence of objects, like his argument about the idea of the self, is that it is a simply a mistaken or fictitious one; we manifestly mistake a discontinuous object for a continuous one. Hume's philosophy is aimed at explaining the psychological principles involved in generating this false but indispensable natural belief. He argues that through philosophical reflections we come to know that the vulgar view, which attributes continued and distinct existence to external objects, is unfounded. Specifically, the vulgar believes that there is no distinction between perceptions and objects and that objects continue to exist distinct and independently from us, whilst we know philosophically that our perceptions are interrupted and depend on the mind. Therefore, the vulgar belief (which is the belief of all of us, at one time or other) in the continued and distinct existence of objects, must be a false idea. This falsehood and the mental mechanism responsible for its generation is exactly what philosophy is supposed to explain. The positive phase of Hume's programme is, indeed, to explain the origins of these mistaken ideas by way of explaining the principles of the mind that generate them.

There is also an important point that should be made clear. For Hume, talking about the continued and distinct existences of objects is, indeed, talking about the *identity* of objects. When we say that we believe in the continued and distinct existence of objects not being perceived by us, we exactly mean that the objects continue to persist over time. To explain the mental belief-determining mechanism involving the continued and distinct existences of objects is to delineate the conditions of identity for external objects. For him, to say that objects have continued and distinct existence outside of the mind is to say that these objects exhibit identity over time and a unity at a given time. Let me explain how this mechanism operates.

According to Hume, the propensity to ascribe identity to the succession of related objects operates uniformly in the mind of the vulgar as well as in the mind of the philosopher. The vulgar is unable to distinguish between objects and perceptions, whereas the philosopher is aware of the distinction. Despite the fact that we, as the philosophers, distinguish between objects and perceptions, we are still unable to resist the

natural propensity of the mind and, thus, “fall into it before we are aware; and tho’ we incessantly correct ourselves by reflection, and return to a more accurate method of thinking, yet we cannot long sustain our philosophy, or take off this bias from the imagination” (THN, 254). This means that the mind operates in a uniform manner, and that the belief in the continued and distinct existence of objects arises in the minds of the vulgar and the philosopher alike. This, in turn, entails that all that matters for Hume’s philosophy of mind is our perceptual experiences and the mechanism that generates certain fundamental philosophical ideas. It is worth emphasizing that this shows Hume’s commitment to a singular perception-based criterion of identity. For him, perceptual experiences and the mechanism involved are the only objects that one must deal with when one is concerned about determining the identity of objects or persons. That is why he declares “that there is only a single existence, which I shall call indifferently *object* or *perception*” (THN, 202). In short, identity is in the mind, and it will be in vain to look for it elsewhere.

Hume’s general schema for the explanation of the psychological causes of our belief in object and personal identity consists of different stages that the mind goes through in order to generate the ideas of such beliefs. I have introduced this mechanism briefly above. Now, I shall outline the general scheme that all the three cases conform to, and then, I will try to show how those stages operate, uniformly and consistently, in those three different situations. In particular, Hume distinguishes between four different successive stages of the mind leading to the production of our belief in object and personal identity. These stages are the followings:

- 1) When we are observing a constancy in certain successions of related impressions that are interrupted and numerically distinct but resembling and qualitatively identical, it triggers a psychological reaction according to which we ascribe identity to the successive associated perceptions due to a ‘propensity’ or ‘inclination’ that mistakes or confuses the resembling but successive series of perceptions with “one constant and uninterrupted perception” (THN, 204). For example, Hume claims that “the perception of the sun or ocean returns upon us after an absence or annihilation with like parts and in a like order, as at its first appearance, we are not apt to regard these interrupted perceptions as different,

(which they really are) but on the contrary consider them as individually the same, upon account of their resemblance” (THN, 199). Due to the propensity to ascribe identity to successive resembling perceptions, we *ascribe identity to constant but interrupted perceptions*.

- 2) However, the palpable interruptions and changes in the successive related perceptions *induce us to consider constant but interrupted perceptions as diverse and different* (THN, 199, 205, 220, 254). We take constant but interrupted perceptions as diverse and different because, on Hume’s account, identity in the strict sense requires uninterruptedness and invariableness of perceptions through a supposed variation of time. This is entailed by his principle of individuation.
- 3) Now, the fact that we take constant but interrupted perceptions both as identical and diverse and different, “we find ourselves somewhat at a loss, and are involv’d in a kind of contradiction” (THN, 199, 205). To reconcile this contradiction, we come to believe that objects continue to exist even when they are not perceived. In fact, “we disguise, as much as possible, the interruption, or rather remove it entirely, by *supposing* that these interrupted perceptions are connected by a real existence, of which we are insensible” (THN, 199, 205, 254). We feign or postulate an ‘unperceived perception’ that fills the gap between the series of interrupted perceptions and connects the perceptions together in order to remove the interruption or variation. As a result, we take an interrupted series of successive impressions as an uninterrupted series. This gives rise to our belief in the continuity of the series.
- 4) When we unite the interrupted and broken impressions by a continued existence we come to believe that objects continue to exist when they are not perceived. The arising belief consists in nothing, Hume claims, but the vivacity and force of an idea, and such an idea may acquire this force and vivacity by its relation to some impressions.

Hume goes on to explain the origins of our ideas or beliefs in the continued and distinct existence of objects and persons. Accordingly, it consists in nothing but the force, liveliness, and vivacity of certain impressions—for Hume, “impressions are naturally the most vivid perceptions of the mind” (THN, 208). To be precise, the idea of the continuing

object or the self acquires its vivacity from the memory of those broken and interrupted impressions, which is transmitted smoothly by that propensity of the imagination. This generates a *feeling* of performing an uninterrupted mental activity, which in turn, gives rise to the *belief* in objects and persons.

In sum, Hume holds that this four-stage psychological belief-determining system, as outlined above, provides the most profound explanation of our belief in the existence of objects and persons. In the first stage, we ascribe identity to a succession of related perceptions due to an “inclination” or “propensity.” The succession of related perceptions consists of interrupted and numerically distinct but resembling and qualitatively identical perceptions. This constitutes the constancy feature of the series of successive but related perceptions. What happens here is that the constancy of the series of perceptions triggers a psychological reaction according to which we ascribe identity to the successive associated perceptions due to a ‘propensity’ or ‘inclination’ that mistakes or confuses the resembling but successive series of perceptions with one constant and uninterrupted perception. It is clear from this that the operation of the propensity depends on an illusion or mistake; the mistake, or confusion, between an uninterrupted series and an interrupted but constant series of perceptions. In the next stage, one realises, by way of philosophical reflection (Hume’s principle of individuation implies that identity in strict sense requires uninterruptedness and invariableness of perceptions through a supposed variation of time), that the series is palpably interrupted. In the final stage, we find ourselves in a contradiction because we take constant but interrupted perceptions as uninterrupted and continuous. To reconcile the contradiction we disguise or postulate an “unperceived perception” in order to fill the interval between the interrupted series of perceptions and, thus, to connect the perceptions together in order to remove the interruption or variation. As a consequence, we take an interrupted series of successive impressions as an uninterrupted series. This ultimately gives rise to our belief in the continuity of the interrupted series. The explication of these stages in Hume’s philosophy heavily relies on the operation of constancy and coherence and their effects on the mind. I must explain these now.

III.2: The Distinction between Constancy and Coherence

According to Hume, our belief in the existence of unperceived objects results from the effects of constancy and coherence. Moreover, it is worth noting that Hume assigns different roles for coherence and constancy in generating our belief in external objects. He denies that they should be treated the same. However, there is disagreement among Hume's scholars about the proper way of characterizing coherence and constancy and the way that they generate our belief in the existence of the external world. The standard interpretation holds that there should be a *unified explanation* of the causes of our belief. Some philosophers, following H. H. Price, have conceded that Hume could have accounted for constancy as an special case of coherence (4). Hume, however, explicitly treats constancy as entirely separate from coherence, and even more important than coherence. This has become the source of the main objection, which is levelled against the standard interpretation according to which Hume was satisfied that his treatment of coherence, and treating constancy as a special case of coherence, is sufficient for the generation of our belief in external objects. Most notably, Jonathan Bennett (5) has argued that although Hume seeks to provide a unified explanation, he realizes that the coherence principle is inadequate alone. As a result, they argue, Hume expresses dissatisfaction with the coherence-based explanation and rejects it in favour of a constancy-based explanation. Yet, the *dissatisfaction hypothesis* (6) faces problems of its own; they must explain why Hume rejects his coherence-based explanation. Hume, however, does not provide any serious argument about the faultiness of the coherence principle. He simply points out that coherence is not adequate to do the job *alone*. Nevertheless, doubts have been recently raised against different versions of the *unified explanation hypothesis*. Some scholars have been arguing that it is unnecessary to try to reduce constancy to coherence. They argue that constancy is not the same as or reducible to coherence. Hume needs both coherence and constancy to do two different things for him (7).

Following this new promising interpretation, I argue that Hume needs one explanation for cases of coherence and a quite different explanation for cases of constancy. The reason is that Hume wants to treat the cases involving coherence as wholly different from the cases involving constancy. Hume utilizes coherence and constancy to explain the identity conditions of different stages of different kinds of

objects as well as to explain their numerical identity. In Hume's view, constancy accounts for the identity conditions of the organic objects, i.e., the vegetables and animals, which undergo *gradual* or *total* changes but maintain constancy during that period of time. For example, we attribute identity to, say, "an oak, that grows from a small plant to a large tree...tho' there be not one particle of matter, or figure of its parts the same. An infant becomes a man, and is sometimes fat, sometimes lean, without any change in his identity" (THN, 257). On the other hand, coherence seems to be a better fit to account for the cases of numerical identity. For instance, when we hear "a noise, that is frequently interrupted and renew'd" we say that it is the same noise, "tho' 'tis evident the sounds have only a specific identity or resemblance, and there is nothing numerically the same, but the cause, which produc'd them" (THN, 258). Or, when a church, which was rebuilt entirely anew, where there is nothing in common between the old and the new churches in terms of their forms, sizes, and materials, coherence is, Hume argues, "sufficient to make us denominate them the same" (THN, 258). Apparently, since there is no constancy involved in cases of these kinds, Hume needs another principle to explain our belief in the identity of the cases that the first object is entirely annihilated before the second comes into existence. Thus, coherence is not only a separate and distinct principle, but also an important and indispensable part of Hume's explanatory framework.

Hence, I reject the unified hypothesis and argue that coherence and constancy are indispensable part of Hume's psychology of identity ascriptions, and that neither is a special case of the other. That is why Hume has provided two separate explanations for the series of perceptions exhibiting constancy and those exhibiting coherence. Similarly, attributing the dissatisfaction hypothesis to Hume faces the same problem. I will argue that the text of the *Treatise* shows that Hume is not forced into incorporating the constancy-based explanation as a special case of his coherence-based explanation; rather, he consciously utilizes both explanations as parts of his psychological belief-determining mechanism. In what follows, I shall first characterise Hume's account of the constancy-based explanation and then account for his coherence-based explanation.

According to Hume, when we have a series of closely similar impressions, with a gap or interval in between, in which all the impressions are alike or similar, and resemble each other closely, the mind confounds the resembling impressions and, as a result of

that, *identifies* the two numerically different impressions as *individually the same*, upon account of their resemblance (THN, 199). I use formal rendering to explain this.

When we have an *un*interrupted series of impressions of an object, the perceptual experience is like the following:

(1) AAA

And the interrupted series of our perceptual experiences will look like this:

(2) ABA

Yet our perceptual experiences in our ordinary lives will be much more like this:

(3) ABCDEA ⁽⁸⁾

In situation (1) there is no interruption and, thus, no constancy involved there. The situation is like gazing at an object without blinking, but in situation (2) we deal with interrupted impressions and constancy is involved to generate the belief in the identity of objects. More specifically, in situation (2) we suppose the A-perceptions on either side of the B-perceptions to be individually the same, although there is an obvious interruption between them. Think of fixing your eyes on a red car driving in front of you down the highway while storing thousands of perceptions in your memory repertoire, or as Hume said, in her bundle of perceptions. Definitely, we are dealing with an interrupted series of impressions, which exhibits constancy in a uniform and consistent manner. We are inclined to ascribe a perfect identity to the A-perceptions, despite the interruption. Obviously Hume thinks that in situation (2) there is constancy, despite the existence of obvious interruption and that the source of the inclination is our tendency to confuse or confound situation (2) with situation (1), where there is no interruption. Indeed, in Hume's view it is a fundamental fact about the mind that "whatever ideas place the mind in the same disposition or in similar ones, are very apt to be confounded" (THN, 203). As a result, "an interrupted series of exactly similar impressions places the mind in almost the same disposition as an uninterrupted series of exactly similar impressions" ⁽⁹⁾. The transition in situation (2) is "almost the same disposition of mind" as that in situation (1), where we are mistaken to think that the two kinds of series have exactly the same effects on us and thus it is very *natural* for us to confound them. Therefore, at the heart of this confusion or confounding there are two distinct but resembling perceptual situations where we wrongly take an interrupted series of resembling impressions to be exactly the

same as of an *un*interrupted series and, therefore, to be a “continu’d view of the same object” or a case of ‘perfect identity.’

In situation (3), we deal with another sort of interrupted series of impressions, though different from (2) in certain ways. We can think of this situation as of an oak tree undergoing changes and constantly getting bigger and bigger. Unlike (2) where the object was invariable and did not change over the course of one having the series of impressions, the object is variable and changing, and that the impressions that one stores in its memory repertoire exhibits difference in terms of contents and constancy, though some resemblance holds between them. In other words, the main difference between (2) and (3) is that the former exhibits constancy and the latter shows coherence in one’s experience. Also, (3) explains those kinds of situations that Hume has in mind when he talks about the porter case (THN, 196). This, I will explain later. In short, for Hume, there are only two categories of series of perceptions: (a) uninterrupted series of perceptions, and (b) interrupted series of perceptions. The latter breaks up into those series that represent the variable objects and those others that represent the invariable objects. The (a) series exhibit constancy and (b) series exhibit coherence. Furthermore, constancy and coherence are certain features or characteristics of the series of impressions, which through affecting the imagination, produces the fiction of continuity, or identity, in objects and selves.

To further elaborate, for Hume, a series of perceptions is constant if and only if they “present themselves in the same uniform manner, and change not upon account of any interruption in [one’s] seeing or perceiving them” (THN, 195). Let’s suppose that A1 is my perception at t_1 gazing at a red car driving down the highway, and suppose that A3 is my perception of the same car at t_3 . In the interval between A1 and A3, that is, at t_2 , “I lose sight of [the car] by shutting my eyes or turning my head” (THN, 194), but still the series of perceptions in this case “present themselves in the same uniform manner.” This situation may explain the paradigmatic case of constancy of the series of perceptions. We can identify three main features for the series of perceptions involving constancy: (1) the series, A1...A3, is interrupted, but the object of our perceptions, the car, or as Hume argues, mountains and houses, are invariable and unchanging; (2) the perceptions in the series are qualitatively identical, but numerically distinct; and (3) perceptions in the series

are *strictly*, resembling. These features constitute the constancy of the series of perception. Nevertheless we should mitigate Hume's notion of *exact*, or strict, resemblance of the series of perceptions into *close* resemblance ⁽¹⁰⁾. I have found this kind of mitigation revealing because Hume seems to be relying on a notion of close resemblance when he explains different cases of constancy (more on this in III.3).

Now, let's look at a different situation. Suppose, A is my perception of the fire at the fireplace at t_1 , and C is my perception at t_3 of the same fire. However, in the interval between A and C, that is, at t_2 , I have been absent from the chamber. Here, the series of perceptions, i.e., A...C, lacks constancy because they do not present themselves in the same uniform manner as in A1...A3. Hume, however, claims that the series of perceptions in this situation exhibits a different form of regularity, namely, coherence. Because as he claims, "When I return to my chamber after an hour's absence, I find not my fire in the same situation, in which I left it" (THN, 195). Instead, because "I am accustom'd in other instances to see a like alteration produc'd in a like time, whether I am present or absent, near or remote" (THN, 195), therefore, I conclude that A coheres with C. More precisely, because I have, on other occasions, observed the alteration of the fire while maintaining a sort of uniformity exhibiting certain features of constancy, I tend to believe that the series is coherent. Accordingly, since I have observed, in other instances, the fire in the chamber while being present there and observing the alteration going uniformly from point A' to B', and then to C', I come to believe that the series of perceptions that I have from the fire in my chamber while I am absent, i.e., from A...C, still present a coherent experience of external objects. The fire example is Hume's paradigmatic case of coherence-based explanation of our belief in identity of objects. Other examples of Hume's coherence-based explanation include the famous porter-letter case, which is slightly different from the fire case. I will explain it in the following section. Similarly, we can identify the three following features for the series involving coherence: (1) the series, that is, A...C, is interrupted, and the object of our perception, the fire, is variable and changing; (2) the perceptions in the series are both *qualitatively* and *numerically* distinct; but (3) perceptions in the series somehow exhibit *remote resemblance*. In short, these features show that the series of perceptions in this situation are not constant. These characterizations, assuming that they hold true, indicate that

constancy is essentially different from coherence.

One fundamental difference between constancy and coherence is that the former consists in a relationship between two individual perceptions, whereas coherence is a relation “between series of impressions; namely, between a fragmentary series now observed and a number of continuous series observed in the past” (11). More precisely, the main difference between constancy and coherence is this: “In the case of Constancy, the original continuous series is a *monotonous* one; it is of the form $A_1 A_2 A_3 A_4 A_5$, where all the items resemble each other very closely. Whereas in the case of Coherence, the original continuous series is of a *variegated* sort; it is of the form $ABCDE$, where the items differ from each other in their qualities or in their spatial relations or in both” (12). For Price, however, this insignificant difference does not amount to a fundamental distinction between constancy and coherence. In fact, what is in common between these two kinds of series of perceptions is that both exhibit, Price argues, a similar characteristic, called, *Gap-indifference*, that is, indifference to the occurrence of gaps (13). He claims that even in the case of constancy we deal with certain series of perceptions, rather than with individual perceptions. Needless to say, after reducing constancy to coherence “the imaginative procedure is essentially the same; it is a passage from an observed partial resemblance to a postulated complete one” (14). When the resemblance is between *series* of impressions, i.e., between a broken series and a complete one, the resemblance between individual perceptions within the same series becomes irrelevant. Hence, Price concludes that constancy and coherence are the “sub-species of a common principle” (15).

I suggest that coherence differs from constancy in respect to the fact that constancy is a relationship between, or a property of, two individual perceptions, whereas coherence is a relation between, or a property of, a *set* of series of inconstant perceptions that cohere with another set of series of inconstant perceptions (16). I think that this distinction is fundamental for understanding the difference between constancy and coherence. It seems that constancy is strictly a relation between two individual perceptions; it does not involve two sets (or more) of series of perceptions. As Price has recognized, the series involving constancy is a *monotonous* one, entailing that individual perceptions within the series resemble each other exactly or closely. Whereas, for Hume,

the resemblance involved in coherence, which Price has characterized as a *variegated* one, is not between individual perceptions within a series but between two sets of series of perceptions. In order to understand the difference between constancy and coherence, it is important to make a distinction between two types of gaps in our perceptual experiences, namely, temporal and spatial gaps.

Hume's mental world is essentially an interrupted, fragmented, episodic, and ever-changing entity. "He is impressed," Price claims, "as no philosopher before him had been, by the *interrupted* and *fragmentary* character of human sense-experience" (17). Basically, every simple blink of the eyes, or shutting them, or falling asleep, or turning the head, or touching this object or that one, hearing a door turning on its hinges or someone walking up the stairs, etc., is enough to cause gaps and holes in our sense-experiences. Hence, for Hume, there is no uninterrupted or invariable series of perceptions; all our perceptions are, at least, *temporally* interrupted. Even, in those objects, or any mass of matter, where its parts are contiguous and connected and continue uninterrupted and invariable, which Hume is inclined to think about them as cases of perfect identity, we can hardly come up with any explanation to justify their perfect identity because they are temporally interrupted. Furthermore, Hume argues that since the spatial characteristics, viz., shape, size, pattern, etc., of some objects alter, whether gradual or small or total changes, it implies that there is *spatial* gaps in these objects and, thus, strictly speaking, they are not continuous. An oak, for example, grows from a small plant to a large tree, an infant becomes a man, who sometimes is fat and sometimes lean (THN, 257). Or, in more extreme cases, like a church that was formerly of brick and fell to ruin and that is rebuilt of free-stone, the first object is completely annihilated before the second comes into existence (THN, 258). So, for Hume, there is no continuity in objects because either they undergo spatial alterations or they are subject to temporal gaps. In other words, he seems to be thinking that there is no 'perfect identity'; "[t]here is properly no *simplicity* in it at one time, nor *identity* in different" (THN, 253). However, these gaps and holes, which certainly interrupt the continuity of our sense-experiences, are subject to, Hume argues, some form of continuity. Despite the fact that our sense-experiences fail to show 'perfect identity,' certain "natural propensity we may have to

imagine that simplicity or identity” (THN, 253). This identity or simplicity is a ‘fictional’ one.

Having this distinction in place, we may say that when we talk about the continuity of external objects, we, indeed, deal with two different kinds of objects: (1) the invariable objects that continue to remain the same in terms of its spatial features, i.e., size, shape, etc.; (2) the variable objects whose spatial features continuously alter. Even though we observe spatial continuity in the first category of objects, such as mountains, houses, and trees (I will explain in a moment why Hume includes trees here), insofar as the mass of matter that makes these objects remain the same (because the mass of matter remains connected, contiguous and unaltered), they are said to be interrupted temporally. Therefore, when we are having a series of qualitatively identical (exactly resembling) perceptions of an object, like a mountain, in the form of, say, A1...A3, with a temporal gap in between, we are forced, by philosophical reflection, to identify the series as temporally discontinuous. Whereas, when we observe an object that continuously alters, like a plant that grows to a large tree, in the form of, say, ABCDE, we deal with a situation in which our perceptual experiences are not only interrupted temporally, but also, and more fundamentally, the *content* of our perceptual experience at A is entirely different from our perceptual experience at B and C, and D, and E, though exhibiting some partial resemblances. This is to say, our perceptual experience of the changing objects is interrupted in terms of the contents of those impressions stored in the mind. In fact, spatial gaps in the changing objects get translated into content gaps in the mind. The key point here is that these objects affect the mind, or the imagination, differently, and as a result of that we have two different kinds of series of impressions, each exhibiting different characteristics: the first series exhibits constancy and the second one exhibits coherence.

As pointed out above, the main difference between the series involving constancy and the series involving coherence is that the former is strictly a relationship between individual perceptions while the latter involves a set of (or more) series of perceptions. The rationale for such a claim is that in the series involving constancy, A1...A3, which is strictly a series among individual perceptions, causal inference are not involved, while in the case of the series involving coherence, ABCDE, causal inference plays a central role.

For Hume, constancy itself, which is a property of a series of perceptions, consists in those perceptions' "being qualitatively identical, or at least similar, to each other" (THN, 194-5). And it is because of the effects of constancy that we come to identify the several resembling impressions in a series to be one continued impression existing during the intervals (periods) of time when it was not before the mind. The mind posits, or postulates, an unperceived perception, *unsensed sensibillum*, in the interval to smooth the passage of the imagination from A1 to A3. Firstly, in this process, there is only one series of impressions involved, and secondly, the relation between each individual perception in the series is not constituted by causal inference. On the contrary, in the case of the series involving coherence, the coherence itself is constituted by the existence of causal inference between the "standard series" of impressions and the interrupted one. A standard series is a sequential series of impressions without any gaps in between, whereas the interrupted series is a sequential series which gaps occur in. To use Hume's example, we may have the following standard series of impressions: (A) a visual impression of a door turning upon its hinges, (B) an auditory impression of a noise, (C) a visual impression of a porter walking up the stairs, (D) the sound of footsteps, and (E) the visual impression of the porter placing an envelope on my desk. So, the standard series in this case runs like ABCDE. The series of impressions in real life may turn out to be in a variety of forms, however. It may turn out that we have the following, for instance, series: (a) A...CDE, (b) ...BCDE, (c) A.....E, etc. Basically, the idea is that these interrupted series somehow maintain coherence because each one bears some partial resemblance to the standard series. For example, in (b) we have an auditory impression of a squeak, and we come to believe in the existence of a door turning upon its hinges. It seems that here we make an inference from hearing the noise of a squeak to the existence of a door opening. This looks like, indeed, a case of causal inference; the hearing of the squeaking sound is the effect of the door turning upon its hinges. Obviously, here we have an auditory impression of the squeaking sound while we have no perception of the door turning upon its hinges. Indeed, the cause of the squeaking sound is entirely absent from series of impressions involved here. Nevertheless, in order to make the inference we need to reproduce the visual impression of the door turning upon its hinges from some previous series, namely, from the standard series. More precisely, in (b) we have an

individual auditory impression of a squeak, but we do not have in this particular series the other impression, namely, seeing a door in motion. Therefore, we have only one individual impression in this series, and it is not enough to render the whole series coherent. As a result, the imagination looks for another individual impression present in the mind from its previous resembling series of impressions to fill the gap. Hume concludes that the standard series provides such a resembling impression. In a nutshell, in the case of coherence two (or more) sets of series of perceptions are involved while in the case of constancy two individual perceptions within a series of impressions is enough to render the series constant. Thus, the coherence of any series of impressions consists in its causal relation with some other previous series. For Hume, to talk about coherence within a single series of impressions is accordingly fruitless.

There is a rival interpretation that I must consider now. This interpretation has two main sources: Firstly, it stems from the so-called unified explanation hypothesis, and secondly, it is reinforced by a tendency that disregards Hume's discussion of coherence and focuses instead on his discussion of constancy (18). As a result, the advocates of this approach claim that constancy is a special case of coherence. To do so, they need to be able to reduce the two sets (or more) of the series of perceptions into *one* single series of perceptions. Tim Black has characterized this interpretation as it follows:

“P and Q are coherent if and only if they are less than exactly similar but nevertheless exhibit a sufficient degree of similarity” (19).

Perhaps this characterization of coherence may be applied to cases like fire in which we observe an interrupted series of impressions that exhibit some resemblance or similarity. As explained above, when Hume left his room he had perception A of the fire in the fireplace, and when he returned he had another perception C of the same fire. In the interval between A and C, he was absent from the chamber and had no perception of the fire. The series of perceptions, i.e., A...C, in this situation lacks constancy because they do not present themselves in the same uniform manner as in A1...A3. However, there is still a remote resemblance between A and C in the sense that they are both perceptions of the same fire at two different stages of the fire. There is a slight possibility that we can treat the case of fire as a case of constancy, and then claim constancy as a special case of coherence. However, that would be a mistaken view of Hume's position. I will explain

below that the case of fire is essentially different from the cases of constancy, and therefore, we cannot subsume the explanation of the belief in objects as it arise from constancy under the explanation of the belief in objects as it arises from coherence.

This interpretation encounters two major difficulties. First, immediately after his discussion of the case of fire, Hume points out that even though when he returns to his chamber after an hour's absence, he does not find the fire in the same situation, in which he left it, "I am accustom'd in other instances to see a like alteration produc'd in a like time, whether I am present or absent, near or remote" (THN, 195). This passage clearly indicates that in order to render the series coherent we need to observe the alteration of the fire on previous occasions. Therefore, the existence of two (or more) sets series of perceptions is essential to have a coherent series in this situation. More fundamentally, the porter-letter case, as I explained above, shows that it is impossible to have a coherent picture of the series without having observed similar series in other instances.

Undoubtedly, the above characterization implies that there should be, at least, two individual, though remotely resembling, perceptions in the series to render the series coherent. Nevertheless, in the porter-letter case we only receive one individual perception, for example, the noise of a door turning upon its hinges while missing the other perception, i.e., seeing a door in motion. Or, we hear the sound of footsteps without seeing the person walking up the stairs. Thus, in cases like these, there is only one individual perception, whereas the rival interpretation requires, as the above characterization shows, at least, two individual perceptions within a series in order to render it coherent. Therefore, for these two reasons, the rival interpretation has to fail.

There is one last point that I shall consider briefly. Hume argues that although the causal inference involved in coherence "may seem to be of the same nature with our reasonings concerning causes and effects; as being deriv'd from custom, and regulated by past experience; we shall find upon examination, that they are at the bottom considerably different from each other, and that this inference arises from the understanding, and from custom in an indirect and oblique manner" (THN, 197). The reason, Hume argues, is that "whenever we infer the continu'd existence of the objects of sense from their coherence, and the frequency of their union, 'tis in order to bestow on the objects a greater regularity than what is observ'd in our mere perception. We remark a connexion betwixt two kinds

of objects in their past appearance to the senses, but are not able to observe this connexion to be perfectly *constant*, since the turning about of our head, or the shutting of our eyes is able to break it. What then do we suppose in this case, but that these objects still continue their usual connexion, notwithstanding their apparent interruption” (THN, 197-8: emphasis added). In short, in the case of usual connexion between cause and effect, the relation among them, is said to be, of a constant conjunction, but in the case of ‘coherence inference’ the connexion, or conjunction, is not perfectly constant. Here, we come to posit, or postulate, an unperceived perception, *unsensed sensibilium*, to render the interrupted series coherent. Some commentators have complained that Hume has failed to distinguish coherence inferences from causal inferences (20). No doubt, Hume’s distinction is puzzling here. Moreover, it is evident that ‘custom’ or ‘habit’ plays a central role in both coherence inferences and causal inferences. But Hume distinguishes the case of causal inference as based on constant conjunction and the case of coherence inference that lacks constancy. Hume thinks that this is a very important distinction.

The current interpretation can settle a passage that has “perplexed commentators” (21). In the transitional paragraph from coherence to constancy, Hume claims that “whatever force we may ascribe to this principle [coherence], I am afraid ’tis too weak to support alone so vast an edifice, as is that of the continu’d existence of all external bodies; and that we must join the *constancy* of their appearance to the *coherence*, in order to give a satisfactory account of that opinion” (THN, 198-199). Stroud, for instance, has read this passage as Hume’s rejection of the coherence-based explanation and thus, concluded that the constancy-based explanation is adequate to account for our belief in the continued existence of bodies (22). I suggest that we may read this passage not as a rejection of the coherence-based explanation, but simply as finding it inadequate, or as Hume himself says, “too weak to support alone so vast an edifice.” If Hume had found it flawed, or something deeply wrong about it, he would have had mentioned it, as he did it in the case of personal identity. There is no textual evidence, however, indicating that Hume was unsatisfied with his coherence-base explanation. Interestingly, in the concluding paragraph of the section of “Of Scepticism with Regard to the Senses,” once again, he emphasizes the roles of both constancy and coherence in causing our belief in the continued existence of bodies. There, he argues, “they are the coherence and

constancy of our perceptions, which produce the opinion of their continu'd existence" (THN, 217). Even more interestingly, despite the fact that Hume immediately says that "the constancy of perceptions has the most considerable effect," he claims that "yet [it] is attended with the greatest difficulties" (THN, 217). Nevertheless, he never talks about any difficulties involved in his coherence-based-explanation. Perhaps one way to read this passage is to say that since Hume needs two different explanations for invariable and variable objects, it follows that constancy and coherence should be treated differently, and that Hume does not treat constancy as a special case of coherence. Therefore, constancy is a main component of Hume's psychological belief-determining mechanism and we cannot do away with it.

Modifying Constancy

For Hume, constancy consists in *exact* resemblance between individual perceptions in a series of perceptions. Hume explicitly expresses this idea in the following passage. Since I can hardly improve on the text, I quote the passage in its entirety:

We find by experience, that there is such a *constancy* in almost all the impressions of the senses, that their interruption produces no alteration on , and hinders them not from returning the same in appearance and in situation as at their first existence. I survey the furniture of my chamber; I shut my eyes, and afterwards open them; and find the new perceptions to resemble perfectly those, which formerly struck my senses. This resemblance is observ'd in a thousand instances, and naturally connects together our ideas of these interrupted perceptions by the strongest relation, and conveys the mind with an easy transition from one to another. An easy transition or passage of the imagination, along the ideas of these different and interrupted perceptions, is almost the same disposition of mind with that in which we consider one constant and uninterrupted perception. 'Tis therefore very natural for us to mistake the one for the other. (THN, 204)

According to this passage, first, the objects that Hume surveys—the furniture of his chamber—remain unchanged and invariable. Second, the content of perceptions after interruption perfectly resembles the content of perceptions before interruption even if observed a thousand times. That is to say, the perceptions in the series are qualitatively identical but numerically distinct. Last, there is strict or exact resemblance among perceptions in the series. This, we may call, Hume's standard account of constancy. Hume, however, discusses other examples and situations in order to explain the property of constancy. These new examples and situation slightly diverge from his standard

account. I shall discuss them now.

In his early characterization of constancy in the *Treatise*, Hume writes:

After a little examination, we shall find, that all those objects, to which we attribute a continu'd existence, have a peculiar *constancy*, which distinguishes them from the impressions, whose existence depends upon our perception. Those mountains, and houses, and trees, which lie at present under my eye, have always appear'd to me in the same order; and when I lose sight of them by shutting my eyes or turning my head, I soon after find them return upon me without the least alteration. My bed and table, my books and papers, present themselves in the same uniform manner, and change not upon any account of interruption in my seeing or perceiving them. (THN, 194-5)

This passage contains almost all the examples that Hume has in mind for the cases of constancy. Here, Hume talks about objects such as mountains, houses, table, books, and papers, showing similarity to the previous example (the furniture and the chamber) insofar as they are all invariable objects. Nonetheless, there is one troubling example here. Hume includes *trees* in the list alongside those of mountains, houses, table, and books and papers, which lie in his sight and always appeared to him in the same order. A tree, unlike other objects mentioned in this passage, is a variable object undergoing gradual changes; it is very much the same as Hume's fire example, the paradigmatic case of coherence. In the section of "Of Personal Identity," Hume argues that "an oak tree, that grows from a small plant to a large tree" endures total change, though we still attribute identity to it while its form, size, and substance—spatial characteristics—is entirely changed (THN, 257). If we take the example of an oak tree as a case of coherence, as this passage may indicate, then the inclusion of the tree example alongside objects like mountains and houses, as the clear examples of constancy, causes an inconsistency in Hume's position. At first sight, it may not be clear why Hume has included objects like trees in the list that is mainly about the cases of constancy. Perhaps Hume thinks of the tree example in this context as a case of constancy because the change in the object is so minimal that can hardly affect the similarity of resembling perceptions in the series. In other words, the perceptions, as Price has said (23), in the series *closely* resemble each others. The content of perception A₁ is so similar to the contents of perceptions of A₂, A₃, A₄, A₅ of the same tree at different times. So it seems that the right interpretation of Hume's notion of constancy is to say that it consists in *close* resemblance rather than *strict* or exact resemblance. There are other examples that may support this construal.

In the paragraph in which Hume gives a “short sketch” of his system concerning constancy, he claims that “when we have been accustom’d to observe a constancy in certain impressions, and have found, that the perception of the sun or ocean, for instance, returns upon us after an absence or annihilation with like parts and in a like order, as at its first appearance, we are not apt to regard these interrupted perceptions as different, (which they really are) but on the contrary consider them as individually the same, upon account of their resemblance” (THN, 199). Again, it might be said that although certain spatial features of the sun or ocean, for instance, their locations or shapes (as they appear to our perceptual system; the sun changes its location in the sky and an ocean constantly changes its shapes), endure changes, the occurrence of these kind of changes can hardly affect their order of appearances to the mind overtime. When we fix our gaze at the sun in the sky at t_1 and again look at it an hour later at t_3 , we are having an interrupted series of perceptions that exhibits constancy. It is similar to the series of impressions that we have of a tree at three different instances. Hence, the series of perceptions involved in the cases of the sun and ocean is similar to the series involved in the tree example in the sense that there are small or inconsiderable changes in the objects overtime, such that the effects of these gradual changes are so trivial to our perceptual system that they can hardly render the series involved inconstant.

Furthermore, in the section of “Of Personal Identity” where Hume discusses the identity of different objects, he discusses cases that are similar to the tree example in the above passage. According to him, any change that is produced *gradually* and *insensibly*, though, strictly speaking, destroys the identity of the object, we still ascribe an identity to the objects. “Supposing some very *small* or *inconsiderable* part to be added to the mass, or subtracted from it,” Hume claims, “tho’ this absolutely destroys the identity of the whole, strictly speaking; yet as we seldom think so accurately, we scruple not to pronounce a mass of matter the same, where we find so trivial an alteration” (THN, 255-56). More specifically, he argues, for instance, that “the addition or diminution of a mountain wou’d not be sufficient to produce a diversity in a planet; tho’ the change of a very few inches wou’d be able to destroy the identity of some bodies” (THN, 256). The “reason can plainly be no other, than that the mind, in following the successive changes of body, feels an easy passage from the surveying its condition in one moment to the

viewing of it in another, and at no particular time perceives any interruption in its actions. From which continu'd perception, it ascribes a continu'd existence and identity to the object" (THN, 256). Therefore, it seems that the occurrence of some minor changes in the objects may not affect the order of their appearances to the mind—that is, even close resemblance within a series of perceptions is enough to render the series constant. In fact, "the passage of the thought from the object before the change to the object after it, is so smooth and easy, that we scarce perceive the transition, and are apt to imagine, that 'tis nothing but a continu'd survey of the same object" (THN, 256). In a nutshell, since there is not a significant difference between the series of perceptions involving objects like mountains, houses, books, etc., and the series of perceptions involving objects like trees, the sun and ocean, it follows that all these objects could be considered cases of constancy.

Reading the passage this way can pose a problem for my interpretation of Hume's principle of constancy, however. As I stated earlier, constancy involves *strict*, or *exact*, resemblance between perceptions within a series of perceptions. There, I claimed that the perceptions in the series involved constancy have the same or similar *contents*. The implication of such claim was that constancy is specified to account for our belief in the continuity and distinct existence of invariable and unchanging objects. Nonetheless, if we hold that the resemblance among perceptions in a series of perceptions involved constancy consists in *close* resemblance, rather than *exact* resemblance, it readily follows that we are no longer able to say that constancy is specified to account for invariable objects alone. In fact, as the above passage may suggest, constancy can account for our belief in the continuity and distinct existence of both variable and invariable objects. More fundamentally, this prepares the ground for the claim that constancy is a special case of coherence. Coherence requires some kind of resemblance between perceptions—close or remote resemblance—in a series (or sets of series) of perceptions. A perception of a fire at t_1 is *remotely* similar to another perception of the same fire at t_3 , and that the same series of perceptions may *closely* resemble another preceding series of perceptions. For instance, there is a remote resemblance between a four-log fire burning at a rate of one log per hour and the same fire two hours later still burning at two-log rate per hour (24). But the whole process and the series of perceptions may closely resemble

another previous series of perception of a fire burning at the same rate and at the same fireplace. In short, it seems that there is no much of a difference between Hume's tree example and his fire example in this regard. Consequently, one could possibly argue that constancy is a special case of coherence.

Now, if we consider the conclusion that is drawn in the last paragraph and the one before it, we may clearly discern a conflict in Hume's position. On the one hand, for Hume, the paradigmatic case of constancy is the one that exhibits strict or exact resemblance among perceptions of a series, and on the other hand, he seems to subsume the series that exhibit close resemblance under the series of perceptions that exhibit strict or exact resemblance. In other words, Hume explicitly subsumes, for instance, the tree example under the category of constancy, and seems to be holding that there is no substantial difference between the fire and the tree examples. I want to resist this argument. I am inclined to think that Hume makes a distinction between constancy and coherence, and that he is not prepared to render constancy as the special case of coherence—that is, he is not prepared to subsume the explanation of the belief as it arises from constancy under the explanation of the belief as it arises from coherence. In principle, it remains true that “the propensity to ascribe identity to related objects, which is essential to Hume's explanation of the belief in body in connection with constancy, plays no role in his explanation of the belief in connection with coherence” (25). Moreover, for Hume, the belief as it arises from coherence involves “a kind of reasoning from causation” (THN, 195). It involves co-operation between custom and the imagination in which when the mind is set “into a train of thinking, is apt to continue, even when its objects fails it, and like a galley put in motion by the oars, carries on its course without any new impulse” (THN, 198). As Loeb points out, this “psychological galley plays no role in Hume's account of the belief as it arises from constancy” (26). Furthermore, as I have shown in previous section, constancy involves only *a* temporally interrupted series of impressions, whereas coherence involves, at least, a *set* of series (or more) of perceptions. This is why causal inference is so fundamental to coherence, but plays no role in constancy.

Given this, it seems that, for Hume, the objects like trees, infants, etc. (almost all organic objects) could be categorized as either the cases of constancy or as the cases of

coherence. It depends how they affect our perceptual system. I should explain this now. According to Hume, if change is produced gradually and insensibly while maintaining a sort of *proportion* to the whole, we are apt to suppose that the object continues the same over time. For instance, the addition or diminution of a mountain to the planet is not sufficient to produce diversity in it. By the same token, the addition or diminution of a few inches to an oak tree, or an infant, does not produce diversity in them, though strictly speaking it completely destroys their identities. The key point is that if the addition or diminution occurs in a uniform manner, that is, if it happens gradually and insensibly, its effects upon the mind after the change are almost the same as its effects on the mind before the change. Thus, an oak tree that I am observing now in my backyard is almost the same as the one that I had observed two days ago, or two weeks ago, or even two years ago. Or, if I observe a ten-year-old child now, and two days later, or two weeks later, or even two years later, I still believe that I am observing the same child. The reason is that the successive changes of the bodies have similar effects on the mind in such a way that the mind feels it is observing the exact same object. In a more fundamental sense, even though constant changes in an object like an oak tree make it cease being strictly the same (perfect identity), we may say that it is the uninterrupted progress of the thought, which constitutes the continuity (imperfect identity) of bodies overtime (THN, 256). In short, there are two elements that are essential for ascribing identity to these kinds of objects: (a) the gradual or insensible changes of objects over time, (b) the uniformity (constancy) of our perceptual experiences of them. Therefore, the continuity that we ascribe to an oak tree in this context is due to the constancy of the series of impressions involved there.

This kind of tendency is most remarkably observed in Hume's noise and river examples. Hume brings up the example of an interrupted noise where we can:

distinguish pretty exactly betwixt numerical and specific identity, yet it sometimes happens, that we confound them, and in our thinking and reasoning employ the one for the other. A man, who hears a noise, that is frequently interrupted and renew'd, says, it is still the same noise; tho' 'tis evident the sounds have only a specific identity or resemblance, and there is nothing numerically the same, but the cause, which produc'd them. (THN, 258)

Clearly, in this case there is no exact resemblance between individual perceptions in the series because the contents of perceptions are distinct (numerical difference), but there is

close resemblance among the contents of perceptions in the series (specific identity). Here, even the vulgar knows that the series is interrupted, and that each noise, or perception for that matter, is distinct from each others, but due to close resemblance she is apt to confound them, and as a result of that to ascribe identity to an interrupted series of numerically distinct but qualitatively identical noises. The case of river is even more lucid than this one. In talking about the identity of a river Hume states:

We may remark, that tho' in a succession of related objects it be in a manner requisite, that the change of parts be not sudden or entire, in order to preserve the identity, yet where the objects are in their nature changeable and inconstant, we admit of a more sudden transition, than wou'd otherwise be consistent with that relation. Thus as the nature of a river consists in the motion and change of parts; tho' in less than four and twenty hours these be totally alter'd; this hinders not the river from continuing the same during several ages. (THN, 258)

Hume unambiguously remarks here that in order to render a series of perceptions constant it is a requisite that the change of parts not be sudden or entire. Even in this case that the change of parts in the river is sudden and entire—in less than twenty four hours all the parts and particles of the body of water constituting the river totally alter— we still observe a sort of constancy in the river. Again, the reason may well be that there is a close resemblance between the perception that I had yesterday, and the one that I have now, and the one that I will have tomorrow. It is this close resemblance among the series of perceptions that cause us to ascribe continuity to the river during several ages.

Now, what if the changes are considerable and inconstant? I observe a small plant in my backyard and then twenty years later see a large oak tree there. How are we supposed to account for the continuity of the oak tree in this new situation? Hume claims that "'tis certain, that where the changes are at least observ'd to become considerable, we make a scruple of ascribing identity to such different objects" (THN, 257). However, Hume, in the immediately following paragraph, claims that "tho' every one must allow, that in a very few years both vegetables and animals endure a *total* change, yet we still attribute identity to them, while their form, size, and substance are entirely alter'd" (THN, 257). He goes on to say, "an oak, that grows from a small plant to a large tree, is still the same oak; tho' there be not one particle of matter, or figure of its parts the same. An infant becomes a man, and is sometimes fat, sometimes lean, without any change in his identity" (THN, 257). It seems that Hume talks about the oak tree, or the infant, in an

entirely new context. I suggest that in this new context he talks about the tree and the infant as cases of coherence rather than constancy, which I explained in the preceding paragraph. Let's say that someone observed a small plant in her backyard and twenty years later she observes a large oak tree in its place. Obviously, the change that occurred is a total one—that is, the form, the size, and substance are entirely altered. It is important to note that the observer does not have a constant series of perceptions in this situation. In fact, the individual perceptions within the series involved in this case are so different that they hardly exhibit resemblance. There is no exact resemblance between the perception of a small plant and the perception of a large oak tree. And since there is no exact or close resemblance among these individual perceptions, it follows that the series of perceptions involved here lack constancy. However, we still ascribe identity to the oak tree even if we lack a constant series of perceptions. In Hume's view, the source of our ascribing identity to the oak tree, despite its entire alteration from a small plant to a large tree, is because of the coherence in the series of perceptions. As I have explained above, in order to render the series coherent, it is necessary that the observer has previous experiences of a small plant growing to become a large oak tree. Indeed, without having another resembling series of perceptions the mind is unable to trigger coherence in the latter series of perception.

After discussing the tree and the infant examples, Hume introduces another case, which clearly shows that when he talks about these examples he is, indeed, thinking of them as cases of coherence. In the immediately following paragraph, Hume talks about the identity (continuity) of a church that is rebuilt entirely anew. There he writes:

In like manner it may be said without breach of the propriety of language, that such a church, which was formerly of brick, fell to ruin, and that the parish rebuilt the same church of free-stone, and according to modern architecture. Here neither the form nor materials are the same, nor is there any thing common to the two objects, but their relation to the same inhabitants of the parish; and yet this alone is sufficient to make us denominate them the same. (THN, 258)

Again, in this case the form, size, and substance of the object are entirely altered, and we still ascribe continuity to the object. Certainly, the ascription of continuity (identity) to the object is not due to the constancy of the series of perceptions involved here.

Constancy requires exact or, at least, close resemblance among individual perceptions in a series. In this case, however, we are aware of the fact that the object is entirely

annihilated before the second one comes into existence. In fact, we know that, strictly speaking, the second object is not the same as the first one. Therefore, it seems that there is no exact or close resemblance between the series of perceptions in the series involved here. As Hume claims, “we are never presented in any one point of time with the idea of difference and multiplicity; and for that reason are less scrupulous in calling them the same” (THN, 258). That is to say, in this case the interruption is so massive that we are not even apt to suppose that the old church and the new one are connected together by resemblance, and thus, continue to be the same. In short, it is highly unlikely that Hume would think of this example as a case of constancy. The reason may well be that when an object is entirely annihilated before the second one comes into existence; strictly speaking, there is no exact or close resemblance between individual perceptions in the series. The reason that we still ascribe identity to the object must be due to the existence of a relation between a previous series of perceptions and the new one. Apparently, if we had no previous experience of observing a church that is demolished and then rebuilt, we would never have come to the conclusion that there is continuity (identity) in it. Hence, the existence of a set (or more) of series of perceptions is necessary to give rise to the belief in the continuity of the church here.

Considering Two Objections

I have argued above that the so-called unified explanation hypothesis, according to which we could deal with constancy as a special case of coherence, should be rejected because Hume needs two different sorts of explanations in order to account for the continuity of different objects. I claimed that Hume is satisfied with both of his coherence-based explanation and constancy-based explanation. On the other hand, I have suggested that when we account for the continuity of certain objects, say, an oak tree, we can explain the belief that arises there as either a case of constancy or as a case of coherence. And the fact that the belief in the continuity of an oak tree could be explained as either a case of constancy or as a case of coherence may suggest that we, indeed, subsume Hume’s explanation of the belief in the continuity of the oak tree as it arises from constancy under his explanation of the belief in the continuity of the same tree as it arises from coherence. Again, this may suggest that constancy is a special case of coherence, and that the unified explanation hypothesis holds to be true. I want to oppose this objection.

Of course, I have defended the view that the belief in the continuity of, for example, an oak tree could either arise from constancy or from coherence. It is said that if the object changes gradually, its effects on the mind is to give rise to constancy or uniformity among individual perceptions in the series. Whereas, when the objects undergo total changes or there is a remote resemblance among individual perceptions in the series involved its effects on the mind is quite different: it exhibits coherence. However, this does not amount to rendering constancy as a special case of coherence. As I have already explained, the main reason has to do with the fact that constancy involves only a single series of resembling perceptions while coherence involves, at least, a set of series (or more) of perceptions. In the case of constancy causal inference does not play any role, whereas in the case of coherence causal inference plays a fundamental role in rendering the series coherent. Hume makes this remarkable distinction in a passage where he identifies coherence with changing objects. “This constancy, however, is not so perfect as not to admit of very considerable exceptions,” Hume writes, “Bodies often change their position and qualities, and after a little absence or interruption may become hardly knowable. But here ’tis observable, that even in these changes they preserve a *coherence*, and have a regular dependence on each other; which is the foundation of a kind of reasoning from causation, and produces the opinion of their continu’d existence” (THN, 195). It seems that this distinction is fundamental in recognizing the fact that Hume needs both constancy-based explanation and coherence-based explanation to account for the continuity of external objects. Therefore, when it is said that the belief in the continuity of, say, an oak tree could be explained either as a case of constancy or as a case of coherence, it does not mean that somehow we could render constancy as a special case of coherence; rather, it only means that we may come up with two different explanations, depending on how objects affect the mind, to account for the belief in the continuity of that object. Price has raised a much more subversive objection against Hume. Let’s turn to his objection now.

Price has claimed that we can render constancy as a special case of coherence because a common imaginative process is operative in both of them. He identifies the constant series of perceptions as a *monotonous* one, that is, of the form $A_1...A_3$, and the coherent series as a *variegated* one, that is of the form, $A...C$. Then, he introduces a

situation that seriously calls into question the validity of this distinction. He says, let's suppose that someone at 1:30 p.m. sees Jones eating cold beef, and at 7:30 p.m. she sees him eating cold beef again, not having observed him at all in the interval. Or as someone at 8 a.m. as going out of town hears the sound of a siren, and at 6 p.m. as she return she hears a very similar sound (27). Clearly, the series in these situations are of the form of A1...A3, because, observing, for instance, Jones eating cold beef at 1:30 p.m. and again seeing him eating cold beef at 7:30 p.m. exactly resemble each other. This is to say, the observer must believe that Jones has been eating cold beef all the afternoon; or that she must have concluded that the siren has been blowing all through the day. But why does she not hold such beliefs? The reason is that she has observed on previous occasions that Jones was doing other things between lunch-time and dinner-time. The implication of this is that the series involved here is no longer a constant series of perceptions of the form A1...A3; rather, it is a coherent series of the form of A...C. Price shows that there are cases that could be rendered as both cases of constancy and coherence. He concludes that we should treat constancy as a special case of coherence.

The objection is powerful. Price might have been right to note that Hume was thinking of a restricted class of cases or that he was thinking of cases where the gap is very short (28). He might even be right to conclude that we must treat constancy as a special case of coherence, but I am tempted to reject this conclusion. First and foremost, Hume himself explicitly rejects the idea of treating constancy as a special case of coherence. Second, I would like to point out that Price's claim that Hume never talks about cases like Jones' eating cold beef, or the siren sound, is not entirely true. As a matter of fact, Hume does talk about the case of a noise that is frequently interrupted and renew'd, where it is the case that the sounds have only a specific identity or resemblance while being numerically distinct. Although it is a clear case of constancy, it suggests that Hume may have been aware of the kind of problem that Price has in mind. Hume also talks about the case of a river that totally alters in less than twenty four hours, while this does not hinder the river to continue to be the same during several ages. It seems to me that this case is very similar to Price's Jones case. Similarly, the series involved in the case of a river is of the form A1...A3. Let's say that Jones observes a river at 7:30 a.m., and again he observes it at 7:30 a.m. the next day. The reason that the series involved

here is of the form A1...A3 is because the individual impressions in the series closely resemble each other. And this is sufficient to render the series constant. I may point out that the river, unlike the case of Jones eating cold beef that remains almost unchanged, alters totally in less than twenty four hours. But the perceptions of the river, or the way that it appears to the mind before the change and after it is entirely altered, is almost the same. Therefore, these remarks indicate that the series is of the form A1...A3.

Yet there is a difference between the series involved in the case of a river and the case of Jones eating cold beef. In the case of a river, unlike the latter, the observer believes in the continuity of the object during the interval. In the case of Jones eating cold beef the observer comes to believe that Jones must have been doing other things in the interval because she has observed him doing other things on previous occasions. But in the case of a river the observer comes to believe that the object continues during the interval because she has observed the river flowing continuously on previous occasions. So the difference is that in the case of Jones eating cold beef the observer believes that Jones must have been doing something else in the interval, whereas in the case of a river the observer believes that the river continuously existed in the interval. This difference is trivial, however. On the contrary, the similarity between the case of Jones eating cold beef and the case of a river is striking. Indeed, not only they are both cases of constancy, but they also both involve coherence. In both situations, the individual perceptions within the series involved exhibit constancy, and at the same time coherence plays a major role to render the series uniform. This may readily allow us to conclude that constancy is a special case of coherence. Price certainly has made that conclusion.

On the contrary, it seems that Hume is aware of the problem and continues to think that cases of constancy are different from cases of coherence. Therefore, Price's claim that Hume considers only restricted cases or that he is only thinking of cases where the gap is very short is not entirely true. The river case shows that Hume was aware of the problem, but he seems to be considering it as "remarkable in their kind" (THN, 257). That is to say, he thinks of them as exceptions to his system. Needless to say, Hume does not think that exceptional cases can seriously affect his philosophical system. Recall, the notorious case of the Missing Shade of Blue, which Hume treats it as an exception to his Copy Principle. I am suggesting that the same thing could be said about the case of Jones

eating cold beef or the case of a river. Therefore, for Hume, these restricted cases are not sufficient to render constancy as a special case of coherence.

To recapitulate, so far I have attempted to explain how Hume's standard account of constancy, which requires *exact* or strict resemblance among individual perceptions in a series, could be modified as requiring only *close* resemblance among perceptions. Noticeably, the intention for making such modification was to show that there is no inconsistency in Hume's position. In particular, I have tried to argue that, for Hume, the belief in the continuity of objects such as trees (almost all organic objects) could be explained either as cases of constancy or as cases of coherence. If objects change gradually its effects on our perceptual system is in such a way that it produces an exact or close resemblance among individual perceptions in the series, but when the change is total we only observe a remote resemblance among the perceptions. The former situation renders the series constant and the latter one renders the series coherent. Therefore, when Hume talks about the tree example alongside other cases of constancy, he most likely has in mind a gradual change in the object, which maintains a close resemblance among individual perceptions in the series involved there, whereas when he talks about an oak tree undergoing total change he seems to be thinking of it as a case of coherence.

I want to highlight a key idea of this project, which is related to my discussion here. Earlier, I bluntly stated that, for Hume, identity is in the mind, and not in the objects. I also suggested that this is where Hume departs from Locke, who thought that identity is to be specified in the objects themselves. The point is this: The argument that objects, say, trees, infants, etc., could be accounted for as either the cases of constancy or as the cases of coherence is a vindication of the claim that identity is in the mind, and not in the objects. Because different objects affect the mind differently, and the mind, through a distinctive psychological mechanism—in particular, by the help of certain characteristics of the series of perceptions, namely, constancy and coherence— generates the belief in the identity (continuity) of objects it follows that the mind itself is the sole authority in the production of the idea of identity. It appears that if the mind is affected in a constant manner (constancy) by some object, say, a tree, the result that it produces is an imperfect identity of the same object. When the mind is affected in an inconstant manner (coherence) the result is still an imperfect identity of the same object, though the route

that the mind takes is quite different from the one that it takes in the case of constancy. So to put it loosely, the mind operates differently when it postulates the identity of different objects. Overall, the identity that it ascribes to objects including itself is an imperfect identity. In a nutshell, Hume's view of identity is almost entirely different from a Lockean notion of identity in which he identified different criteria of identity for different objects. For Hume, the mind categorically lacks the idea of perfect identity, but it generates an idea of imperfect identity. He thinks that that is the way that the mind functions, and that we should be satisfied with this inadequate condition of the human mind.

I hope that it is clear now that there is a difference between the series involving constancy and the series involving coherence. For my purposes, making the distinction is extremely significant; it tells us that the constancy-based explanation is specified to account for the kind of series of perceptions that the contents of the perceptions admit of a close resemblance, whereas the coherence-based explanation accounts for those series that the contents of the perceptions are either different or remotely resembling. I think that Hume needs both the constancy-based explanation and the coherence-based explanation in order to provide a full account of the continued and distinct existence of objects. This implies that the unified explanation hypothesis ought to be rejected. Now, I should explain the principles of constancy and coherence in some detail.

III.3: Hume's Principle of Constancy

On different occasions, Hume describes constancy as a "gross illusion" (THN, 217) or tells us that it is a "deception" (THN, 202) to ascribe identity to resembling but interrupted perceptions. When we observe constancy in certain succession of related perceptions, which are interrupted and numerically distinct but resembling and qualitatively identical, the experience triggers a psychological reaction causing us to ascribe identity to the successive associated perceptions due to a 'propensity' that mistakes the resembling but successive series of perceptions with one constant and uninterrupted perception. Simply put, we mistake, or confuse, resemblance with identity; we attribute identity to a series of interrupted but resembling perceptions. Hume's intention is to provide an explanation for the source of this mistake or illusion. he says that "to enter, therefore, upon the question concerning the source of the error and

deception with regard to identity, when we attribute it to our resembling perceptions, notwithstanding their interruption; I must here recall an observation, which I have already prov'd and explain'd" (THN, 202). He refers here to a maxim that he has already established in the section of "Of the Ideas of Space and Time." There he claims that "we may establish it as a general maxim in this science of human nature, that wherever there is a close relation betwixt two ideas, the mind is very apt to mistake them" (THN, 60).

In that section, Hume provides a detailed physiological account of the propensity to ascribe identity to the succession of related objects. Although a physiological account of the propensity violates the earlier promise that he had made not to delve into anatomy (THN, 8), he seems to be tempted here to link the physiological basis of the propensity with its psychological one. Basically, the idea is that to excite any idea the mind "dispatches the spirits into that region of the brain, in which the idea is plac'd; these spirits always excite the idea, when they run precisely into the proper traces, and rummage that cell, which belongs to the idea. But as their motion is seldom direct, and naturally turns a little to the one side or the other; for this reason the animal spirits, falling into the contiguous traces, present other related ideas in lieu of that, which the mind desired at first to survey" (THN, 61). He goes on to say that "this change we are not always sensible of; but continuing still the same train of thought, make use of the related idea, which is presented to us, and employ it in our reasoning, as if it were the same with what we demanded" (THN, 61). Hume then links this physiological account of the propensity to its psychological account and claims that the psychological relation of resemblance "is the most fertile source of error; and indeed there are few mistakes in reasoning, which do not borrow largely from that origin" (THN, 61). In other words, resemblance is the primary source of the propensity to mistake, or to confound, an idea for a related one. He, therefore, concludes that "[r]esembling ideas are not only related together, but the actions of the mind, which we employ in considering them, are so little different, that we are not able to distinguish them. This last circumstance is of great consequence; and we may in general observe, that wherever the actions of the mind in forming any two ideas are the same or resembling, we are very apt to confound these ideas, and take for one another" (THN, 61). This is the kind of background setting for Hume's notion of constancy that he capitalizes on later.

Hume ends the section of “Of Scepticism with Regard to the Senses” by emphasizing the “illusory” and “deceptive” nature of the belief in the continued and distinct existence of bodies as arising from constancy. There, he claims that such belief in the existence of body is entirely unreasonable. He says, “I cannot conceive how much trivial qualities of the fancy, conducted by such false suppositions can ever lead to any solid and rational system. They are the coherence and constancy of our perceptions, which produce the opinion of their continu’d existence; tho’ these qualities of perceptions have no perceivable connexion with such an existence. The constancy of our perceptions has the most considerable effect, and yet is attended with the greatest difficulties” (THN, 217). Hence, Hume in his introductory remarks and in his conclusion characterizes constancy as a feature of the series of our perceptions that involves an “illusion.” Simply put, the belief that arises from constancy is illusory and deceptive, though that is the only source of our belief in body. I think that such is the reason why Hume characterizes constancy as a ‘gross illusion’ or a ‘deception.’

Hume reiterates his earlier claims about constancy and resemblance in the section of “Of Scepticism with Regard to the Senses.” Here, he claims:

Nothing is more apt to make us mistake one idea for another, than any relation betwixt them, which associates them together in the imagination, and makes it pass with facility from one to the other. Of all relations, that of resemblance is in this respect the most efficacious; and that because it not only causes an association of ideas, but also of dispositions, and makes us conceive the one idea by an act or operation of the mind, similar to that by which we conceive the other. This circumstance I have observ’d to be of great moment; and we may establish it for a general rule, that whatever ideas place the mind in the same disposition or in similar ones, are very apt to be confounded. The mind readily passes from one to the other, and perceives not the change without a strict attention, of which, generally speaking, ’tis wholly incapable. (THN, 203)

Again, the effect of constancy, or the propensity of the mind, is to confound any two ideas that place it in similar dispositions; it mistakes the idea of diversity for the idea of identity.

Hume before discussing the details of his account of constancy provides us with the summary of that account; particularly he outlines the psychological stages that give rise to the belief in the continuity of objects. Here is Hume’s most vivid and clear statement of the principle of constancy:

When we have been accusom’d to observe a constancy in certain impressions, and have

found, that the perception of the sun or ocean, for instance, returns upon us after an absence or annihilation with like parts and in a like order, as at its first appearance, we are not apt to regard these interrupted perceptions as different, (which they really are) but on the contrary consider them as individually the same, upon account of their resemblance. But as this interruption of their existence is contrary to their perfect identity, and makes us regard the first impression as annihilated, and the second as newly created, we find ourselves somewhat at a loss, and are involv'd in a kind of contradiction. In order to free ourselves from this difficulty, we disguise, as much as possible, the interruption, or rather remove it entirely, by supposing that these interrupted perceptions are connected by a real existence, of which we are insensible. This supposition, or idea of continu'd existence, acquires a force and vivacity from the memory of these broken impressions, and from the propensity, which they give us, to suppose them the same; and according to the precedent reasoning, the very essence of belief consists in the force and vivacity of the conception. (THN, 199)

This summary passage contains all the psychological stages that are involved in rendering a series of perceptions constant. In the rest of this section, I explain these psychological stages that constitute constancy.

In the first stage, when the mind views any object which subsists without variation or interruption, “the faculties of the mind repose themselves in a manner, and take no more exercise, than what is necessary to continue that idea, of which we were formerly possess, and which subsists without variation or interruption. The passage from one moment to another is scarce felt, and distinguishes not itself by a different perceptions or idea, which may require a different direction of the spirits, in order to its conception ” (THN, 203). Hume immediately points out that the question that “what other objects, besides identical ones, are capable of placing the mind in the same disposition...is of the last importance” (THN, 203). Then, he says that we may certainly conclude that it is the resembling perceptions that are naturally confounded with the succession of related perceptions. When the mind is considering the succession of resembling objects the mind’s disposition is almost the same disposition when it views any object with ‘perfect identity.’ “A succession of related objects places the mind in this disposition,” Hume claims, “and is consider’d with the same smooth and uninterrupted progress of the imagination, as attends the view of the same invariable object” (THN, 204). He goes on to argue that “[t]he very nature and essence of relation is to connect our ideas with each other, and upon the appearance of one, to facilitate the transition to its

correlative. The passage betwixt related ideas is, therefore, so smooth and easy, that it produces little alteration on the mind, and seems like the continuation of the same action; and as the continuation of the same action is an effect of the continu'd view of the same object...the thought slides along the succession with equal facility, as if it consider'd only one object; and therefore confounds the succession with the identity" (THN, 204).

This smooth and easy passage between related ideas triggers a psychological reaction according to which we ascribe identity to the successive related perceptions due to a 'propensity' or 'inclination' that mistakes or confuses the resembling but successive series of perceptions with one constant and uninterrupted perception. Hume links this propensity to ascribe identity to a succession of related objects to constancy as the property of sense impressions. I quote him:

We find by experience, that there is such a *constancy* in almost all the impressions of the senses, that their interruption produces no alteration on , and hinders them not from returning the same in appearance and in situation as at their first existence. I survey the furniture of my chamber; I shut my eyes, and afterwards open them; and find the new perceptions to resemble perfectly those, which formerly struck my senses. This resemblance is observ'd in a thousand instances, and naturally connects together our ideas of these interrupted perceptions by the strongest relation, and conveys the mind with an easy transition from one to another. An easy transition or passage of the imagination, along the ideas of these different and interrupted perceptions, is almost the same disposition of mind with that in which we consider one constant and uninterrupted perception. 'Tis therefore very natural for us to mistake the one for the other. (THN, 204)

What we can gather from this is that constancy makes us to overlook the numerical difference between our interrupted perceptions. In fact, it makes us to ignore the *gap* between the two resembling perceptions. Essentially, the illusion that involves in constancy consists in this act of ignoring the gap between resembling perceptions and confounding them with each other. The result of this confusion is to identify a succession of related perceptions with an uninterrupted idea; we confound the idea of diversity with that of identity. Thus, the idea of identity arising thereby is an illusory one.

In the next stage, Hume argues that we are aware of the interruption and we can see that it is contrary to the perfect identity of the different perceptions. However, we cannot always succeed in overlooking the gap. "The smooth passage of the imagination along the ideas of resembling perceptions makes us ascribe to them a perfect identity." Hume continues: "[t]he interrupted manner of their appearance makes us consider them

as so many resembling, but still distinct beings, which appear after certain intervals” (THN, 205). Therefore, on the one hand, the perceptions seem the same, but on the other hand, they are certainly distinct insofar as there is an interruption there. In fact, the palpable interruptions in the successive related perceptions induce us to consider constant but interrupted perceptions as diverse and different (THN, 199, 205, 220, 254). We take constant but interrupted perceptions as diverse and different because, on Hume’s account, identity in strict sense requires uninterruptedness and invariableness of perceptions through a supposed variation of time. Hume’s principle of individuation implies this, and because we take constant but interrupted perceptions both as identical and diverse and different at the same time “we find ourselves somewhat at a loss, and are involv’d in a kind of contradiction” (THN, 199).

The source of this conflict, or contradiction, is found in Hume’s distinction between the philosopher’s view and the vulgar’s view of the continuity of bodies. The unthinking and unphilosophical part of mankind, that is, all of us, at one time or other, suppose that their perceptions to be their only objects, and that is why “’tis to these interrupted images we [the vulgar] ascribe a perfect identity” (THN, 205). In contrast, we, as the philosophers, are reflectively aware of the interruption of the resembling perceptions, and that “naturally leads us to regard these resembling perceptions as different from each other, we here find ourselves at a loss how to reconcile such opposite opinions” (THN, 205). In a different sense, here our natural propensity comes in conflict with our philosophical reflection. It is, indeed, due to our natural propensity (the vulgar’s position) that we ascribe a perfect identity to the smooth passage of the imagination along with the ideas of the resembling perceptions. Whereas, the philosophical reflection makes us consider the resembling perceptions “as so many resembling, but still distinct beings, which appear after certain intervals” (THN, 205).

According to Hume, when the mind experiences this opposition between the notion of the identity of resembling perceptions and the interruption of their appearance, “the mind must be uneasy in that situation, and will naturally seek relief from the uneasiness” (THN, 206). Hume talks about a strategy that the mind adopts to relieve itself from the uneasiness. Accordingly, since the uneasiness arises from the opposition of two contrary principles—interruption and continuity—the mind has to sacrifice one to the

other in order to relieve itself. Hume argues that the mind cannot yield up the opinion that it is “the smooth passage of our thought along our resembling perceptions that makes us ascribe to them an identity” (THN, 206). Therefore, it turns “to the other side, and suppose that our perceptions are no longer interrupted, but preserve a continu’d as well as an invariable existence, and are by that means entirely the same” (THN, 206). The final stages results from the mind seeking relief from this kind of uneasiness. According to Hume, uneasiness is a kind of pain and the mind has an aversion to pain (THN, 438-39, 574-75). This brings us to the final psychological stage. In this final stage, to resolve the contradiction “we disguise, as much as possible, the interruption, or rather, remove it entirely, by supposing that these interrupted perceptions are connected by a real existence, of which we are *sensible*” (THN, 199). To explain this disguise, Hume describes the nature of ‘feigning,’ or postulating, the existence of certain *unexperienced* perceptions, or as some commentators have called *unsensed sensibilia* (29), “a continu’d being, which may fill those intervals, and preserve a perfect and entire identity to our perceptions” (THN, 208). Now, the fundamental question is that how the imagination is led to posit, or postulate, additional particulars in order to fill the gap? This is a question about the nature of the *feeling* that involves in generating the belief in the continued and distinct existence of objects. It is very important to know exactly the nature of these unexperienced perceptions that fill the gap and join the two interrupted perceptions together. We must know “*from whence arises such a belief?*” (THN, 208). Indeed, additional argument is required to explain why the act of postulating constitutes belief.

Let’s look at these so-called *unsensed sensibilia*, the unperceived perceptions, a little farther. As Hume emphasizes, postulating, or supposing, the existence of an unperceived perception in order to fill the intervals not only involves *feigning*, but more importantly it involves *believing*. According to Hume’s theory of belief, believing consists in nothing but the vivacity and liveliness of an idea, which gets its vivacity and liveliness through being related to some present impression. This principle is a Humean maxim that he has already proved in the Book I of the *Treatise*. It is, therefore, conceivable that there must be some impression, or impressions, that by way of filling the gap, gives rise to our belief in the continuity of objects. “The mind falls so easily from the one perception to the other,” Hume claims, “that it scarce perceives the change, but

retains in the second a considerable share of the vivacity of the first. It is excited by the lively impression; and this vivacity is convey'd to the related idea, without any great diminution in the passage, by reason of the smooth transition and the propensity of the imagination" (THN, 208). Strictly speaking, the vivacity and liveliness of the actual impressions before the gap and the one after the gap is transmitted to the idea of similar but unsensed perceptions between the gaps making the imagination slides along a train of successive or associated perceptions. What exactly happens here is that not only do we think and feign the unperceived perceptions, but we actually believe that they exist. As a result, when the mind slides along a train of successive perceptions that are qualitatively identical but numerically distinct it generates the *feeling* of being connected by one continuous and uninterrupted mental activity. "The passage betwixt related ideas is, therefore, so smooth and easy that it produces little alteration on the mind, and seems like the continuation of the same action" (THN, 204). Here, then we have a propensity of the imagination that feigns the continued and distinct existence of all sensible objects and since that propensity arises from some lively impressions, it makes us believe the continued and distinct existence of bodies. This is Hume's core idea about the mind's propensity of identity ascriptions and generation of the belief in objects.

Constancy and the Identity of Material Substrata and the Self

I am inclined to think that Hume's theory of identity conforms to a general scheme of explanation. I must show now how these stages operate in the production of the belief in the continuity of material substrata, and the self. Hume provides his most detailed account of the propensity to ascribe identity to related objects in I.iv.2 of the *Treatise*, where he spends almost eleven pages to explain the vulgar's belief in bodies. His accounts of the effects of constancy in producing the ideas of material substrata and the self are brief, and as I claim, depended upon his account in I.iv.2 of the *Treatise*. In particular, he talks about the propensity to ascribe identity to related objects on pages 219-220, and on pages 253-254 in a similar fashion as he talked in pages 202-210. That's why his treatment of these cases is relatively short and to a certain extent incomplete. Moreover, there are explicit textual evidences that Hume refers the reader to his accounts of personal identity in section of "Of personal Identity" in order to grasp his full account of the principle of constancy. For instance, in page 206 where he discusses the strategy

that the mind adapts to seek relief from the uneasiness that arises from the opposition of two contrary principles, in the footnote 40n he refers us to the section of “Of Personal Identity.” So it seems that scholars like Loeb are quite right when they claim that “Hume intends his subsequent discussions of the identity-ascribing propensity to be read against the background of the initial discussion of the propensity at I.iv.2” (30).

Needless to mention, Hume’s usage of technical terminology is strikingly similar in all these cases. He often talks of the propensity to ascribe identity as leading to the “supposition” (THN, 198, 208) or “fiction” (THN, 205, 209) of the continued and distinct existence of external objects, and that when we suppose, we, indeed, “feign” (THN, 208, 209) the existence of bodies. In a similar way, when the ancient philosophers “supposed” (THN, 220) the existence of material bodies as remaining unchanged, they “feigned” (THN, 220) the existence of a material substratum; this is one of their “fictions” (THN, 219, 222, 224, 226). Likewise, Hume writes that the belief in the existence of a soul, or mind, is a “fiction” (THN, 254, 255); again, we “suppose” (THN, 253) the existence of “a *soul*, and *self*, and *substance*” to “feign” (THN, 254) the continued existence of the perceptions of our senses.

Hume gives his explanation of the belief in material substrata in the section of “Of the Ancient Philosophy.” In this section, those schematic stages by which we arrive at the belief in the continued and distinct existence of bodies are clearly distinguishable as well. Hume’s main question here is why, given that “our ideas of bodies are nothing but collections form’d by the mind of the ideas of the several distinct sensible qualities, of which objects are compos’d, and which we find to have a constant union with each other” (THN, 219), do we still “regard the compound, which they form, as ONE thing, and as continuing the SAME under very considerable alterations” (THN, 219). As before, Hume’s answer lies in the way that our perceptions affect the mind. The schematic psychological stages in this case go as follows: First, when the perceptions of “several distinct *successive* qualities of objects are united together by a very close relation, the mind, in looking along the succession, must be carry’d from one part of it to another by an easy transition, and will no more perceive the change, than if it contemplated the same unchangeable objects” (THN, 220). Again, the smooth and uninterrupted progress of thought deceives the mind, and makes us ascribe an identity to the changeable succession

of connected qualities (THN, 220). As a result, we readily consider these related successive sensible qualities as one continued object, existing without any variation.

In the second stage, we encounter a “kind of contrariety in our method of thinking” (THN, 220). Basically, identity requires that an object “remains invariable and uninterrupted thro’ a suppos’d variation of time,” but after comparing and surveying the different conditions of the successive sensible qualities before and after an interval during which it undergoes considerable change or variation, we realize that the progress of thought is broken, and that the identity of the object is entirely destroyed (THN, 220). This gives rise to a contradiction in our thinking and presents us with the idea of diversity rather than with the idea of identity. This triggers a psychological reaction because it causes uneasiness here. Finally, “to set themselves [the ancient philosophers] at ease” (THN, 224), and “in order to reconcile which contradictions the imagination is apt to feign something unknown and invisible, which it supposes to continue the same under all these variations; and this unintelligible something it calls a *substance*, or *original and first matter*” (THN, 220). Therefore, they postulate, or suppose, a material substratum for these successive sensible qualities.

Specifically, Hume discusses the cases of substantial forms, accidents, and occult qualities and that how the peripatetic philosophers have come up with these ‘fictional’ ideas that are “entirely incomprehensible, and yet is deriv’d from principles as natural as any of those above-explain’d” (THN, 222). Hume explains, though very briefly, how these fictional ideas are generated by those schematic psychological stages as explained above. For instance, when we observe a succession of related parts in compound object it has almost the same psychological effect as if the imagination conceives a simple object without change or variation. Here again, the smooth and easy passage of thought along the succession of related sensible qualities places the mind in the same disposition as if it attends the view of the same invariable object. “Hence the colour, taste, figure, solidity, and other qualities, combin’d in a peach or melon, are conceiv’d to form *one thing*; and that on account of their close relation, which makes them affect the thought in the same manner, as if perfectly un compounded ” (THN, 221). The mind, however, does not rest here. “Whenever it views the object in another light, it finds that all these qualities are different, and distinguishable, and separable from each other” (THN, 221). The outcome

is a contradiction and “these philosophers set themselves at ease, and arrive at last, by an illusion, at the same indifference, which the people attain by their stupidity, and true philosophers by their moderate scepticism” (THN, 224). Therefore, the contradiction “obliges the imagination to feign an unknown something, or *original* substance and matter, as a principle of union or cohesion among these qualities, and as what may give the compound object a little to be call’d one thing, notwithstanding its diversity and composition” (THN, 221). Hume sees the same psychological mechanism operative in the generation of the ideas of accidents and occult qualities as well.

Generally speaking, Hume’s account of the belief in the identity of sensible qualities, as arising from the constancy of our series of perceptions and the identity-ascribing propensity, is relatively short. Loeb, again, rightly notes that “Hume can make do with a relatively brief discussion of the propensity’s role in producing the beliefs in substrata; they closely parallel the three-stage reaction that produces the belief in the continued existence of perceptions, discussed at length at pages 199-210” ⁽³¹⁾. Hume treats the subsequent cases of the double existence of perceptions, material substrata and the self as instances of the identity-ascribing tendency. As a matter of fact, Hume writes: “We shall afterwards see many instances of this tendency of relation to make us ascribe an *identity* to *different* objects” (THN, 204).

To move to Hume’s account of personal identity, the propensity to ascribe identity to related perceptions is, again, responsible for the production of the belief in the identity of the self. As before, the mind goes through different schematic stages to produce the belief in the identity of person. Here, Hume spends almost two pages to establish the idea of the mind, or the self, as “nothing but a bundle of collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement” (THN, 252). These successive perceptions constitute the mind alone; “there is properly no *simplicity* in it at one time, nor *identity* in different” (THN, 253). The reason that there is no simplicity and identity in the mind is because the principle of individuation entails that identity, in a perfect sense, requires an object to remain invariable and uninterrupted thro’ a suppos’d variation of time (THN, 253). Since the mind is constituted by successive related perceptions (series of perceptions) it follows that we somehow confound the idea of identity with that of related objects (diversity). In

the first stage, the identity-ascribing propensity “facilitates the transition of the mind from one object to another, and renders its passage as smooth as if it contemplated one continu’d object” (THN, 254). Indeed, it is the smooth passage of the imagination along the ideas of the resembling perceptions that places the mind in the same disposition as if we are experiencing an uninterrupted and invariable object.

In the next stage, philosophical reflection provides us with the clear idea that identity, in a strict sense, requires that the idea of the self remains uninterrupted and invariable for a supposed variation of time. Once more, we face a contradiction here; “at one instant we may consider the related successions as variable and interrupted, we are sure the next to ascribe to it a perfect identity, and regard it as invariable and uninterrupted” (THN, 254). The text is strikingly clear:

Our propensity to this mistake is so great from the resemblance above-mention’d, that we fall into it before we are aware; and tho’ we incessantly correct ourselves by reflection, and return to a more accurate method of thinking, yet we cannot long sustain our philosophy, or take off this bias from the imagination. Our last resource is to yield to it, and boldly assert that these different related objects are in effect the same, however interrupted and variable. (THN, 254)

This passage clearly shows that, in the final stage, the arising contradiction causes uneasiness, and, for the mind to set itself at ease, it has to “disguise” or “feign some new and unintelligible principles, that connects the objects together, and prevents their interruption or variation” (THN, 254). Therefore, “we feign the continu’d existence of the perceptions of our senses, to remove the interruption; and run into the notion of a *soul*, and *self*, and *substance*, to disguise the variation. But we may farther observe, that where we do not give rise to such a fiction, our propension to confound identity with relation is so great, that we are apt to imagine something unknown and mysterious, connecting the parts, besides their relation” (THN, 254-5).

To sum up, in all these cases, the beliefs in objects, material substrata, and the self arise as a result of the constancy of the series of impressions (perceptions). In each case, the propensity to ascribe identity to related objects leads to a contradiction, or uneasiness, in the mind, and in order to set itself at ease, the mind feigns, or postulates, an unperceived perception. In short, it is the operation, or function, of this identity-ascribing propensity, as arising from the constancy of the series of our perceptions, which is responsible for the generation of the beliefs in body, material substrata, and the self. I

should mention two related points now: Firstly, these so-called schematic stages are also observable in the explanation of the belief in the existence of double vision (THN, 210-215) and the local conjunction of taste with matter (THN, 235-38), which for the sake of brevity, I will not discuss here ⁽³²⁾. Secondly, I should point out that I have not been concerned above about the epistemic/justificatory status of the identity-ascribing propensity as it gives rise to the beliefs in body, material substrata, and the self. I have not been concerned whether the identity-ascribing propensity results in a doxastic stable state or if it leads to an instable belief in body, material substrata and the self. Perhaps, there might be some tensions in Hume's account of the psychological propensity to ascribe identity to related objects. It may turn out, as Loeb has forcefully argued, that "Hume has no uniform account of the psychological properties of the propensity to attribute identity to related objects" ⁽³³⁾ resulting in an underlying instability in our beliefs in material substrata and the self ⁽³⁴⁾. Nevertheless, it is a fact that, in Hume's view, constancy and the resulting propensity to attribute identity to related objects are genuinely responsible for the production of our beliefs in objects, material substrata, and the self. And that the belief as arising in this manner can function to "conceal" the interruption of an object, and disguise the contradiction—the underlying instability in our belief system—and hence gives rise to stability in our beliefs in objects, material substrata, and the self. So much for Hume's principle of constancy. Now I shall consider Hume's principle of coherence.

III.4: Hume's Principle of Coherence

Hume argues that coherence must be combined with constancy in order to explain our tendency to believe in the existence of objects. He thinks that the imaginative process involved in coherence is different from the one involved in the case of constancy. Coherence plays a supplementary role; "it gives us a notion of a much greater regularity among objects, than what they have when we look no farther than our sense" (THN, 198). Coherence, unlike constancy which involves resemblance between *individual* perceptions, involves resemblance between *series* of perceptions—that is, between interrupted and fragmentary series observed now and other series observed in the past. Moreover, there is another difference between them: the identity-ascribing propensity, which was essential to Hume's explanation of the belief in the body, material substrata,

and the self as arising from the constancy of perceptions, plays no role in his coherence-based explanation. Furthermore, Hume tells us that coherence involves “a kind of reasoning from causation” while in the case of constancy causal inference plays no role whatsoever. These differences are quite substantial. I suggest that these differences explain why Hume does not treat his constancy-based explanation as a special case of coherence-based explanation.

Hume argues that constancy alone cannot account for the continuity of all objects. “Bodies often change their position and qualities,” he says, “and after a little absence or interruption may become hardly knowable” (THN, 195). Even in these changes objects “preserve a *coherence*, and have a regular dependence on each other; which is the foundation of a kind of reasoning from causation, and produces the opinion of their continu’d existence” (THN, 195). As I argued, Hume’s account of coherence is clearly a kind of argument from analogy. Hume’s argument from analogy consists in a resemblance between a set of series (or more) of impressions—between a fragmentary series now observed and other continuous series observed in the past—and not between individual perceptions ⁽³⁵⁾. More importantly, Hume characterizes coherence, unlike constancy, as involving *changing* bodies. These preliminary points also show that there are substantial differences between Hume’s constancy-based explanation and his coherence-based explanation of our belief in the continued existence of objects.

I have shown earlier that in the case of constancy the mind postulates or feigns the existence of an *unsensed sensibillum* in order to ascribe identity to an interrupted series of perceptions. This is not unique to constancy, however. The same imaginative process is operative when “our imagination is led to fill in the missing parts of the present fragmentary series from our memory of the previous continuous ones. And in so filling them in, we are, of course, postulating the existence of particulars which we have not on this occasion sensed (though we *have* sensed particulars of that kind in like circumstances in the past). We are postulating, indeed, the existence of an unsensed sensibilia” ⁽³⁶⁾. Price provides us with an excellent example describing how the mind comes up with an unsensed sensibilia. “I see a black cat in one corner of the room,” he says, “then I turn to read *The Times* for half a minute, then I look up again and see a black cat in the opposite corner of the room, while the first corner is empty. But on many previous occasions I

have watched black cats walk all the way across rooms from one corner to the other. And owing to this analogy between this broken series and those continuous ones, I imagine that here also there has been a continuously existing cat which has walked continuously all the way across” (37). Simply stated, the observer postulates that the cat existed in those moments that he had no impressions of the cat walking from one corner to the other.

Hume’s own example in which we could observe the postulating of an unsensed sensibilia is the porter-letter case. In this case, which is comprised of several broken series combined together, Hume explains how we come to suppose that objects exist and can subsist even when we are not observing them. Hume while sitting in an upstairs room hears a noise as of a door turning upon its hinges (auditory impression), but does not have visual impression of the door. Then, a little after see a porter, who gives him a letter that is said to come from his friend two hundred leagues distant. The following broken series are distinguishable in this case: (a) The noise and the motion of the door on its hinges; (b) the body of the porter and the stairs; (c) the letter and the continued existence of posts and ferries. In all these broken series he postulates the existence of some objects that he lacks visual impressions of them. In (a) he supposes that the door turns upon its hinges, and in (b) he postulates the existence of the stairs, even though he does not observe the porter coming up the stairs, and he does not observe the stairs at all. Also, in (c) he assumes the existence of the sea, posts and ferries existence without having any visual impression of them.

Hume describes his belief in the existence of these unobserved objects as arising from customary constant conjunction. He states: “I never have observ’d, that this noise cou’d proceed from anything but the motion of a door; and therefore conclude, that the present phenomenon is a contradiction to all past experience, unless the door, which I remember on the other side of the chamber, be still in being” (THN, 196). The claim here is that his observation of constant conjunction of the noise of a door turning on its hinges preceding the motion of the door in the past is the background upon which he infers the existence of the door’s motion. Likewise, he supposes the existence of the stairs and the seas, posts and ferries according to the same principle. Nevertheless, Hume immediately points out that these conclusions concerning the existence of unobserved objects “are contradictions to common experience, and may be regarded as objections to those

maxims, which we form concerning the connexions of causes and effects” (THN, 196). The reason is that because he is “accustom’d to hear such a sound, and see such an object in motion at the same time,” which he has not “receiv’d in this particular instance both these perceptions” (THN, 196). Hume goes on to argue that, “[t]hese observations are contrary, unless I suppose that the door still remains, and that it was open’d without my perceiving it: And this supposition, which was at first entirely arbitrary and hypothetical, acquires a force and evidence by its being the only one, upon which I can reconcile these contradictions” (THN, 196-97).

Price has argued that the contradiction that Hume refers to in the passage is, indeed, a contradiction between a causal maxim, that is, the constant conjunction of our past experiences, and a broken series. He seems to be saying that the contradiction lies between the observation of the noise not being preceded by the motion of the door (the broken series) and the observed cases in which the noise always has been preceded by the motion of the door. Simply put, we are accustomed to always observing the noise and the door in motion in a constant manner, but in the broken series we only observe the noise without observing the door opening. As a result, we encounter a situation that flies in the face of our customary expectation derived from the constant conjunctions of the noise and the door opening. In order to remove this contradiction, Price says, we have to “suppose that objects can exist and events occur when I am not observing them: i.e. that my fragmentary sense-data are continued into unsensed sensibilia” (38). Price concludes that the imaginative process in postulating unsensed sensibilia cannot be a case of ordinary causal reasoning. Because, he claims, “it is something which ensures the truth of the very causal rules upon which such reasoning depends. It is not itself causal reasoning, because it is something more fundamental, without which causal reasoning would not stand; for without it all the major premises used in such reasoning would be utterly precarious, and any drowsy nod would refute them” (39). In short, while the causal inference is nothing except the customary transition of the imagination, the postulation (inference) in the present situation, e.g. that of the cat, the opening door, the porter and the letter, “we find ourselves imagining a regularity *greater than that which we actually observe*; and this cannot be accounted for by mere custom” (40).

Similarly, Loeb (41) and Bennett (42) have suggested that the contradiction that

Hume has discovered about his causal maxim must be applicable to any instance of causal inference as well. In fact, Loeb explicitly states that “strictly speaking, there is no contradiction between the experience of an observed regularity and the occasion for a causal inference. The fact that the noise is not observed to have been preceded by the motion of the door in a given case is not an exception to the regularity that, *in observed cases*, the noise always has been observed to be preceded by the motion of the door”⁽⁴³⁾. Loeb then suggests that the contradiction may arise for the following reason. He says, “[t]he noise has always been observed to be preceded by the motion of the door, *except* in the case that serves as the occasion for the causal inference. For that reason, the *observed* conjunction of the noise and the motion is not, strictly speaking, constant”⁽⁴⁴⁾. This statement is very ambiguous. It seems that it is saying that in the broken series in which we infer the existence of the unobserved moving door as the cause of the observed noise, the inference *itself* relies on the observation (experience) of previously observed series of perceptions exhibiting constant conjunction between the noise and the door opening. He claims that, strictly speaking, “there is no observed *constant* conjunction even prior to the present instance”⁽⁴⁵⁾. Because “the present instance of hearing a noise as of a door opening is not the *only* exception to an observed regularity between the noise and the motion... [Indeed], just as, in the present case, Hume does not observe the motion of the door, there have been previous occasions when he has heard the noise but not observed the motion—because he turned his head, shut his eyes, and so forth”⁽⁴⁶⁾. In fact, Loeb relies on a passage in which Hume states, “we remark a connexion betwixt two kinds of objects in their past appearance to the sense, but are not able to observe this connexion to be perfectly *constant*, since the turning about of our head, or the shutting of our eyes is able to break it” (THN, 197-98). In short, the reason that the observed conjunction of the noise and the motion is not constant is that there have been previous occasions that Hume has heard the noise without observing the motion. More fundamentally, Loeb’s reading relies on a point that Hume has made in which “any occasion for a causal inference itself destroys an observed constant conjunction, if that conjunction is supposed to extend to the occasion for the inference itself”⁽⁴⁷⁾. Thus, for Loeb, the reason for Hume’s claim that the case of the door and letter are contradictions to common experience and objections to causal maxims is that any inference from the observed noise to the existence of the

unobserved moving door as its cause is contradictory to the constant conjunction between causes and effects.

What we can gather from these arguments is the following: First, in the case of coherence, like constancy, we come to postulate the existence of an unsensed sensibillum; second, the conclusion derived from the mechanism activated by coherence involves causal inference; thirdly, the causal inference involves more than a series of perceptions; fourthly, it follows that coherence involves more than a series of perceptions. Last, since each instance of casual inference itself destroys the constant conjunction (constancy) of the series of perceptions observed on previous occasions it entails that there is a *gap* within the regularity arising from the former. This last point is linked to a question, which has largely remained unanswered, that is, why does Hume think that the conclusion that we derive from causal inference is “considerably different” from the conclusion that we derive from the mechanism activated by coherence?

Hume states that on the occasion of a causal inference we postulate the existence of an unobserved object to give “force and evidence” to that “supposition, which was at first entirely arbitrary and hypothetical” (THN, 196-97). The arbitrary and hypothetical supposition that Hume talks about it in this context seems to be the mere psychological habit” or “custom” that the mind develops after having repeated experiences of constant conjunction between objects. Hume claims that by postulating the existence of the unobserved object the mind acquires a force and evidence that elevates that arbitrary and hypothetical supposition to something real and actual; something over and above the mere inductive generalization. In fact, he explicitly says, “the extending of custom and reasoning beyond perceptions can never be the direct and natural effect of the constant repetition and connexion, but must arise from the co-operation of some other principles” (THN, 198). Hume leaves no doubt that the greater regularity that we arrive at in unobserved objects can never be founded upon the regularity we derive from our perceptions. He writes:

’Tis not only impossible, that any habit shou’d ever be acquir’d otherwise than by the regular succession of these perceptions, but also that any habit shou’d ever exceed that degree of regularity. Any degree, therefore, of regularity in our perceptions, can never be a foundation for us to infer a, greater degree of regularity in some objects, which are not perceiv’d; since this supposes a contradiction, viz. a habit acquir’d by what was never present to the mind. But ’tis evident, that

whenever we infer the continu'd existence of the objects of sense from their coherence, and the frequency of their union, 'tis in order to bestow on the objects a greater regularity than what is observ'd in our mere perceptions. (THN, 197)

To explain the nature of this greater regularity Hume introduces a “so curious and questionable” principle of his science of human nature, namely, the Inertia Principle (48). Accordingly, “the imagination, when set into any train of thinking, is apt to continue, even when its object fails it, and like a galley put into motion by the oars, carries on its course without any new impulse” (THN, 198). Hume’s earlier statement of the principle takes place in his explication of mathematical notion of equality in I.ii.4. There, Hume tells us that we are sensible that the addition or removal of any minute parts to two geometrical figures alter their equality, but we “suppose some imaginary standard of equality, by which the appearances and measuring are exactly corrected...For as the very idea of equality is that of such a particular appearance corrected by juxtaposition or a common measure, the notion of any correction beyond what we have instruments and art to make, is a mere fiction of the mind, and useless as well as incomprehensible. But tho’ this standard be only imaginary, the fiction however is very natural; nor is any thing more usual, than for the mind to proceed after this manner with any action, even after the reason has ceas’d, which first determin’d it to begin” (THN, 48). The inertia of imagination is operative in the case of rendering the time proportionable and many other subjects. “A musician finding his ear becoming every day more delicate,” Hume says, “and correcting himself by reflection and attention, proceeds with the same act of the mind, even when the subject fails him, and entertains a notion of a complete *tierce* or *octave*, without being able to tell whence he derives his standard. A painter forms the same fiction with regard to colours. A mechanic with regard to motion. To the one *light* and *shade*; to the other *swift* and *slow* are imagin’d to be capable of an exact comparison and equality beyond the judgments of the senses” (THN, 48-9).

Accordingly, the same imaginative process that renders equality so “correct and exact a standard of that relation...makes us easily entertain this opinion of the continu’d existence of body” (THN, 198). Objects exhibit a kind of coherence when they appear in the senses, but as Hume argues, “this coherence is much greater and more uniform, if we suppose the objects to have a continu’d existence” (THN, 198). Again, when “the mind is

once in the train of observing an uniformity among objects, it naturally continues, till it renders the uniformity as complete as possible” (THN, 198). Hume, once more, emphasizes the idea that the greater regularity among objects arises from some principle other than the mere constant repetition of objects.

The greater regularity among objects is due to some principle over and beyond the mere customary repetition as inherent in the nature of habit. Loeb, for instance, notes:

If we infer the existence of the unobserved moving door as the cause of the observed noise, we can maintain, consistently with what we have observed and inferred, that the motion *always* precedes the noise. We can maintain this even though we have not observed the door’s motion in every instance in which we heard the noise. The claim that the motion always precedes the noise is distinct from the claim that the motion *always has been observed* to precede the noise. If we decline to infer the existence of the unobserved moving door, we cannot maintain that the motion of the door *always* precedes the noise. The effect of the inference to the existence of the moving door is to enable us to maintain...that objects possess a greater degree of regularity than they have been observed to possess, and a perfect regularity in particular. (49)

Loeb provides an interesting argument for Hume’s distinction between the greater regularity as arising from the inertia principle and the mere regularity inherent in the nature of habit. He thinks that we can distinguish between two sorts of regularities in Hume’s discussion, that is, *gappy regularities* and *statistical regularities*, and that we are unable to assimilate gappy regularities to statistical regularities. To elucidate, let’s assume that on one-third of the occasions that we have observed the noise of a door opening, we have not observed the motion of the door. Now, if we treat this as a case of statistical regularity (probability of causes), we would certainly believe that there is *only* a two-thirds probability that the motion of the door preceded the noise. However, “this is not what we believe; we believe that it is overwhelmingly probable, with a likelihood approaching one, that a door moved there a moment ago” (55). The point is that no matter what the likelihood of our observed statistical regularity may be, it never supplies the kind of absolute certainty that involves in our belief in the existence of a door in motion. Hence, we conclude that gappy regularities cannot be assimilated to statistical regularities, and this in turn, leads to the conclusion that the imagination bestows a much greater degree of regularity upon object than what they are observed to possess.

Price has argued that we may approach the question of custom and the degree of its regularity from two different directions. First, it may be said that there is no other

types of regularity over and above the customary repetition of our experiences as inherent in the nature of habit. And that the greater regularity that Hume talks about in the inertia principle “is simply the ordinary everyday inertia which is inherent in all habits. That we imagine a regularity greater than we actually observe is therefore no paradox; it is exactly what we should expect, and Hume is making a fuss about nothing” (51). Given this, it is unnecessary to provide another explanation for the regularity of the imagination, and thus, the inertia principle becomes redundant. However, Price says that we may legitimately ask whether Hume’s ‘habits of the imagination’ are really those ordinary habits that we acquire by the mere repetition. Then, he suggests that the imaginative process occurring here, which is a kind of argument from analogy, involves “something more, namely, a tendency to *persist in* our inductive generalizations even in the absence of favourable evidence; i.e. a tendency to continue supposing that there is order in the world even when order is not at once obvious, and a consequent tendency to form hypotheses such that (if they are true) this order extends as far as possible” (52). This tendency not merely to generalize but to extend our generalizations as far as possible, Price argues, seems to be simply Hume’s way of referring to what Kant called an Idea of Reason. If it turns out that Hume’s Inertia Principle is something like the Kantian Idea of Reason, then it seems that not only is the tendency to extend our generalization as far as possible something over and above the mere repetition of our experiences, but it is also something real and important. Thus, the Inertia Principle is a fundamental act of the imagination, which involves only cases of coherence, and this in turn, implies that coherence is fundamentally different from constancy.

Conclusion

To recapitulate, the purpose of the present chapter was to explain Hume’s psychology of identity ascriptions. In particular, I wanted to show how the imaginative process of postulation or assimilation of an unsensed sensibilia, which fills up the gaps between the series of perceptions, produces the idea of identity. I have shown that such process of imaginative supplementation operates in both cases of constancy and coherence. More specifically, I have argued that the idea of the continued and distinct existence of external objects neither arises directly from the senses nor from reason; instead, it arises from the imagination. It is also said that for Hume there is no an idea of

the self considered as something having ‘perfect identity’ and simplicity. It is not something that we could directly observe and experience either by the senses or by reason; rather, there is a propensity to ascribe an ‘imperfect identity’ to successive perceptions in the mind. Therefore, both the identity of the self and the external object is something that we merely attribute to them as a result of various operations of the imagination. The rationale goes something like this: If both the self and the external objects lack such ‘perfect identity’ while at the same time we have a great propensity or inclination to ascribe an ‘imperfect identity’ to them, it should follow that the mere existence of such an inclination which attributes ‘imperfect identity’ to the self and the external body requires some explanatory pattern to justify it. Hume believes that that explanatory pattern is a functionalistic pattern of justification.

Essentially, our sense-perceptions provide us with fragmentary, interrupted, and episodic experiences of the world around us, including the idea of ourselves as the cognizant of these realities. However, the mind exhibits a higher function as well. It exhibits a great propensity to ascribe continuity (identity) to these fragmentary and episodic mental objects. In Hume’s view, the imaginative postulation of unsensed sensibilia provides us with the mechanism for such identity-ascription. At the core of the imaginative postulation of unsensed sensibilia lies the idea that the former is essentially an unreflective, automatic, and sub-doxastic process. This, I assume, explains to a great extent Hume’s functionalistic theory of mind. It is not the case that we become aware of an interruption in our sense-experiences and, then, consciously postulate an unsensed sensibilia to fill the gap. On the contrary, the process of postulating or assimilating an unperceived perception in the gap is an automatic and unconscious activity of the mind. It occurs so rapidly and automatically that we not only fail to notice the occurrence of gaps themselves, but we also fail to notice the very imaginative process of postulation as well. It could be said that even the vulgar sometimes notices that she experiences interruption in her sense perceptions, and as a result of that she may realize that she consciously postulates an unsensed sensibilia. Or, as Hume argues, one may through philosophical reflection come to some sort of awareness about these things, but that does not mean that the imaginative process of postulation is a conscious act of the mind. The philosopher may enjoy reflective awareness of these kinds of things, but even he easily fails to attend

to them. Philosophical reflection makes us momentarily aware of the interruption and the process of postulation, but it does not play any role in the working of the process of postulation itself. It is not the case that the philosopher consciously postulates an unsensed sensibilia whenever she encounters a gappy situation. Accordingly, the postulation is done by the kind of operation that is inherent in the act of the imagination. According to Hume, this is nature's remedy for the fragmentariness of our experiences of the external world and the self. In short, we may sometimes be aware of the fact that we somehow postulate, or supplement, our sense perceptions data, but the actual process conceals itself from our conscious mind. It occurs beneath our consciousness, that is, in the sub-doxastic states of our belief system. Therefore, we may find that we have supplemented our sense data, "but we do not notice ourselves actually doing it. Still less do we do it deliberately and with an intellectual effort, after a process of questioning and wondering ... On the contrary, the utmost effort is needed to *undo* it, and to contemplate the data in all their nakedness and fragmentariness. And this artificial suspension of supplementation never lasts long" (53). Needless to say, all these processes occur below the level of our consciousness, and we may label it as sub-doxastic state of our consciousness.

In the previous chapters, I tried to show that Hume's theory of the self can be interpreted in terms of a causal or functional theory of mind. In particular, I have distinguished between two different senses of functionalism in Hume's theory of the self, namely, teleological functionalism and psycho-functionalism. The former was rooted in his teleological characterization of the mind according to which the bundle of perceptions persists over time by maintaining functional continuity, similar to the kind of functional continuity which constitutes the identity of organic objects and artifacts (Chapter I). And the latter consisted mainly in his representational theory of ideas (Chapter II). Also, I have examined Hume's psychology of identity ascriptions (Chapter III). Yet, one of the main objectives of this work is to show that Hume is committed to a version of psychological-functional continuity account of personal identity. Specifically, I claim that Hume is committed to the view that functionalism implies personal identity. This is an important question in contemporary debate on personal identity. I turn to this in the following chapter.

IV

Hume's Functionalism and Personal Identity

And in this view our identity with regard to the passions serves to corroborate that with regard to the imagination, by the making our distant perceptions influence each other, and by giving us a present concern for our past or future pains and pleasure. (THN, 261)

There is a consensus among philosophers that persons, or selves, somehow persist over time. However, they take different approaches in explaining what personal identity involves, or what constitutes it. Cartesians held that personal identity consists in an immaterial self, or soul, that is irreducible into time-sliced entities, and persists over time and even survives death. There are a few people, like Richard Swinburne, who still defends something like the Cartesian ego/soul theory of the self. Locke argued that consciousness-memory makes personal identity. Many of present-day philosophers hold that some version of the Lockean psychological continuity constitutes personal identity. Hume of course not only recognizes the idea that persons, or selves, persist over time, but also offers a genetic explanation about its persistence over time (Chapter III). In fact, Hume allows that we have an idea about ourselves as unified mental subjects, though he denies the idea of the self as an indivisible, immaterial and unchanging substance. In particular, Hume holds that the belief in the existence of ourselves as unified mental subjects persisting over time is caused by a *feeling*. This feeling *mistakenly* makes us believe that our individual perceptions are connected by a single train of mental activity, or mental continuity, even though, strictly speaking, there is no such a mental continuity that *really* connects our individual perceptions.

Roughly, Hume characterizes such mental continuity in terms of some form of functional continuity, similar to the kind of functional continuity that constitutes the identity of organic objects and artifacts. But more fundamentally, his account of character

traits, self-concerns, and self-interests in Book II of the *Treatise* shows that he is very much concerned about continuity of the self over time. Given that, Hume is committed to the functionalist theory of mind it follows that, for him, functionalism must imply personal identity. In fact, Hume is the only modern philosopher who links the traditional mind/body problem to the question of personal identity. He seems to be suggesting that functionalism as a theory about the mind/body problem has not only important and relevant ramifications for personal identity, but also it implies personal identity.

Contemporary philosophers of mind treat the problem of personal identity as a separate issue from the popular mind/body theories. It is assumed that these theories—logical behaviourism, identity theory, and functionalism—either have no relevance for the problem of personal identity or almost no ramifications for it. This means that the psychological theorist of personal identity can be a materialist and the physical theorist can hold a mind/body dualism. It is commonly believed that the traditional mind/body problem, which is strictly a view about the metaphysical status of mental states and their relations with the physical states, has no implications on the question of personal identity, which is a psychological concern about our persistence over time. However, Sydney Shoemaker has famously argued that functionalism implies a kind of psychological-functional continuity account of personal identity. He claims that functionalism can solve the problem of personal identity. Shoemaker's suggestion that functionalism implies a version of the psychological-functional continuity criterion of personal identity has raised positive as well as negative reactions. Some philosophers, like Eric Olson and Garry Fuller, have rejected it altogether while others, like Lawrence Davis and Nicholas Agar, have offered a partial defence of Shoemaker's thesis. In particular, both Olson and Fuller have rejected the claim that functionalism implies that psychological/functional continuity is necessary and sufficient for personal identity. Whereas, Davis thinks that functionalism does imply a sufficient condition for personal identity, but not a necessary condition.

Hume, like Shoemaker, maintains that not only functionalism is the right theory about the nature of mental properties, but also that functionalism tells us important things about personal identity. He not only explains the basic tenets of a functionalistic theory of mind, but he also explains how functionalism implies a theory about the identity

condition of persons over time. However, he does not succeed to carry out his project, and that seems to be one of the main sources of his self-criticism in the Appendix. In fact, I want to argue that one possible source of Hume's inconsistency, as I understand, has to do with the question of functionalism and its implication on personal identity over time. I suggest that Hume might have detected circularity in his original characterization of personal identity; something similar to the one Butler levelled against Locke claiming that consciousness already presupposes personal identity; or even something similar to what Eric Olson has raised against Sydney Shoemaker's psychological-functional continuity account of personal identity claiming that the identification of a mental state in terms of its combinatory relations with other mental states that are members or states of *the same person*, invokes or presupposes a notion of personal identity. Basically, the claim is that if Hume is right to say that something like functional continuity would suffice for persons to persist through time, then he must show that we can have a *complete* account of how one's mental states produce the idea of a persisting self without making assumption about the identity (continuity) condition of their subject or bearer. And of course, psycho-functionalism, including Hume's, identifies a mental state in terms of its functional relations to other mental states that are the states of *the same person*. This is straightforwardly circular.

Hume could have circumvented the circularity if he had a notion of a simple substance in which perceptions inhere or that the mind perceived some real connection among perceptions. I will argue that Shoemaker's functional continuity account according to which a person is identified as an ontologically independent substance upon which mental properties depend, and that constantly and autonomously perpetuates itself over time, is the right kind of theory that Hume, and functionalism for that matter, needs for his account of personal identity over time. If Hume, like Shoemaker, had identified the self as an independent substance that perpetually and spontaneously causing its successor states, then he would not have found himself in a situation to denounce his original characterization of the self as a fiction. Or if he conceived psychological-functional continuity in terms of *immanent* causation according to which things retain their properties over time while undergo changes, he wouldn't had to face circularity. By and large, Hume plays a distinctive and significant role in showing us that functionalism

tells us important things about the self and its identity over time. Hence, I suppose that Hume's contribution to the problem personal identity and its psychological-functional continuity over time is significant, though it stumbles upon some serious difficulty.

In IV.1, I will examine Hume's account of character traits, self-concern, self-interest, and personal identity. My primary concern is to show that Hume's concept of character traits and self-interests provide a connecting link between his functionalist theory of mind and his account of personal identity. In IV.2, I provide a more detailed account of the psychological-functional continuity theory of personal identity. In particular, I present and defend Shoemaker's psychological-functional continuity account of personal identity. In IV.3, I discuss Olson's criticism of Shoemaker's psychological-functional continuity account of personal identity. In IV.4, I explain Hume's account of the reification of perceptions and the impersonal description of the self. I will show that these accounts lead Hume to difficulty. In IV.5, I claim that Hume's functionalistic account of personal identity is circular.

IV.1: Hume on Character Traits, Self-Concerns, and Personal Identity.

Hume makes a famous distinction between "personal identity, as it regards our thought or imagination, and as it regards our passions or the concern we take in ourselves" (THN, 253). The self as it regards our imagination is nothing but a collection of ontologically different perceptions bundled together. Accordingly, there is no an underlying substance in which these perceptions inhere. There is no a single impression of the self surviving the whole course of someone's life. Hume then concludes that personal identity is a fiction of our imagination. His treatment of the problem of personal identity in Book I of the *Treatise* is, indeed, an explanation of our belief in that fictional self. However, towards the end of the section of "Of Personal Identity", Hume states that a present self who is not properly identical with a past self or future self may share the same concerns, characters, and interests. Hume argues that "in this view our identity with regard to the passions serves to corroborate that with regard to the imagination, by the making our distant perceptions influence each other, and by giving us a present concern for our past or future pains and pleasures" (THN, 261). There Hume claims that "I cannot compare the soul more properly to any thing than to a republic or commonwealth, in which the several members are united by the reciprocal ties of government and subordination, and

give rise to other persons, who propagate the same republic in the incessant changes of its parts” (THN, 261). He then concludes that “as the same individual republic may not only change its members, but also its laws and constitutions; in like manner the same person may vary his *character* and dispositions, as well as his impressions and ideas, without losing his identity” (THN, 261: emphasis added). Hence, “whatever changes he endures, his several parts are still connected by the relation of *causation*” (THN, 261: emphasis added).

The text indicates that there is a certain kind of *continuity* in a person’s mind in the sense that the person’s identity is bound up with the kinds of concerns, and interest that she takes for her past or future pains and pleasures. This continuity obviously does not require the mind itself to be a substance in order to remain the same over time, but it requires some kind of *causal connection* between the perceptions that make one’s mind at different times. Hume argues, for instance, that when the ideas of pains and pleasures return upon the mind they affect it by producing “new impressions of desire and aversion, hope and fear” (THN, 8). Or when one anticipates the possibility of some painful event it produces fear and trembling in the mind (THN, 444). The existence of this kind of causal connection provides the ground for a sense of person’s identity that is bound up with the kinds of values, character traits, interests, concerns, and projects that the person is associated with in her entire life. This is to say, character traits, self-concern, and values contribute substantially to the continuity of the bundles of perceptions. Pitson, for instance, notes that “character is central to our sense of self—of our remaining the same sort of self or person over time and so, in this sense, retaining an identity” (1). The self extends its concerns to its past and future, and takes interests in the kind of character traits that it intends to associate with or disassociate from (2). Similarly, Jane McIntyre has argued that Hume’s theory of the self as it regards passions could be read as his response to the problem of self-concern and character traits. “On Hume’s view, my past encompasses those perceptions, thoughts, and actions related to a present self by resemblance and causation,” McIntyre argues, “what Hume’s account of the passions needs to explain is how those past perceptions, thoughts, and actions affect my present feelings, and therefore why they are of interest and importance to me. Likewise, his account must explain why I sometimes act out of concern for a future self—a future

collection of perceptions—that will bear to me, now, the relation I bear to my past” (3). She claims that Hume’s theory of the self in Book I of the *Treatise* is incomplete, and it finds completion only after he gives an account of self-concern in Book II. Now the question that arises here is that what constitutes Hume’s account of character traits and self-concern?

Hume links the notion of character to a wide variety of different mental qualities and principles. In particular, he connects these qualities and principles to a person’s categories of vice and virtue, personal character, social character or group (national) character (4). Among the categories of a person’s vice and virtue Hume includes the character traits of reputation and name (THN, 316), the character of eloquence (THN, 322), the character of judgment and veracity (EHU, 123), the character of companionable virtues of good manners and wit (EPM, 280: EPM, 262), the character of eloquence and sound reasoning (EPM, 263), the character of the possession of delicacy of taste and sentiment to discern beauty and deformity in objects (EPM, 260), and the character of a person’s temperament and emotional propensities (EPM, 250). Examples of personal character include curiosity of the mind (EMPL, 300), avarice, a universal vicious passion (EMPL, 113), ambition, friendship, vanity, and generosity (EHU, 83). Social character includes those traits of character like national character (EMPL, 197). Hume thinks that certain qualities of characters are more associated with some nations rather than with others. These characters are determined by moral and physical causes such as, for example, the form of the government, economy and its geographic climate (EMPL, 202-204). These wide varieties of character, on Hume’s account, refer to certain mental qualities and principles that make a person the kind of person she is.

A fundamental feature of these qualities and principles is their continuity (durability), as compared to the nature of Hume’s self as a momentary thing. These character traits are relatively continuous and permanent qualities of the mind, whereas the bundle of perceptions is said to be a transient and perishing entity. “’Tis not the present sensation alone or momentary pain or pleasure,” Hume claims, “which determines the character of any passion, but the whole bent or tendency of it from the beginning to the end” (THN 381). What character traits and self-concern do for Hume is this: It provides a causal link between the self and its various past, or future, perceptions, thoughts, and

actions in such a way that we may say that those thoughts, actions, and perceptions belong to her. On different occasions, Hume links character traits to a person's motives, dispositions, intentions, and mental qualities. Accordingly, a person's motives and intentions causally link its present state of mind to its past or future actions, and thoughts. Hume writes that "an intention shows certain qualities, which remaining after the action is perform'd, connect it with the person, and facilitate the transition of ideas from one to the other" (THN, 349). These character traits are "mental causes in the form of recurrent perceptions which belong to the bundles or systems of perceptions in which our minds consist, and they play a crucial role in providing continuities among these perceptions which contribute to the sense of self which most of us have" (5).

Hume's explanation of linking a person's characters and concerns with its past or future states is embedded in his classification of them with the category of the passions, particularly with his account of the indirect passions of pride and humility (6). In general, Hume, believes that the good and bad qualities of our actions and manners naturally produce pain or pleasure in the *subjects* that observe them (THN, 285). This occurs either *directly* or *indirectly*. The direct passions "arise immediately from good or evil, from pain or pleasure" (THN, 276) while the indirect passions—pride and humility, and love and hatred—depend on the existence of other additional qualities. When someone feels a sense of pride, for instance, by being related to a good deed, it reintroduces the idea of the self into the mind. Similarly, when someone feels a sense of humility it also reintroduces the idea of the self. Hume argues that these indirect passions have the self as their object, and that is why they reintroduce the idea of the self into the mind. He offers an underlying mechanism to explain this (THN, 286). He calls it the 'double relation of ideas and impressions' (THN, 286). Hume argues that we must first distinguish between the cause and the object of these indirect passions. He claims that the self cannot be the cause of pride or humility, but only their objects (THN, 277). Whereas the cause of pride is the subject in which certain qualities are found. For example, a beautiful house is the subject of my pride because the quality of beauty is the cause of the pride of which I am the object. In general, Hume tells us that the natural or immediate causes of pride and humility are certain 'qualities of either mind or body' (THN, 303) such as one's wit, courage, integrity, beauty, strength, and agility (THN, 279), which belongs to the self.

Hume, however, argues that the objects of these passions and their causes are closely related. In fact, he says that ‘the true system’ that strikes us ‘with an irresistible evidence or force’ works something like this: “that cause [the house], which excites the passion, is related to the object [the self], which nature has attributed to the passion; the sensation, which the cause separately produces, is related to the sensation of the passion: From this double relation of ideas and impressions, the passion is deriv’d. The one idea is easily converted into its correlative; and the one impression into that, which resembles and corresponds to it: With how much greater facility must this transition be made, where these movements mutually assist each other, and the mind receives a double impulse from the relations both of its impressions and ideas?” (THN, 286). This is a very abstruse passage. It is in need of some elaboration.

According to this passage, the cause of my feeling of pride, say, in my beautiful house, arises when the subject (the house) and a quality (beauty in this case) together produce an idea, *A*. The subject (the house) brings to the mind, by the association of ideas, the idea of *myself*, or idea of *B*. In fact, the subject brings to the mind the idea of myself because it is always the object of pride, but Hume says that the idea of *A* also gives rise to impression of *A*, a distinct pleasure that I take in my having a beautiful house. This pleasure, impression *A*, by the association of impressions, brings to mind the resembling pleasure of pride, say, impression *B*. In conclusion, idea *B* (idea of myself) is related to impression *B* (the idea of the pleasure of pride) as the object of this impression (7). An outline of Hume’s complex argument goes like this: “(1) The beautiful house produces in me an impression and then an idea of itself. (2) This idea, because it is an idea of my house, also gives rise to a related idea, the idea of myself. (3) At the same time, this idea of a beautiful house produces in me an impression, a feeling of pleasure. (4) This first impression gives rise to a second impression to which the first is related by resemblance...(5) This second impression of pleasure has the second idea (the idea of myself) as its object and is the impression or feeling that constitutes the core of the passion of pride. In short, by means of a double relation of impressions and ideas, the pleasant feeling associated with pride is itself associated with an idea of the constant object of pride, namely, the self ” (8).

This mechanism, or system, as Hume says, provides a causal basis for the continuity, or durability (to use Hume's own terminology) of persons' motives, intentions, concerns, dispositions, interests, etc. What is important about this mechanism of explanation is that "the self which provides the object of the indirect passions of pride and humility is a relatively *enduring* self" (9). It is in light of this account of the indirect passions that Hume is able to explain why one's past or future actions continue to concern her, despite the fact that the self is not a single substance or strictly identical through time. The self extends its concerns, dispositions, interest, etc., to the future as well as the past by establishing a causal relation between its intentions and actions. Typically, "even the most ordinary plans can require that our present actions are governed by a future interest that we see as our own. We think of ourselves in the future, and are motivated (at least sometimes) by the prospect of future pain or pleasure, punishment or reward" (10). Hume argues that there is a causal link or necessity between intentions and actions, and that such necessity "has universally, tho' tacitly, in the schools, in the pulpit, and in common life, been allow'd to belong to the will of man, and no one has ever pretended to deny, that we can draw inferences concerning human actions, and that those inferences are founded on the experienc'd union of like actions with like motives and circumstances" (THN, 409). Thus, the present self (Hume's momentary self) is causally related to its past or future by the kind of concerns and interest that it takes about its actions and characters via its motives and intentions.

I suggest that insofar as Hume's account of character traits and self-concerns are so central to his conception of personal identity—in implying the continuity of ourselves over time—indicates that his theory of the self as a bundle or collection of different perceptions "do, after all, possess a certain kind of *structure*" (11). The bundle may not be *strictly* unified, but it can have a certain kind of continuity over time. It seems to me that the structure that unites the bundle over time is a functional continuity. Pitson remarks that "an immediate connection between myself now (or in the past) and my future self is established by the fact that my motives and intentions—which obviously include ones that have previously been formed—function as causes of my actions. It is the existence of this kind of relation that makes them *my* actions, part of the continuing narrative in which I as a person consist along with their mental causes. ...This makes a vital connection

between the actions one performs at present and oneself in the future. This aspect of self-concern could be expressed in terms of an identification with a self in the future: an identification which essentially depends upon the operation of passions” (12).

Having examined the issues of character traits, self-interests, and self-concerns, I want to draw some conclusions now. First, we have seen that, on Hume’s account, the notion of character traits is closely connected to such aspects of persons as their motives, dispositions, concerns, interests, mental qualities, and principles of mind. And as I explained, a distinctive feature of such character traits is their durability or continuity over time. In this sense, character traits give a certain kind of continuity to the bundle of perceptions in which selves consist. Therefore, character traits are vital to our remaining the same sort of person over time, and thus retaining an identity. Second, Hume thinks that when one thinks about oneself in the future or past, one thinks about the actions that follow from her intentions, motives, and characters. In this respect, there is a relatively significant mental continuity between different perceptions which make up Hume’s bundle theory of the self. In particular, this sort of mental continuity consists in causal connection between a person’s intentions, characters, concerns, etc. and her actions over time. In a nutshell, there are various kinds of causal connections among the mental states that make up a person’s mind at different times and it is the existence of these connections that leads us to regard that person as remaining the same. Of course, this kind of continuity does not require either that the person should have a continuing consciousness (or memory) of each of her past mental states or that the mind itself should be some kind of substance. Nor even such causal connections require that one should continue to have the same mental dispositions, concerns, interests, etc. that are typically distinctive of one’s as a person. Provided that, whatever changes of these different kinds of dispositions, traits, characters, etc. the mind or self undergoes, its different parts “are still connected by the relation of causation” (TH, 261). Accordingly, it is by this sort of causal connection in the form of recurrent perceptions belonging to the same bundle of perceptions that gives us a sense of moral responsibility (agency) for our actions insofar as they proceed from certain causes in our characters.

A central idea of this chapter is that Hume wants his account of character traits, self-concerns, and self-interest to be viewed as the basis for his claim that the

functionalist theory of mind implies personal identity. In fact, I would like to argue that the Humean idea that a present self extends its concerns and interest to a future self or a past self is very similar to Shoemaker's functionalistic notion of the self as a self-perpetuating substance. In what follows, I should explore this important topic.

IV.2: Functionalism and the Psychological Continuity Account of Personal Identity

Strictly speaking, functionalism is a theory about the nature of mental states such as desire, feeling, and belief. More specifically, it defines mental states in terms of their causal relations to one another and to sensory inputs and behavioural outputs. Shoemaker uses functionalism in this general sense. In particular, there are two interrelated characteristics that are important for him. First, mental states are constituted by their causal roles, so that "it is of the essence of a mental state to be caused in certain ways and to produce, in conjunction with other mental states, certain effects (behaviour or other mental states)" (13). Second, since functionalism characterizes mental states solely in terms of causal relations to other states, it does not specify the "intrinsic" nature of such states. This refers to the abstract nature of mental states. The abstractness of functional characterization of mental states implies that such mental states could be realized either physically or non-physically. If it is realized physically, materialism is true. If it is realized non-physically, then a mind-body dualist agrees with functional characterization of mental states without being committed to materialism. Moreover, such abstractness "allows that the same functional state could be realized in a variety of different ways... in very different physical states in creatures of different species (humans and dolphins, or, more radically, humans and Martians)" (14).

Gary Fuller has argued that the "psychological continuity accounts of personal identity can be described in terms of three dimensions: the range of continuity, the amount of continuity within the range, and the degree of abstractness of the appropriate causal connections" (15). The traditional Lockean memory account of personal identity is based on a narrow range of continuity. Shoemaker's psychological continuity account appeals to a broader range of continuity; it includes not only memory continuity, but also continuity of "interests, tastes, talents, and traits of personality character" (16). He also maintains that a significant amount of continuity is required. He allows that a person may survive certain kinds of amnesia, the total and irreversible loss of almost all memory

events, but not a total “brain zap,” a total destruction of personal traits and continuity. Moreover, Shoemaker uses the notion of causal connection on a very abstract level. In the case of brain-state transfer procedure, BST-procedure, in which the total state of a brain is transferred to a new brain, he argues that BST-procedure is person-preserving. In fact, this means that the notion of continuity that is involved here does not require the persistence of the body or brain or “of any bodily organ, or of any matter at all” (17).

These features make Shoemaker’s version of psychological-functional continuity account of personal identity clearer. His account involves a broad psychological continuity that is related by almost any causal mechanism. Shoemaker thus takes functionalism to imply that a “reasonable amount of broad psychological continuity with any cause is necessary and sufficient for personal identity” (18). This is the core idea behind the psychological-functional continuity account of personal identity, which is endorsed by Shoemaker. The urgent question now is this: What does constitute the necessary and sufficient conditions for such psychological-functional continuity criterion of personal identity? Let me explain this.

At the core of Shoemaker’s psychological continuity account of personal identity is what he calls a *functionally appropriate successor state* (FAS) to a psychological state. “Psychological continuity,” he argues, “is just the playing out over time of the functional natures of the mental states characteristic of persons. To the extent that it consists in psychological similarity between different person-stages, this is due to the fact in many cases what is required as the successor state of a mental state is just another token of the same state” (19). He goes on to claim that “to the extent that it consists in ‘memory-continuity,’ this is because it belongs to the nature of certain states (sense experiences and intentional actions) that they give rise to successor states of the sort I have called memories from inside, and because it belongs to the nature of these to perpetuate themselves, i.e., to produce successor states having the same or closely related contents” (20). These passages capture Shoemaker’s main concept of the psychological-functional continuity account of personal identity. It requires further elaboration.

As said above, the functionalist theory of mind defines mental states in terms of their causal relations to sensory inputs, behavioural outputs, and other mental states. Indeed, the individuation of such mental states is constituted in terms of their

causal/functional roles in a complex causal network of relations and states. Let's say, for example, that my belief that it is very cold now is brought about by certain sensory inputs, or sensory-experiences. It happened that I hold another belief that a hot cup of coffee keeps me warm. Combine these two beliefs with certain desires, namely, the desire to keep warm. According to functionalism, this leads to some behavioural outputs, i.e., taking a cup of coffee when I feel cold. Obviously, this is an oversimplification of certain functional states. Of course, in reality the causal network is much more complex and the functional roles that are played are brought about as a result of massive causal relations among mental states. However, this example clearly shows that there are no other states that count as the belief that one needs a cup of coffee on a cold day unless that belief satisfies this characterization. To believe that on such a cold day one needs a hot cup of coffee, on the functional account, is a state that is caused in this way and which has these kinds of effects when it is combined with such other states. In other words, it is central to the functional nature of mental states that they bring about their effects in conjunction with other functional states. It also shows that functionalism defines each mental state in terms of its relation with other mental states.

According to functionalism, the functional characterizations of mental states are not limited to oversimplified versions of certain states within that complex causal network of relations and states. In fact, functionalism assumes that our whole belief-making system functions on the basis of this kind of characterization of our mental states. If this is true, which seems to be the case, then our beliefs lead us to another set of beliefs which they entail or support, and that one, in turn, leads us to some other sets of beliefs, desires, which eventually lead us to make certain decisions or form certain intentions. This is a very general and vague characterization of what really occurs in the conduct of our daily lives, which require reasoning and deliberations, as well as making decisions, and forming intentions. But what are the implications of these for the psychological-functional continuity account of personal identity? Shoemaker argues that since certain functional states extend over time, or as he himself put it, "play out over time" it follows that those functional states guarantee, and satisfy, the requirements of the continuity criterion of personal identity. "I form the intention to do something tomorrow," he states, "and when tomorrow comes I do it, from that intention. But while there is a sense in

which I retained the same intention throughout, there is also a sense in which the content of my intention was constantly changing; it began as the intention to do something twenty-four hours hence, evolved into the intention to do it twenty-three hours hence, and eventually became the intention to do it *now*" (21). This is the central idea of the psychological-functional continuity account of personal identity.

At the heart of the psychological-functional continuity account of personal identity lies the notion of 'mental content.' The claim is that there is always a kind of continuity in the contents of a person's mental states that ensures the dependence of her mental states in a later time on her contents at an earlier time. Shoemaker, for instance, points out that "[w]hile the content of a person's mental state (belief, intention, etc.) will to a certain extent depend on the nature of all of his states at earlier times, there will often be a particular state at an earlier time on whose content its content especially depends—and it is of that state that we will call it the 'successor state'" (22). On this account, the continuity of a person's content of mental states is understood to be analogous to what Quine calls 'eternal sentences.' Quine defines 'eternal sentences' as sentences that their "truth value stays fixed through time and from speaker to speaker" (23). Similarly, Shoemaker argues that in the same way that the truth-values of an eternal sentence, e.g., there is no highest prime number, is constant and unchanging over time and is no subject to the speaker's viewpoint, to a certain extent the content of one's belief that there is no highest prime number is fixed through time and is not subject to change. As a result, the content of one's mental states in a successor time is the same as one's content in an earlier time. In fact, the idea is that the content of one's mental states in a successor time is a *token* of the same belief that one had in an earlier time.

Shoemaker's psychological-functional continuity and connectedness (C&C) view of the self is based on a unique notion of substance. He places himself in the tradition of those who hold that the self is a substance, but at the same time he is reductionist about the self. Shoemaker thinks that it is possible to combine some key intuitions of the substantial view of the self with some central intuitions of the reductionist view of the self. According to him, the traditional conception of substance depicts substances as ontologically *independent* entities in ways in which other entities are not (24). "Modes" and "affections" or states are all entities that whose existence is logically dependent

(parasitic) on the entities that they are modes, affections, and states of. Hence, entities that are ontologically independent, that is, on which other entities are dependent, and that are not dependent on other entities themselves, are individual substances. The basic idea is that an individual substance does not require the existence of any particular mode, affection, or state to exist. However, it does not mean that an individual substance exists without having any modes, affections and states at all. On the contrary, the modes, affections, and states depend on particular individual substances for their existences. In this sense, ontological independence is a kind of *criterion* of substantiality.

Shoemaker further argues that the independence criterion of substantiality is related to the notion of individual substance as a *subject of properties*. He makes a distinction between two different senses of substance, i.e., the “subject of properties sense” and the “parcel of stuff sense” (25). In the first sense, while the continuants, e.g., tree, bodies, minds, undergo changes, they can survive changes in their compositions and that count as substance. By contrast, substance in the parcel of stuff sense strictly requires the continuity of the portions of the stuff (material or immaterial) that constitute the object. Substance in this sense is thought of as a “quantity” of some material or immaterial stuff that is composed of the same stuff at every moment of its existence (26). Specifically, when Shoemaker claims that a self is a substance in the subject of properties sense, he means that if and only if the two following conditions are met: “(1) statements of the form “S thinks, experiences, etc. such and such” are sometimes true, and (2) such statements are not analyzable in a certain way” (27). To be analyzable means that personal identity could be analyzed in terms of some sort of psychological continuity and connectedness. To say that, for example, S experiences pain is to say that pain is adjectival on S. And to say that pain is adjectival on S is to say that pain in any way is logically dependent on it. This suggests that the psychological-functional relation between mental particulars and mental subject is more than what the grammar or gerundial phrases designate (more on this later). Hume’s bundle theory fails to meet this requirement. The analysis that would fail to consent to (2) would be the kind of analysis “whose *analysans* does not refer to or quantify over persons or subjects of mental properties” (28). Of course, Hume’s bundle theory of the self or Parfit’s impersonal description of personal identity analyse persons in a way that make no commitment to

mental subjects as constituents of the world. These conditions are not required in the parcel of stuff sense. The implication of the conception of individual substances as subject of properties for the independent criterion of substantiality is that minds or selves will be ontologically independent substances relative to mental particulars, such as thoughts, sensations, etc. In other words, these mental particulars, that are modes and affections of, are adjectival on minds or selves (29). Mental particulars are dependent on minds or selves as the subjects that experiences them, or have them. Shoemaker thinks that this can qualify minds or selves as substances.

Shoemaker thinks that a notion of causation involves in the independence criterion of substantiality. This has to do with a “feature of the ‘continuants’ that are paradigm individual substances that their persistence through time involves there being relations of causal and counterfactual dependence of their properties at later times on their properties at earlier times” (30). For example, this particular plant in the garden would not have the shape that it has now, if it did not have that shape a day ago, or a week ago. In part, the shape that it has now is because of the latter. In other words, the earlier state of an object plays a causal role in its later existence. In particular, Shoemaker holds that the causation involved here is what W. E Johnson has called “immanent” causation, that is, the causation that is internal to the thing’s career. It contrasts with the “transeunt” causation involved when one object acts on another one (31). It is by virtue of this sort of immanent causation that plants and animals, as the paradigm cases of individual substances, retain their properties over time while undergoing the changes that are characteristic of their nature. Individual substances in this sense are relatively autonomous self-perpetuators (32). That is, these individual substances, e.g., trees, mountains, etc., autonomously perpetuate themselves by retaining their same properties over time. The inanimate objects retain their properties in a less dynamic manner, while the animate objects continue to exist in a more dynamic way by continuing to be the kind of things that they are. In the case of the self, the dynamic of self-perpetuation is much more complex and intricate. Likewise, it is also plausible to say that minds or persons are individual substance as long as they live up to this degree of independence and autonomous self-perpetuation. In all these, the notion of immanent causation plays the key role in this sort of autonomous self-perpetuation. By and large, it is quite plausible

that “it is to the extent that something is viewed as an autonomous self-perpetuator that we find it natural to regard it as an individual substance” (33)

Shoemaker applies the independence criterion of substantiality to the case of personal identity. First, he argues that persons or selves are independent entities, as the subject of experience, and that mental particulars are adjectival on selves *qua* mental subjects. Secondly, he claims that selves or minds are relatively autonomous self-perpetuators. In regard to the first point, Shoemaker argues that the claim that mental particulars are adjectival on mental subjects must go beyond simple grammatical grounds. We cannot say that mental particulars are adjectival on mental subjects simply because such mental particulars are designated by gerundial phrases. At the core of the dependent (adjectival) status of mental particulars on mental subject is the idea that mental particulars are *logically*, or causally, dependent on the subject of experience. So, strictly speaking, adjectival dependence means only logical dependence. Logical dependence means that there is a relation between the subject of experience and the experience itself, and that this relation in certain ways goes beyond superficial grammar. We have two distinct entities with independent ontological status. The view that mental particulars are adjectival on, or logically dependent upon, selves *qua* mental subjects is the opposite of the sense-datum theory of perception, or the ‘act-object’ theory of perception.

Shoemaker rejects both the so-called ‘act-object’ theory of perception and the sense-datum theory of conception. According to these theories of perception, mental particulars do enjoy an ontological status. That is, these theories typically hold that mental particulars or perceptions being related to their bearers in a certain way, but at the same time they deny that the relation between mental particulars and their subjects amounts to logical dependency of the former on the latter. Hence, in Shoemaker’s view, proponent of the act-object theory, which is a main source of bundle theories, “drops the dependence claim and attempts to regard these mental particulars as the mental building blocks out of which the mind is built” (34). On bundle theory, we may allow that “S’s experiencing pain at t” designates a mental particular that is adjectival on S. However, it is important to note that, on such an account, the designation will be something like: the inclusion of a perception of pain in S (a certain bundle of perceptions). Clearly, on

Hume's account, a perception of pain does not logically depend on the bundle *qua* mental subject. On Shoemaker's account of the self, there is a logical relation (dependency) between experiencing and experiencer, and between perceiving and perceiver, whereas, on the bundle account, perceptions are not logically dependent on the perceiver.

The logical dependence of mental particulars on the mental subject has a close relation with functionalism. Functionalism claims that the identity of a mental state is determined by, or depends upon, the contents of other mental states. One cannot have the belief that the cold war is over, or that the United States is the most powerful country in the world, without believing a vast number of other things. Hence, "the identity of beliefs is partly determined by their inferential connections—what beliefs they tend to give rise to when combined with other beliefs. The identity of mental states generally is partly determined by the ways they combine with other states to influence behaviour...and to generate other mental states" (35). The important point is that the existence of a mental state of a certain kind "brings with it the truth of a vast number of conditional propositions about what other states would be apt to exist, or what behaviours would be apt to occur, were that state to be combined with—were it to be coinstantiated with—other states of certain kinds" (36). In short, the adjectival dependence of mental particulars on mental subjects goes beyond mere superficial grammatical dependence; it includes logical dependence of one on the other. The logical dependence in this sense satisfies the independence criterion of substantiality of the self.

The second independent criterion of substantiality, i.e., the autonomous self-perpetuation of individual substances, could be applied to the case of personal identity as well. Again, the idea is that the kind of psychological continuity view of personal identity that the functionalist conception of mental states implies is compatible with the notion of individual substances as relatively autonomous self-perpetuators. According to the functionalist account of mental states, one important functional role of a particular mental state is to give rise to other mental states. It is typical of the mind that when it wants to trigger a particular mental state, it involves a set of other mental states; it includes not only conscious mental states, but also all, or part, of the desires, preferences, intentions, beliefs, hopes, etc. constituting the whole bundle (system) of the mind. For example, a functional account of mental states tells us that the belief that it is raining is accompanied

with the desire to stay dry, and that in part is related to another belief that umbrellas can keep us dry. Yet these mental states produce the decision to use an umbrella when it is raining. Sometimes happenings of these kinds involve conscious exercise of agency, but in other times it does not involve such deliberate act. “The “cognitive dynamics” of mental states is such that over time they change in certain ways,” Shoemaker says, “depending on what other mental states accompany them. An expectation of something as being in the remote future evolves into an expectation of something immediately forthcoming, given normal awareness of the passage of time” (37). It is this kind of cognitive dynamics in which mental states give rise to other mental states in a later time, and that the notion of a self that starts at a particular time with a certain set of mental states (bundles) is expected to continue over time. The important point is that “the later stages of the series are the consequence of the earlier stages in it playing the functional roles that are constitutive of their being the kind of mental states they are” (38). When we apply this to the case of personal identity, we can say that these stages belong and constitute the history of an individual person, or an individual mental subject. That history is both necessary and sufficient for calling a mental subject a substance, that is, an autonomous self-perpetuator, or person-preserving. In conclusion, the notion of a self or mind is somehow compatible with the idea that the self is a substance, and that such a notion of the self as a substance, in the sense explained above (adjectival dependency and autonomous self-perpetuator), is consistent with the psychological continuity account of personal identity.

I think that Shoemaker’s suggestion that functionalism implies a version of the psychological account of personal identity is the right theory. In particular, functionalism implies that functional continuity is both necessary and sufficient for personal identity. However, Olson rejects the claim that functional continuity is both necessary and sufficient for our identity over time. Indeed, he argues that functionalism does not tell us anything about personal identity. In the following section, I shall present Olson’s claim that functionalism implies nothing about personal identity.

IV.3: Functionalism Tells us Nothing about Personal Identity

Olson has rejected Shoemaker’s psychological-functional continuity criterion of personal identity. In particular, He argues against the thesis that the functional nature of mental

state, in conjunction with other mental states of the same person, gives rise to appropriate successor states affecting one's future behaviours. Olson, who is a self-proclaimed *Animalist*—a version of the Physical Continuity Criterion of personal identity—has claimed that the kind of organism that we are, or the kind of thinking human animals that we are, provides both necessary and sufficient conditions for our continuity over time. More precisely, what the psychological continuity criterion of personal identity says about our persistence over time is not true of human organisms: no version of psychological continuity is either necessary or sufficient for the kind of human animals that we are. Olson thinks that the identity of human animals does not consist in mental continuity. In a sense, he seems to be holding that what matters in our survival is the continuity of the sort of organism that we are. He explicitly claims that “[t]he fact that each human animal starts out as an unthinking embryo and may end up as an unthinking vegetable shows that no sort of mental continuity is necessary for it to persist” (39). It is not necessary because human embryos and human vegetables have no psychological properties, and we as beings with psychological properties cannot be continuous with organisms that have no psychology at all. Hence, we as animal organisms can survive without having any psychological continuity. Therefore, psychological continuity is not *necessary* for our identity.

Furthermore, he argues that hypothetical cases like that of cerebrum transplantation, or fission cases, indicate that no sort of mental continuity is *sufficient* for identity either. In fission cases, when the surgeon transplants each of one your hemisphere into two different empty-heads, it results in two numerically distinct persons each one of them psychologically continuous with you. In fact, in this situation we have two different persons each claiming to be you. This shows that psychological continuity is not sufficient for personal identity because it implies that any future person who is psychologically continuous with you must be you. But you cannot be identical with two numerically distinct persons because one thing cannot be numerically identical with two things. For this reason, the fission problem is one source of worry for the psychological-functional approach. However, Olson's rejection of the sufficiency argument for the psychological-functional continuity account of personal identity heavily relies on, what he calls, *the problem of the thinking animal*, or the two-many-minds problem. Let me

explain this.

Olson's strategy is to reduce the psychological continuity theory of personal identity, and the claim that functional continuity is sufficient for the identity condition, into a *reductio ad absurdum*. According to him, when we detach a cerebrum, for instance, from the body what, indeed, happens is no more than removing and then transplanting it to another body, exactly like when we remove and then transplant a liver from one organism to another. The empty-headed body left behind is also an organ. It is merely the transference of an organ from an animal organism to another. Since the psychological continuity account claims that when one's cerebrum is transplanted its psychological properties go along with its transplanted cerebrum, and leaves behind its animal organism, it says that one is "a thing and another thing." That is to say, the psychological continuity account ends up embracing the existence of two sorts of entities: a psychological person and a human animal organism (40). Essentially, Olson thinks that even the empty-headed human animal organism that is left behind can have a mental life, or think, or be conscious (he considers having propositional attitudes as thinking), though lacks psychological persistence conditions. As a result, "people will come in two kinds: "animal" people who are animals and have the persistence conditions of animals, and "psychological" people, who aren't animals and have psychological persistence conditions. There would be two such people, one of each kind, wherever we thought there was just one. There would be both an animal person and a psychological person numerically different from it sitting in your chair and reading this" (41). Hence, it seems that there are two sorts of conscious, or intelligent, beings: a person who is not an animal organism, and a human animal organism that is not a person. Olson concludes that it is useless to deny that human animals are people. Therefore, if human animals can have mental life, or can think, it implies that there can be rational, intelligent beings with non-psychological identity conditions. This he takes it to be a *reductio ad absurdum* of the psychological continuity account of personal identity.

This *reductio ad absurdum* is entailed by the familiar "too-many-thinkers" or "too-many-minds" problem, or as Olson calls the *problem of the thinking animal*. The trouble with the too-many-thinkers problem is that it claims that human animals seem to think, or have consciousness. This, in turn, is inconsistent with the psychological

continuity view of personal identity according to which the persistence conditions lie exclusively in mental continuity. More precisely, the problem of too-many-minds claims that if persons and their coincident biological organisms share the same physical properties, and given that mental properties somehow supervene on physical properties, it follows that the biological organism will share the mental properties of the person. If the biological organism shares the mental properties of the person, it must count as a person as well. And when the animal is thinking, it seems that it thinks your thoughts, and has every experience and mental properties of yours. Therefore, it is surely you. Let's assume that you are not that animal. The animal thinks, and you think too. And since you cannot be that thinking animal, it follows that there would be two beings thinking your thoughts: a thinking animal and you, a thinking non-animal. This would mean that there are two persons where there is supposed to be only one; for every thought there will be two thinkers, which are numerically different. In short, we end up having two sorts of mental existences, namely, an animal person (human organism) and a psychological person, each having mental life of its own. Obviously, this looks as an inconsistent position.

In a nutshell, Olson claims that there are only "three alternatives to our being an animal: (1) there is no human animal where you are; (2) there is an animal there, but it doesn't think in the way that you do; or (3) there is an animal there, and it thinks exactly as you do, but you are not it" (42). With regard to the existence of animals, he assumes that no one would deny the existence of animals because its denial would rule out the existence of most of the things we might be. Concerning the second alternative, there are two possibilities: either the human animals do not think at all or that they think but not the way we do. Olson goes on to argue that the first possibility amounts to the endorsement of eliminative materialism, obviously a position that the proponents of the psychological continuity account would reject altogether. The second possibility says that human animals have mental properties but different from us, for instance, they are conscious but never self-conscious (43). Obviously, proponents of the psychological continuity account of the self can argue that human animals, the organism that we are, cannot think, or can be conscious. Shoemaker, for instance, has argued that organisms are unable to think because they do not have the right sort of mental properties. For Shoemaker, mental properties have characteristic causal roles, and that is something that

physical properties characteristically lack. According to him, whatever thinks, or is conscious, must persist by virtue of some sort of psychological or mental continuity. Hence, the psychological continuity view rules out the fact that persons are material entities (44). In any case, Olson maintains that Shoemaker's response does not meet the problem of animal-thinking. The final alternative, which holds that there is an animal and thinks exactly as you do but it is not you, ends up facing the too-many-thinkers problem. He finds all these alternatives unconvincing, and thus concludes that we are that thinking animals.

If Olson is right by any account, the result of cerebrum transplantation shows that there are two people thinking what you think right now. Strictly speaking, it claims that there are two rational, intelligent but numerically different beings, namely, our human animal, with non-psychological continuity condition, and your psychological person, with psychological identity condition, sitting in your chair and reading this dissertation now. This shows that the psychological continuity account leads to the problem of "too-many-minds." And this problem says that there are two persons sitting in you each having mental properties of its own, but somehow psychological or mentally continuous with you. Obviously, Olson thinks that this is absurd. But more importantly, he thinks that this proves that psychological-functional continuity account fails to provide sufficient criterion for personal identity over time. Because the "too-many-minds" problem shows that you are psychologically continuous with two persons, and one thing cannot be the same or continuous with two numerically distinct things.

So far I have tried to show that Olson argues that the psychological-functional continuity account fails to provide both necessary and sufficient conditions for the identity of persons over time. The fact that we all begin our lives as unthinking embryos and may end up in a vegetative state shows that mental continuity is not necessary for our identity. And the fact that in fission, for instance, you may end up being psychologically continuous with each one of your offshoots, indicates that psychological-functional continuity is not sufficient for your identity over time. The cerebrum transplantation or the problem of the thinking animal further shows that the psychological-functional account fails to provide sufficient condition of your identity. Olson, however, argues that the problem with Shoemaker's account runs deeper than this. I will explain this.

Olson believes that Shoemaker's *causal-powers argument* will have some surprising consequences: (a) no living organism could have mental properties; (b) no mental properties would supervene even in a weak sense on a thing's microstructure and surroundings; and (c) it rules out our being material things ⁽⁴⁵⁾. "The psychological-continuity view stands or falls with the claim that human animals can't think," Olson claims, "[i]f they can, then we are those animals, and thus have non-psychological identity conditions. If we aren't animals, that can only be because human animals can't think. And no other account even begins to explain why human animals can't think" ⁽⁴⁶⁾. This, of course, not only rules out our being material things, but more importantly, it implies that material beings like us cannot think at all. This poses a problem for those who think that, including the proponents of the psychological continuity account, we as material beings can think but the human animals cannot. This conclusion is inconsistent with Shoemaker's causal-powers argument, which purports to explain the conditions of our identity in terms of a materialistic view of the mind. On the one hand, the causal-powers argument implies that no animal organism can think, and on the other, it supposes that other material things like us can think.

Olson claims that the conclusion that no organism can think undermines Shoemaker's functionalist account of personal identity in certain specific ways. He argues that according to Shoemaker's version of psychological continuity account it is impossible for persons to have physical properties at all, and therefore, it is incompatible with materialism, a position that Shoemaker adheres to. In other words, the problem that it creates is that it implies that "none of a thing's mental properties supervene on its microstructure together with its spatio-temporal surroundings in even the very weak sense that things with the same microstructure and surroundings must be the same with respect to those mental properties" ⁽⁴⁷⁾. However, as Olson points out, functionalism does tell us that mental properties do supervene, at least in a weak sense, on the thing's physical microstructure and its spatio-temporal surroundings. Olson's argument is simple. He argues that "having the right microstructure and surroundings does not even reliably *cause* a thing to have any mental properties: certain human animals have the same microstructure and surroundings as we have, yet never have any mental properties" ⁽⁴⁸⁾. Or that there are microphysical duplicates that, while identical in their surroundings

and microstructure, are completely different in their mental properties. So it cannot be that a thing's mental properties supervene on its physical microstructure and spatio-temporal surroundings. Olson believes that this is a huge blow to Shoemaker's causal-powers argument.

Furthermore, Olson claims that Shoemaker's causal-powers argument entails that no electronic digital computer or robot can think, though functionalism allows that thinking could be realized in different ways. Let's say that we have a thinking computer, which supposedly exhibits the sort of mental continuity that functionalism requires. Now, if we transfer the data, or the mental contents, from one computer to another one, it ought to create mental continuity between these two computers, supposing that there is mentality involved there. Of course, Shoemaker accepts that in the cases like that of the "brain-state-transfer machine," or BST-procedure, mentality continues between the state of the recorded brain and the state of the imposed brain, even though the bodies are changed (49). Olson argues that if we are allowed to say this, we can surely say that the data transfer from one computer to another one could create mental continuity between these computers. He concludes that "even if a computer could *produce* thought, the subject of that thought would have to be something other than the computer itself: something that could literally move from one piece of hardware to another without any matter moving" (50). Olson calls it a most dubious suggestion because it implies that in such cases mental properties do not supervene on physical properties, something that functionalism cannot rule it out. So he thinks that this conclusion is inconsistent with Shoemaker's materialist position.

Moreover, Olson thinks that Shoemaker's functionalist account of the self rules out the possibility of immaterial thinking substance. First, it is obvious that mental continuity is insufficient for an immaterial substance to persist over time. It is said that if psychological continuity is necessary and sufficient for a thinking thing to persist, then no such substance could have persisted over time. More importantly, "an immaterial substance might somehow relate to an insubstantial thinking thing that could "move" from one such substance to another in something like the way that you or I can move from one human animal to another according to the psychological-continuity view" (51). If such substance moves so freely from one thinking thing to another, the problem that it

causes is that we no longer are able to say anything about its subject of experience.

Olson is right to say that these consequences, which follow from Shoemaker's causal-powers argument, are inconsistent with the functionalist theory of mind. It is a commonly held view among the functionalist that mental states are ontologically neutral to their physical, or non-physical, makeup. It defines mental states in terms of their functions, but it remains neutral about how they are realized. It allows mental states to be realized differently: by a human brain, Martian brain, electronic digital states or by non-physical state of immaterial substance. Olson, however, argues that these surprising consequences show that Shoemaker's functionalist account "severely restricts the metaphysical nature of the *subjects* of those states, as well as implying that there must be such subjects" (52). If it is conceded that Shoemaker's causal-powers argument rules out the possibilities of immaterial thinking substances, artificial intelligence, and animal thinking, then it is plausible to say, as Olson does, that Shoemaker's version of functionalism severely restricts the metaphysical nature of the subjects of mental properties. To put it differently, on the one hand, Shoemaker's causal-powers argument claims that it is part of the nature of mental properties that they are individuated by the conditional causal powers they bestow on their bearers, and on the other hand, he thinks that animals, immaterial substances and computers can think. These two do not go comfortably together.

Olson argues that the problem runs even deeper. The reason that functionalism implies nothing about the identity conditions of thinking beings is that the very characterization of mental properties in terms of the causal powers they give their bearers "can't give a complete account of what those causal powers are without making assumptions about what it takes for their bearers to persist" (53). Let's grant that functionalism's main intuition that mental states are characterized in terms of their causal roles they play is true, but why should we believe that "the characteristic causes and effects of a thing's mental states always be states of that thing, and no other? Why should anyone who isn't already a psychological-continuity theorist accept that? For someone to be hungry is at least in part for him to be in a state typically caused by someone having low blood sugar, and apt to combine with someone's belief that there is food before him to cause someone to eat. Why *must* it be the same being all four times? Why couldn't my

being hungry cause someone else to eat in the way that it ordinarily causes me to?” (54). According to Olson, the problem runs even deeper than this. Let’s consider the *branching* cases of mental continuity. If the surgeons transplant each of my cerebral hemispheres into two different heads, the two beings are going to be mentally continuous with me at some future time. But the two beings cannot be numerically identical with one (me). This leads to a contradiction. In order to avoid the contradiction, Shoemaker and other psychological theorists call for a *non-branching* restriction, the claim that one is identical with any past or future being if and only if non-branching mental continuity is guaranteed. It is like to say that if one cerebral hemisphere is transplanted and the other one is destroyed it guarantees the person’s survival. Olson remarks that “with the possible exception of Shoemaker, no one finds it easy to believe that you survive if one future person inherits your mental features but not if two do” (55). In any case, the problem that branching creates is that we end up having two different mental subjects related to the donor in the same way. Whereas the causal-powers argument claims that mental properties related in that way must belong to the same subject, and not to two different ones. Functionalism allows, for instance, that the intentions that you have could continue without interruption in each of the two recipients and cause actions in the same way that it causes action on you. So we see a conflict here between functionalism, which allows that each of the fission offshoots of you would be mentally related to you, and the non-branching restriction clause that apt the existence of causal relations between mental states as belonging to the same subject.

This conflict seems irresolvable. On the one side, the causal-powers argument emphasizes that mental properties must be characterized in terms of the causal powers they bestow on their bearers. This means that there is a *unique* sort of causal link between mental properties and their subjects; the sort of ordinary causal links that exist between mental states in one’s brain. So it is my being hungry, in conjunction with my beliefs that food is in the fridge, for example, that causes me to eat. It is, indeed, this sort of characteristic way of relating mental states and actions to the same subject that guarantees one’s survival. On the other hand, functionalism clearly predicts that mental properties in each of your fission offshoots are causally related to you. It could be said that the state of being hungry in one of the fission offshoot in connection with the state of

belief in the other offshoot that the food is in the fridge causes it to eat. Functionalism will not rule this out. In a sense, this makes reservation for the sort of claim that my being hungry might be caused by your belief that there is food in the fridge. This conflict poses serious difficulties for Shoemaker's functionalist account of personal identity. Certainly, Olson is justified in saying that "we can't give a complete account of what those causal powers are without making assumptions about what it takes for their bearers to persist" (56). It certainly requires an assumption about the existence of an "appropriate" causal link between mental properties and their subjects. Therefore, it would be wrong to "say that any being whose [mental] states are caused in *any* way by your current mental states is you. It is only when the causal link is "appropriate" (57). Apparently, the non-branching causation would provide the kind of appropriate causal link required here. Therefore, it is quite plausible to say that Shoemaker's psychological continuity account of personal identity does not follow from functionalism alone, but from functionalism together with an assumption about the persisting self as a sort of individual substance; an experiencing subject that is the bearer of mental properties. And that the characteristic causes and effects of a person's mental states are always states of that person. Shoemaker summarizes his view about the appropriate causal link between mental properties and personal identity as follows:

What the functionalist view claims is that it is of the essence of a mental state to be caused in certain ways, and to produce, in conjunction with other mental states, certain effects (behavior or other mental states). But of course, it is in conjunction with other mental states *of the same person* that a mental state produces the effects it does; and its immediate effects, those the having of which is definitive of its being the mental state it is, will be states (or behavior) on the part of the *very same person* who had the mental state in question (58).

Again, he writes:

The functional role of a mental state will consist in part in its generating appropriate successor states, and affecting future behavior, usually in conjunction with other mental states. and those successor states, and that behavior, will be on the part of the possessor of the mental state. ...Now, whatever else person are, they are subjects of mental states. And as subjects of mental states, they had better have the persistence conditions that go with this. It had better be the case that when mental states generate their appropriate successor states, or their appropriate behavioral expressions, those successor states belong to the same person who had the states that generated them, and those behaviors are on the part of that person (59).

These passages show that Shoemaker's psychological continuity account of personal identity is deeply committed to the non-branching restriction clause. They also indicate that the sort of appropriate causal connection between the relevant successive mental states of a person (diachronic unity) is grounded in a mechanism that involves the persistence of the person's actual brain, or body for that matter. If this true, which seems to be the case, it means that any theory about the nature of mental states may not imply a robust theory about the identity conditions of persons. That is to say, functionalism might not tell us anything serious about personal identity over time.

Olson concludes that the causal-powers argument fails. Alternatively, he argues that if we concede that functionalism is right to say that the characteristic causes and effects of one's mental states must always be states of the same being, then *Organisms* are the best candidates. In Olson's view, the animal organisms are in such a state that exhibit the causes and effects characteristic of mental states. "We could then conclude," he claims, "as philosophers of mind typically do, that organisms are the bearers of mental properties" (60). Therefore, the organism is necessary and sufficient for our continuity over time.

Olson criticisms are challenging and difficult to meet. I think that he is right to say that Shoemaker's psychological-functional continuity account presupposes personal identity. Likewise, Hume's functionalistic account of the self, I think, encounters the same difficulty, and that such difficulty is a possible source of self-criticism in the Appendix. It seems to me that this major problem in Hume's account of the self arises as a result of his impersonal description of the self and the reification of perceptions. I must consider these issues now.

IV.4: Hume's Reification of Perceptions and the Impersonal Description of the Self

Derek Parfit is the best-known contemporary philosopher who defends an impersonal description of the self. The impersonal description thesis tells that "because we ascribe thoughts to thinkers, we can truly claim that thinkers exist. But we cannot deduce, from the content of our experiences, that a thinker is a separately existing entity. And ... because we are not separately existing entities, we could fully describe our thoughts without claiming that they have thinkers. We could fully describe our experiences, and the connections between them, without claiming that they are had by a subject of

experiences”⁽⁶¹⁾. The Impersonal description of personal identity is constituted by ‘non-branching psychological continuity and connectedness.’ This psychological continuity consists in there being a chain of psychological connections. What matters, indeed, is not personal identity as a perfectly determinate entity; rather, what is important in the survival of persons is the sort of continuity or persistence of our psychological traits over time, which relates the earlier stages of one’s existence with the later stages of her existence. Therefore, strictly speaking, personal identity consists in the continuity of certain psychological connections or relations.

It is my view that Hume’s functionalistic account of the self fits perfectly within this kind of impersonal description of personal identity. Hume, like Parfit, believes that because we ascribe thoughts to thinkers, there must be thinkers, but he also holds that we cannot deduce from the content of our experiences that the thinker is an ontologically distinct existence to which all the experiences belong. According to him, we can account for our thoughts and experiences without claiming that they belong to a separate subject of experience. For Hume, the causal/functional account of the mind is the perfect theory, which explains our thoughts and experiences without making any claim about their subject. Functionalism, according to him, provides an impersonal description of personal identity.

The impetus for impersonal description of persons is to avoid the circularity objection that is brought up against the psychological continuity criterion of personal identity. Parfit shows that one way to avoid circularity is to define personal identity in terms of Relation-R hypothesis according to which the continuity of a person is determined by the existence of certain causal and/or functional relations *per se* rather than by relating it to an independent subject of experience, whether that be a Cartesian Ego or Shoemaker’s independent subject of experience. Consequently, we can define a person in terms of the causal/functional roles that the earlier stages of a person’s psychological existence and her later stages play in her overall spectrum of life.

According to the impersonal description thesis, we can compare persons with entities such as clubs, nations, commonwealth, etc. There is no doubt that this sort of reductionist account of the self is derived from Hume’s analogy of the human soul to a republic or a commonwealth⁽⁶²⁾. This reductionist characterization of personal identity

consists in the fact that persons can be given impersonal descriptions, that is, descriptions which have no actual reference to any specific person, or a particular subject of experience. Because “we are not separately existing entities, we could fully describe our thoughts without claiming that they have thinkers” (63). When we ascribe the perceptions to a mental subject, in fact, the subject of experiences is mentioned “only in the *content* of the thought, this sentence does not ascribe this thought to a thinker” (64). In other words, “persons are *logical constructions* out of entities whose existence does not require that they be states of persons or other ‘subjects’—entities whose existence is not ‘adjectival’ on mental subjects, in the way seeings are adjectival on seers, deeds are adjectival on doers, or (for a non-mental example) dents are adjectival on dentable surfaces” (65). Hume thinks of mental entities, i.e., ‘perceptions,’ as the building blocks out of which persons are logically constructed (66). Hume has famously claimed that perceptions are ontologically independent existences, and that their existences are independent and in a sense more fundamental than the subjects that experience them. This is called the reification of perceptions. I think that Hume’s reification of perceptions is closely related to his impersonal description of the self. I should explain this now.

Hume’s claim for the reification of perceptions is extremely problematic. Harold Noonan has shown this more clearly. The reification of perceptions thesis shows that Hume is a relational (non-substantial) theorist about the self (67). Noonan remarks that Hume rejects one’s introspective awareness of any self or mental substance. Hume in a well-known passage asserts:

When I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe any thing but the perception” (THN, 252).

Hume then concludes that “[personal] identity depends on the relations of ideas; and these relations produce identity” (THN, 262), and that “there is properly no *simplicity* in it at one time, nor *identity* in different” (THN, 253). Presumably, this is the most explicit statement made by Hume suggesting that neither personal identity consists in a notion of substance, nor that personal identity really matters. Some commentators have said that Hume fails here to recognize the fact that he is, indeed, observing something (68). It is

pointed out that if Hume has no idea of a self, then how he knows that he is not observing one when he stumbles on some particular perception. Specifically, it is claimed that the self that Hume professes to be unable to observe is the one that he finds to be stumbling onto (69). Noonan himself has argued along the same line; he holds that the reification of perceptions results in Hume's inability to explain the relation between the perceptions and their subjects. It seems that there must be some sort of Shoemakerian substantial self in Hume's bundle theory. Of course, there is an appealing aspect in this kind of argument. Because when Hume says that "I never can catch myself at any time without a perception" he could be read as saying that 'nothing but perceptions is ever observed.' As Noonan points out, this 'assertion would have committed him to denying that *anyone* ever observes anything but perceptions, and so would have gone far beyond the evidence available to him. For how could he know that?" (70).

Noonan argues that Hume's denial is not an empirical one. On the contrary, he claims that Hume's denial of the introspective inaccessibility of the self is the result of his conception of all perceptions as relational mental states that have no intrinsic properties to be called substance. "Hume starts from a conception of mental states," Noonan argues," according to which to be in a mental state is for a certain relational statement to be true of: for one to be in any mental state is for one to be *perceiving* a certain sort of *perception*" (71). He then suggests that "for if to be in any mental state is to possess a relational property of the type: perceiving a perception of type *x*, then *no* mental state can be an *intrinsic* property of its subject. [And] given that the only states of which one can be introspectively are mental, then, introspective awareness of a self would require awareness of it without any awareness of its intrinsic properties. But surely it makes no sense to speak of observing something introspectively if the thing has no intrinsic properties whatsoever which one can observe by introspection" (72). So Noonan claims that since no mental state can be an intrinsic property of its subject, it follows that we lack introspective knowledge of the self. More fundamentally, he argues that such Humean conception of mental states—as not being an intrinsic property of its subject—implies *nothing* whatsoever about the relation between their subjects and their objects, namely, Hume's perceptions. What Hume's conception of the mental implies is that "these perceptions are *things* to which the subject of mental states is related in

somewhat the way in which, as we ordinarily think, we are related to trees or tables” (73). He claims “that just as trees and tables might exist unperceived, so a pain or a dizzy spell or thought might occur without occurring to anyone” (74).

In Noonan’s view, the root of the problem is Hume’s reification of perceptions, or to think of them as ontologically independent existence. Here is what Hume writes:

Now as every perception is distinguishable from another, and may be consider’d as separately existent; it evidently follows, that there is no absurdity in separating any particular perception from the mind; that is, in breaking off all its relations, with the connected mass of perceptions, which constitute a thinking being. (THN, 207)

Noonan reads this passage as saying that any perception, like having a pain or a dizzy spell, “*might have been the only thing in the universe*,” that is, “a pain or a dizzy spell or a thought might occur ‘loose and separate’ without belonging to anyone” (75). There is another passage that, he thinks, completes Hume’s reification of perceptions as substances. Hume writes:

Since all our perceptions are different from each other, and from everything else in the universe, they are also distinct and separable, and have no need of anything else to support their existence.

They are, therefore substances, as far as this definition explains a substance. (THN, 233)

According to Noonan’s reading, this passage establishes that perceptions are, indeed, substances, that is, they are something that may exist by themselves. This causes a problem about the status of the relation between perceptions as ontologically (logically) independent existences and the self as the thing that has those perceptions. When Hume claims that perceptions are ontologically independent existences, he “thus denies the only possible basis for regarding the self, *qua* perceiver, as ontologically prior to its perceptions... [and] once perceptions are reified as substances no other conception of the self makes any sense at all” (76). In sum, Hume’s enterprise is misconceived from the start, and thus it is radically defective.

What we can gather from the reification of perceptions thesis is two things: (a) Hume’s reification of perceptions as ontologically independent existences indicates that his conception of the perceptions (the bundle) is a relational one; and (b) a relational conception of mind is doomed to fail because it is unable to account for the continuity of such ‘loose and separate’ mental states—a pain or a dizzy spell—which may exist as the only thing in the universe. Observation (a) says that Hume’s theory of mind is a

causal/functional one, but without making any commitment to substance, and (b) claims that the functionalist theory of mind, as conceived by Hume, fails since it is unable to account for the continuity of personal identity over time. These two opposing views give rise to a tension in Hume's psychological-functional continuity account of personal identity, and, I think, it shows the kind of philosophical labyrinth that he finds himself in. I turn to this now.

IV.5: Tension in Hume's Functionalistic Theory of Personal Identity

The tension that I find in Hume's functionalistic theory of the self is embedded in his reification of perceptions and the impersonal description of the self on the one hand, and his psychological-functional continuity account of personal identity on the other. Specifically, on the one hand, Hume's reification of perceptions and the impersonal description of the self leave him with a bundle that lacks an experiencing subject to which all the perceptions belong, and on the other hand, his psychological-functional continuity account of personal identity requires an experiencing subject to which all the perceptions belong and continues over time. In fact, I think that this tension eventually gives rise to circularity in Hume's theory of personal identity. Hume's psychological-functional continuity account of the self requires an experiencing subject that continues over time and that all the experiences belong to. Yet, at the same time his causal/functional characterization of perceptions refuses to allow that perceptions belong to a continuous subject that is distinct and separate from the experiences themselves. And since Hume is committed to the general idea that the self somehow continues over time, this means that the specification of mental states in terms of their causal roles must be states of *the same person*. This clearly presupposes a notion of personal identity.

The circularity that I ascribe to Hume is very similar to the one Butler raised against Locke or the one Olson levelled against Shoemaker. Hume does not mention anything about circularity in the text of the *Treatise* or in the Appendix. However, Butler's best known criticism of Locke's consciousness-memory account of personal identity, in which he claimed that consciousness or memory cannot be part of a definition of personal identity because it already presupposes personal identity, had surfaced a few years before the publication of the *Treatise*. Specially, if we take into consideration the fact that personal identity "has become so great a question in philosophy, especially of late years

in *England*” (THN, 259), then it would be very surprising if Hume was unaware of Butler’s circularity objection (77). Furthermore, a similar criticism is found in Samuel Clarke’s debate with Anthony Collins where he says that by consciousness Locke must mean “That Consciousness by which I *not only remember* that certain Things were done many Years since but also *am conscious that they were done by Me, by the very same Individual Conscious Being* who now remember them” (78). So, historically speaking, it is quite possible that the circularity was known to Hume, though it is not clear why Hume does not discuss it. In fact, Hume pays no attention to a whole host of issues concerning personal identity which were being hotly debated at the time. For instance, he never discusses Locke’s notion of “sameness of consciousness” or his prince and cobbler thought experiment (79).

As I have shown in IV. 3, Olson rejected the psychological-functional continuity account of personal identity because he thought that it implies nothing about how mental properties are related to their bearers; the nature of mental properties does not tell us anything about the identity conditions of their bearers. Olson denied that there is connection between functionalism, as a theory about the nature of mental states, and the identity conditions of the bearers of those mental states. The reason that he denied such a connection is that he does not think that we are justified in believing that a person is a substance in the subject of properties sense. He argued that to think about a person as an individual substance as the subject of properties presupposes the existence of personal identity. Olson’s objection is similar to the circularity criticism that Butler leveled against Locke claiming that consciousness (memory) presupposes personal identity, and therefore, cannot constitute personal identity. Olson is right to claim that, I think, the psychological-functional continuity account of personal identity “doesn’t follow from functionalism alone, but only from functionalism together with an assumption about personal identity: the assumption that the persistence of thinking beings has something to do with psychology, for instance” (80). Accordingly, the psychological-functional continuity account of personal identity clearly presupposes a mental subject to which the psychological properties belong. In a nutshell, on the one hand, it seems that functionalism *without* psychology does not tell us much about personal identity over time, and on the other, functionalism *with* psychology does presuppose personal identity

over time.

I think that this is the problem that Hume finds himself in. Basically, the claim is that if Hume is right to say that something like causal/functional continuity would suffice for persons to persist through time, then he must show that the specification of mental states in terms of their causal roles must be states of *the same person*. But Hume cannot hold this conclusion unless he presupposes the existence of a persisting self. Apparently, Hume identifies a perception (mental state) in terms of its causal/functional role in combination with other mental states that are members or states of *the same person*, and the effects (outputs) that such mental states produce belong to *the same person* or mind as well. If one characterises mental properties in terms of their membership in a single mind, one indeed, invokes or presupposes a notion of personal identity; one cannot have a complete account of what those properties are without making assumption about the persistence conditions of their bearers. This strongly suggests that Hume's attempt to define personal identity in terms of the causal/functional relations among perceptions (mental states) will lead to circularity. Because, on the one hand, he holds that the causal/functional chains of perceptions belong to the same person, and on the other, he identifies personal identity in terms of the existence of causal/functional relations among mental properties. This is straightforwardly circular.

In sum, if we combine Hume's impersonal description of the self with his reification of perceptions, the result is quite subversive. There is no doubt that an impersonal description of the self in which each perception is so radically reified, and thus, exists distinctly and separately from other perceptions in the bundle, circumvents the circularity that is inherent in Hume's account of the self. In fact, an impersonal self is a non-substantial self, and a non-substantial self is a self without a mental subject. Therefore, it seems clear that if there is no mental subject to which all mental states belong, then there is no way that we could presuppose it. Thus, circularity circumvented. However, if there is no mental subject to which mental states or perceptions belong and that each perception itself is a distinct subject, then it is meaningless to ascribe personal identity to the bundle of perceptions itself. This is symptomatic of a tension in Hume's functional continuity account of personal identity. Hume's reification of perceptions and the impersonal description of the self is the main source of this tension in his

psychological-functional continuity account of personal identity. In fact, if we could reduce, as Hume claims, the mind into a single perception, for instance, the life of an oyster, that “have only one perception, as of thirst or hunger” (THN, 634), in that situation we conceive that perception alone. That single perception does not belong to any mental subject or bundle of perceptions. Simply stated, we have no notion of the self involved here. And if this single perception does not give rise to a notion of personal identity, “the addition of other perceptions can never give you that notion” (THN, 634) either. Thus, in Hume’s view, our experiences or the bundles of perceptions do not belong to a subject or thinker that is distinct and over and beyond those perceptions themselves. So there might be some sort of continuity in the bundle, but that cannot be a psychological continuity in the sense that all mental entities belong to a mental subject. Therefore, it seems that the bundle of perceptions itself does not tell us anything about its continuity condition. In fact, it does not imply personal identity. Hence, I may suggest that Hume’s reification of perceptions and the impersonal description of the self undermine entirely his theory of personal identity; it undermines his functionalistic account of the mind, which implied a psychological-functional continuity account of personal identity over time. If this is true, then, Hume’s functionalism does not imply personal identity. So it seems to me that Hume’s reification of perceptions and the impersonal description of the self are the main sources of a tension in his account of the self, which eventually makes him to confess it later on.

Conclusion

Overall, I think that there is much less reason to reject the general idea that functionalism implies a sort of psychological-functional continuity account of personal identity over time. It seems to me that Shoemaker’s basic insight—the claim that the functionalist view about the nature of mental states can provide us with a continuity account of the self over time—is the kind of account that we need for a theory of the self. Moreover, I have argued that we can find a similar project in Hume’s philosophy. In particular, Hume’s account of personal identity as it regards passions provides sufficient ground for making that comparison.

However, this view becomes problematic when we tie functionalism with a notion of mental subject or a substance. There is one way out of this problem: functionalism

without mental subject or substance. This is essentially a Humean project, at least, as I conceive it. Hume finds, indeed, the same problem in Locke's theory of personal identity and tries to repair it. What he does is that he provides a psychological-functional continuity account of identity without having tied it to any notion of substance. In fact, his reification of perceptions and the impersonal description of the self are intended to avoid this problem. As a result, Hume's functionalistic theory of the self makes no commitment to the notion of substance. He, unlike Locke and Shoemaker, specifies the identity conditions of our persistence in terms of the mere mechanical operations of our perceptions (mental states) and the sort of causal powers that they impose on each other. Hume does not think that there exists an independent and distinct entity like a substance or an experiencing subject over and above the very causal/functional processes taking place within the bundle. In other words, Hume does not think that the mere existence of such causal/functional processes *per se* gives causal powers to its bearers. An impersonal description of the self does not need a mental subject to which the perceptions belong. Nevertheless, I think that Hume's impersonal description of the self together with his reification of perceptions stores a much deeper problem for his functionalistic account of personal identity. It gives rise to circularity.

Hume can circumvent the circularity by adopting a substantial view of the self, something that Shoemaker is committed to. That, however, would substantially undermine his philosophy as a whole. Hume got the gist of the psychological-functional continuity account of personal identity right, but he makes a mistake to think that such continuity is possible without making commitment to a substantial view of the self. The functionalist account of personal identity is badly circular. Given that this is the case, Hume supposedly has the two following options at his disposal: (a) either he has to allow that all perceptions are united in the same single bundle by either some real connexion or by inherence in some simple substance; or (b) he rejects that they are necessarily members or states of the same single bundle (mind). If he holds the former, it circumvents the circularity, even though it entirely undermines his philosophy. Presumably, Hume realizes that he cannot posit a theory of the self without a simple substance in which perceptions inhere, and that is why he confesses it in the Appendix. And if he holds the latter it leads to the problem of 'bundle incoherence' or *other minds*

problem. The problem of 'bundle incoherence' or other minds problem arises when the compound of perceptions is too 'loose' to be united in one unified mind (bundle). It claims that because the causal/functional definition of perceptions does not specify the persistent condition of their bearers, it follows that we have no necessary and sufficient criterion to secure the coherence of the bundle. In fact, it says that it is quite possible that qualitatively identical perceptions occur in other minds. Clearly, Hume's account of sympathy recognizes that one can share the mental states of other minds. In this chapter, I have discussed situation (a). In the following chapter, I will discuss situation (b).

V

Hume's Second Thoughts about Personal Identity

But all my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought or consciousness. I cannot discover any theory, which gives me satisfaction on this head. (THN, 635-36)

They are the coherence and constancy of our perceptions, which produce the opinion [idea] of their continu'd existence. (THN, 217)

Hume prefers to determine the justificatory status of a belief with reference to generic properties of belief-forming mechanisms... beliefs are assessed epistemically with reference to the underlying intellectual character or disposition that produces them. The justificatory status of a belief derives from the justificatory status of the mechanism that produces it. (Loeb 2002, 13)

Hume seems to think of constancy *and* coherence as supplying individually sufficient and disjunctively necessary conditions of the impressions to which we ascribe an outer existence. (Wright 1983, 64)

Hume's idea of the self not only becomes a 'labyrinth' for himself, but a major interpretative dilemma for his readers. In his original characterization of personal identity (Chapter I), Hume takes pride in explaining why we regard the self as an entity having identity through time. Yet, in the Appendix, he feels frustrated and thinks that he has gone astray. On his second thought, he talks about an emerging inconsistency or contradiction. There, he expresses strong dissatisfaction with his earlier account of personal identity and brings against himself the charge of inconsistency.

In the section of “Of Personal Identity,” Hume characterizes the self as a bundle of perceptions to which we attribute identity and simplicity by virtue of the association of its members. He rejects the idea that there is an impression from which the idea of the self is derived. Therefore, the “self or person is not any one impression, but that to which our several impressions and ideas are suppos’d to have a reference” (THN, 251). Supposedly, if an impression produces the idea of the self, Hume argues, “that impression must continue invariably the same, thro’ the whole course of our lives; since self is suppos’d to exist after that manner. But there is no impression constant and invariable. Pain and pleasure, grief and joy, passions and sensations succeed each other, and never all exist at the same time. It cannot, therefore, be from any of these impressions, or from any other, that the idea of the self is deriv’d; and consequently there is no such idea” (THN, 251-52). As a result, he declared that “when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe anything but the perception” (THN, 252). But then he asks after what manner these perceptions belong to self, provided that all these perceptions are different, distinguishable, and separable from each other, and may exist separately, and have no need of anything to support their existence. His answer to this question is that we “are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement” (THN, 252). So he claims that when we reflect upon ourselves we only find succession of ordinary but related impressions and ideas. We can never catch ourselves at any time without perceptions; we never encounter a self, but only some collection of perceptions (Chapter I).

However, Hume argued that we are inclined to believe that we are simple and identical, and persist over time. According to him, the idea of the self is a ‘fiction’ that results from the confusion of the succession of related perceptions (objects) with that of the uninterrupted and invariable objects. What gives us such a great natural propensity to ascribe an identity to these successive perceptions, and to suppose ourselves possessed of an invariable and uninterrupted existence through the whole course of our lives, is that we generally confound ‘perfect identity’ with the ‘succession of related perceptions.’

Hume argued that resemblance and causation are the cause of this confusion or mistake that make us substitute the notion of identity with diversity (THN, 254). So we “feign the continu’d existence of the perceptions of our senses, to remove the interruption: and run into the notion of a *soul*, and *self*, and *substance*, to disguise the variation” (THN, 254). Our propensity to this mistake is so powerful that we fall into it before we are even aware of it; and though we may correct this mistake, and return to a more reflective method of thinking, “we cannot sustain our philosophy, or take off this bias from the imagination” (THN, 254). In short, in Hume’s view, the mind is a collection of ontologically distinct existences, each separable and distinguishable, lacking ‘real connexions’ among them; and we feel connexions between them because of the effects of the associative principles, viz., resemblance, causation, and continuity in time. The source of the idea of the self is a “smooth and uninterrupted progress of the thought along a train of connected ideas” (THN, 260). He concludes that identity is not something inherent in the perceptions or in the relations connecting them together, but merely a quality that we attribute to them because of their union in the imagination. Therefore, the idea of the self is neither derived from a simple and unchanging impression nor it is constituted by some ‘real connexions’ among perceptions. Such an idea is a fictional existence that we merely attribute to the succession of our related perceptions. This was Hume’s official position in the main text of the *Treatise*, and he never wavers it until he arrives at the Appendix.

In the Appendix, however, Hume expresses dissatisfaction with his original account and finds himself “involv’d in such a labyrinth” that he neither knows how to “correct” his original account “nor how to render them consistent”. He finds a serious defect in his account of how we come to think of our perceptions as bound together in a ‘real simplicity and identity.’ He claims that “all my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought or consciousness” (THN, 636). He does not explain why he is dissatisfied with his earlier account of the “uniting principles,” yet asserts that “there are two principles, which I cannot render consistent; nor is it in my power to renounce either of them, viz. *that all our distinct perceptions are distinct existences*, and *that the mind never perceives any real connexion among distinct existences*” (THN, 636).

This statement is problematic. It is widely acknowledged that the two principles to which Hume ascribes inconsistency are not, indeed, inconsistent. In fact, as it has been pointed out, the second principle is derived from the first one. Therefore, one must look for the source of the inconsistency in another place. It is fairly clear from the earlier passage that these two principles are inconsistent with his previous account of the unity or identity of the self. Surprisingly, however, Hume does not say much about this newborn problem. As a result, Hume's self-doubt and his failure to situate the inconsistency lead him to plead the privilege of a sceptic. This has not only created an interpretive dilemma, but has also left us with a theory of the self that is shattered into pieces.

Hume's worries about his earlier account of personal identity in the Appendix have puzzled most of his readers and interpreters, and a dozen interpretations have been proposed. Yet, no consensus as to what exactly Hume's worry is has emerged. The Appendix continues to baffle the scholars. In this chapter, I begin to assess some of the most recent and relevant explanations of Hume's second thoughts about personal identity. Then, I present my own account of the source of Hume's inconsistency.

According to my interpretation of the inconsistency, Hume's functionalism, as presented in the main body of the *Treatise*, stores a problem for him and when he arrives at the Appendix he realises the problem and confesses that he is unable to resolve it. In particular, I claim that the problem that leads to the inconsistency has two main possible sources: The first possible source of Hume's inconsistency has to do with the question of functionalism and its implication on personal identity over time. Hume might have detected circularity in his original characterization of personal identity; something similar to the one Butler levelled against Locke claiming that consciousness already presupposes personal identity; or even something similar to what Olson has raised against Shoemaker's psychological-functional continuity account of personal identity claiming that the identification of a mental state in terms of its combinatory relations with other mental states that are members or states of *the same person*, presupposes a notion of personal identity. I attempted to show that if something like functional continuity suffices for persons to persist over time, then Hume must show that we can have a *complete* account of how one's mental states produce the idea of a persisting self without making

assumption about the identity condition of their subject. But since functionalism, including Hume's, identifies a mental state in terms of its functional relations to other mental states that are the states of *the same person*, it clearly presupposes a notion of personal identity. This seems to be straightforwardly circular. I have discussed this difficulty in the preceding chapter.

The second source of Hume's inconsistency, which I discuss in the present chapter, is caused by the failure of his principles of constancy and coherence as the condition (criterion) of personal identity. Hume's account of constancy and coherence, as laid out in the section of "Of Scepticism with Regard to the Senses," establishes that the principles of constancy and coherence unite our series of successive perceptions in the mind by a process of imaginative postulation, or supplementation, of an unperceived perception, or an unsensed sensibillum. I contend that the principles of constancy and coherence may successfully account for the arising belief in the idea of the continued and distinct existence of external objects and the idea of personal identity, but they fail to explain our belief in *other minds* (selves). In fact, Hume's account of *sympathy* in Book II of the *Treatise* establishes that we *actually* share the mental states of other minds (selves) around us. This shows that Hume treats the belief in other selves, like the beliefs in external objects and persons, as an important part of his psychology of identity ascriptions. Therefore, he must use the same psychological belief-determining mechanism, particularly the principles of constancy and coherence, to account for our belief in other minds. Nevertheless, coherence and constancy do not play any role whatsoever in the explanation of the arising belief in other minds. Further, the absence of constancy and coherence in the explanation of the belief in other minds shows that not only the belief in other minds is not explained adequately by Hume's psychological belief-forming mechanism, but also the absence of constancy and coherence undermines even their roles as the necessary and sufficient conditions for the beliefs in objects and personal identity. This problem becomes even more pressing, if we note that Hume's theory of mind as defended in Book I of the *Treatise* is incompatible with his account of the self in Book II.

In V.1, I present some of the most recent and relevant interpretations of Hume's inconsistency. In V.2, I discuss Hume's principles of constancy and coherence as the

criterion of personal identity. In V.3, I present and defend my own preferred explanation of Hume's second thoughts on personal identity.

V.1: Interpretations of Hume's Second Thoughts

A dozen interpretations have emerged as a result of Hume's failure to spell out the inconsistency. In what follows, I will present some of these explanations. The list of the explanations that I will reproduce in the following pages is not exhaustive ⁽¹⁾. However, there is a common theme which ties all of these different interpretations together. This common theme that runs in all of these different explanations is derived from Hume's famous confession claiming that he is unable to explain the principles—resemblance, causation, and contiguity—that unite our successive perceptions in our thought or consciousness (THN, 635-36). I also think that what causes Hume to worry so much is the insufficiency of these principles to unite the ontologically distinct perceptions into a single and coherent bundle (mind). I turn to these interpretations now.

Inadequacy of Resemblance and Causation

A. H. Basson and S. C. Patten propose that Hume's confession is due to the inadequacy of causation and resemblance in explaining the principles that unite our successive perceptions in the mind. Stroud also considers this explanation. They suggest that resemblance and causation are not sufficient and necessary to associate the ideas in the mind and, thus, bring about coherence in the bundle. On the one hand, Basson argues that these two relations are not *sufficient* because a perception in one person's mind may cause a resembling perception in some other minds. On the other hand, Patten claims that these two relations are not *necessary* because not all of our perceptions are related by causation and resemblance, especially the impressions of sensation ⁽²⁾. He notes that, an impression of a coffee cup, for instance, neither resembles nor causes the succeeding impression of a pen. He concludes that "various mental states are somehow non-associatively related and it is the presence of these other relations which accounts for the unity of a single mind" ⁽³⁾. Furthermore, Stroud argues that resemblance and causation are neither sufficient nor necessary to bind perceptions together in one mind. It is said that resemblance is not sufficient because it cannot guarantee a coherent conception of the mind. Consider a bundle of perceptions composed of all those actual perceptions of the Eiffel Tower (Stroud's example), which resemble each other, but certainly do not

constitute one mind simply because the resembling perceptions belong to different minds. So, resemblance is not sufficient to give rise to a coherent notion of mind. It is also said that causation is not sufficient because it is possible that a perception in one mind might cause a perception in some other minds, which again would cause incoherence in the bundle. Moreover, the associative principles are not necessary because they fail to account for all our perceptions. That is to say, not all of our perceptions are related by resemblance and causation. The natural relations of resemblance and causation are inadequate for Hume. Stroud argues that, in the case of causality, Hume cannot find any objective “conditions that can be said non-circularly to be both logically necessary and sufficient for the truth of ‘A caused B’” (4). He suggests that this is also true of the idea of personal identity. The upshot is that if mental states or events are nothing but the presence of certain objects in the mind called “perceptions” while lacking objective conditions to connect, or relate, a particular perception to a particular mind, then, perceptions are detached from the ‘minds’ or the ‘subjects’ that have them. It follows that perceptions are ‘loose’ and ‘floating,’ independent of all the rest and independent of other minds. Stroud concludes that Hume, by placing exclusive emphasis on mental objects, while ignoring the subjects in whom those mental objects exist, pushes the theory of ideas to the edge of an abyss and realizes that only a notion of substance or real connection among perceptions can save him (5). This conclusion suggests that in order to have a coherent bundle, there must be some objective conditions to bind perceptions in one simple mind. Hume’s theory of ideas, especially the principles of resemblance and causation, lack that condition and thus is inadequate to do the job for him.

These philosophers identify serious misgivings in Hume’s theory of the self. These misgivings, however, can account for certain aspects of Hume’s “labyrinth.” The idea that resemblance and causation fail to account for the identity of the self is crucial to Hume’s philosophy of the self. As we will see in a moment, David Pears, Robert Fogelin, and Don Garrett all have held some versions of the inadequacy explanation, one way or another. More importantly, however, this cluster of explanations began to shift Hume scholarship on the question of personal identity to the question of *other minds*, which is fundamental to Hume’s functionalistic theory of the self. Nevertheless, they are not conclusive and Hume seems to be able to account for some of these objections. In

response to the sufficiency argument (Basson and Stroud), Hume can allow that a perception in one mind could cause another perception in the other mind. But these perceptions either (a) belong to a large single bundle or (b) they belong to two distinct bundles. Now if we say there is only one single direct causal relation between the two bundles “then no doubt the latter alternative would remain the more natural, but should the direct interrelations multiply in number and complexity, then the former would indeed become more natural, as Hume’s theory suggests”(6).

In response to Patten, who raises concerns about the ownership of impressions of sensations, Garrett has made some useful points concerning the impressions of sensations, though it remains unsure if these points could do Hume any benefits at all. First, he points out that a bundle of perceptions is not supposed to involve *direct* and *immediate* causal relations between each perception and every other in the bundle. All that is needed is a web of interrelations between perceptions in which all ideas and impressions of reflections are causally dependent, or derived, from impressions of sensations. This web of interrelations between perceptions produces *indirect* causal interactions between impressions of sensations as well. On Garrett’s view, we cannot surely say that there is no causal relation or resemblance between the impressions of sensations. Secondly, he claims that all the impressions that we perceive at any given time is “nearly always *partly*” dependent on our wills (THN, 399). These impressions of wills in turn are caused by other perceptions. Thus, he concludes that it is not true to say that one perception is the whole cause of another perception in order to have the required causal relation. Also, he argues that, since the “vulgar”—that is, most of us—cannot distinguish between objects and perceptions it follows that the mind cannot distinguish causal interactions among objects from causal interactions among perceptions (7).

The criticisms that Garrett has made against Patten and Stroud face difficulties of their own. Responding to his first point, it is unclear how the web of interrelations among perceptions can account for the production of an indirect causal relation among impressions of sensation. Garrett needs to produce a mechanism to explain how the ideas and impressions of reflections that are causally dependent on prior impressions of sensation can give rise, even indirectly, to impressions of sensation. It is a basic principle in Hume’s empiricist theory of ideas that ideas and impressions of reflections cannot

produce impressions of sensations. So it is insignificant how complex the web of causal interactions between perceptions is or those ideas and impressions of reflections interact causally in numerous ways; this claim does not establish the fact that the complex web of perceptions produces even an indirect causal relation between distinct impressions of sensation. It is not known how Garret is going to account for causal and resembling relations between, for instance, an impression of a cup of coffee and the succeeding, but distinct impression of a pen. More fundamentally, even if Garrett was right in arguing that a complex web of causal interactions among perceptions could somehow explain the causal relations among the impressions of sensation, it would have violated Hume's empiricism in the sense that it is no longer the impressions of sensation that is the cause of ideas and impressions of reflection. It is the other way around. This in turn violates the causal thesis entailed in Hume's Separability Principle. Furthermore, both Patten and Stroud explicitly argued that the chain of causal connection, on Hume's theory of ideas, runs "vertically" from the impression up to the idea, and then to other ideas and impressions (impressions of reflections). What is lacking, however, is a causal chain that runs "horizontally" between the incoming perceptions (impressions of sensation) that we receive constantly through our sensations (8). So it is not true that each of our perceptions is caused by and resembled by other perceptions of ours. I think that Patten and Stroud got it right and that Garrett is unable to account for this Humean misgiving. Garrett, in his original paper on "Hume's Naturalistic Theory of Representation", quite convincingly argues that impressions of sensation can represent bodies and their qualities (9). There is no indication, however, that representation between perceptions runs "horizontally." In other words, a particular impression of an object, say a building, as in Stroud's example, does not have any causal and resembling relation with another particular impression, say, a tree. In fact, it is evident from our experiences that impressions of sensation constantly enter into our consciousness (bundles of perceptions) independently of other impressions and the perceptions that already exist in the bundles. As Garrett rightly pointed out, there is no representational content in the latter impression that comes from the former one. In consequence, Stroud and Patten are right in saying that resemblance and causation fail to account for the unity of the bundle.

In response to Garrett's second reason, I must point out that the passage that

Garrett refers to does not indicate that impressions of sensation depend upon our impressions of will or volition. Hume says that the *will* is “nothing but the *internal impression we feel and are conscious of, when we knowingly give rise to any new motion of our body, or new perception of our mind*” (THN, 399). The definition of the *will* captured in this passage refers to it only as an impression of reflection (internal impression), and not as an impression of sensation. However, the will, only as an impression of sensation, can serve Garrett’s purpose here. Now, if Hume defines the will as an impression of reflection, as the text indicates, this brings us back to Garrett’s previous points—that is, ideas and impressions of reflection are indirectly causally related to impressions of sensation—to which I have already raised some objections. In conclusion, I am not convinced that Garrett’s objections to Patten’s and Stroud’s explanations concerning the failures of resemblance and causation do them justice. In particular, his objections do not undermine Patten’s and Stroud’s main thesis that our impressions of sensation are not related by causation and resemblance.

Patten and Stroud are right to say that impressions of sensation are not related by causation and resemblance. However, it seems to me that this is not the only cause of a serious inconsistency in Hume’s theory of the self, though it clearly shows that resemblance and causation are inadequate to produce a unified self. Hume adopts his Theory of Ideas from his predecessors, most notably from Descartes and Locke, without any further attempt to defend it, or even to give a careful exposition of it. Hume is able to see that his commitment to the Theory of Ideas is central to his construction of an empirical (natural) cognitive science of the mind. In this sense, he is constructing a naturalistic philosophy of mind, which is essentially Cartesian. According to Descartes’ Representationalism, concepts (ideas) are mental particulars that represent objects and things in the mind. To have the concept implies that one is able to think about whatever it is the concept of—that is, Descartes’ representationalism is inherently *intentional*. Representationalism in this sense is essentially “vertical.” The concepts, or ideas, are always the representations of objects, which are mediated through impressions of sensations. In this sense, having a “horizontal” representation between impressions of sensation is entirely pointless, if not meaningless, insofar as such representations would lack intentionality. Even if we had assumed that there was a causal and resembling

relation between an impression of a cup of coffee and an impression of a pen, it would not have been correct to say that the content of the latter impression is the representation, or is *about* or *of*, the former impression. If we had said that, then, the whole intentionality issue would have fallen apart—which Hume would hesitate to allow. So, for Hume, the paradigmatic and definitive form of representation is the “vertical” one that runs from an impression of sensation up to the idea and, then, to the other ideas and reflective impressions. Hume, of course, thinks that the “horizontal” representation is less important for his theory of ideas perhaps because he thought of impressions of sensation as the primary *medium* through which the bodies and their qualities are represented in the mind. However, as Patten and Stroud have shown, he later realizes that the relations of resemblance and causation are inadequate to bind perceptions in the mind. I will explain later that what is specifically the cause of Hume’s worries is the failure of the principles of constancy and coherence as the criterion of personal identity.

The Necessary Ownership of Perceptions

David Pears, like the previous explanation, locates the cause of Hume’s labyrinth in the inadequacy of causation and resemblance to unite our successive perceptions into a bundle. He claims that the main deficiency in Hume’s theory of personal identity is that the associative principles he uses to explain personal identity do not draw the boundary between one mind and another in the right place. In other words, Pears’ claim is that although these relations that bind perceptions in the mind might explain the mind’s sense of its own *unity* (that which holds a single mind together) for Hume, they cannot guarantee the unity (independence) of one’s mind from other minds. It cannot draw a boundary between the contents of two distinct minds. He calls this the problem of the necessary ownership of perceptions. The key to the problem of the necessary ownership of perceptions lies not in the relations which unify a mind but in, what he calls the criterion of identity of a sensation. Pears, in Wittgensteinian fashion, links the criterion of identity of a sensation to the association of a person’s mind with a particular body. The reason why someone’s perceptions could not have been someone else’s is that the identity-line of a sensation starts at the point where someone has it. He claims that, when an impression of sense occurs in one’s series of perceptions (bundle), it does not hold its place in the bundle by possessing any property, or by standing in any specific relation to

any other perception in the bundle. This suggests that the “sensation is anchored by anything other than the fact of its contingent occurrence at the point where it occurred.” Pears goes on to claim that, “nevertheless, there is a strong [Wittgensteinian] modal statement that can be made about it: It, *the very same sensation*, could not conceivably have occurred in any other series. This modal statement is based on the special character of the identity-line of a particular perception: the identity-line of a particular perception starts at the point where it occurs in a particular series, and no sense can be attached to the speculation that *it, the very same sensation*, might have occurred in some other series” (10). However, in Hume’s philosophy, the only possible relations among perceptions are contingent relations, which cannot satisfy the requirement of modal statement. The upshot is this: the three relations that, according to Hume, contingently connect the perceptions of a single mind cannot serve as the ground for the strong modal denial of alternative ownership. If each person’s perceptions were anchored by a separate substantial (nuclear) self, Hume’s inconsistency would have disappeared. But, of course, that is something unthinkable for Hume. So Hume’s problem, according to Pears’ explanation, is that he is in no “position to explain the strong modal statement that none of [his] perceptions could have conceivably belonged to anyone else” (11).

Pears’ explanation locates the inconsistency in the right place. It locates the source of Hume’s labyrinth in the failure of the associative principles—resemblance, causation, and contiguity—binding the bundle of perceptions in one single mind. In particular, Pears argues that Hume’s omission of the principle of contiguity, especially spatial contiguity, leaves Hume unable to draw the boundary between one person’s mind and another person’s mind. This omission creates the problem of the necessary ownership of the perceptions. Pears’ argument is certainly not a trivial one; it states something that Hume, as I will argue, might have in mind when he talks about the labyrinth in the Appendix. As I will show, later in this chapter, what Hume realizes on his second-thought about personal identity is the problem of the *incoherence* and *inconstancy* of the bundle of perceptions, which is essentially the problem of the criterion of identity. Pears’ explanation is concerned about the ownership of perceptions, whereas my account is concerned about the coherence and the constancy of bundles of perceptions.

There are three main problems with Pears’ explanation. First, he maintains that

Hume's simple, though fundamental mistake, is that his neglect of the body results in a lame account of the criterion of personal identity over time. In particular, he points out that Hume excludes the principle of contiguity in his psychological analysis of personal identity. He claims that Hume only uses temporal contiguity in a partial manner.

Although Hume's original account of personal identity somehow excludes the body, it seems unlikely that concerns over this problem has caused Hume to recant his theory in the Appendix. At various places in the *Treatise*, Hume asserts that his project is committed to a natural/psychological analysis of some of our fundamental beliefs such as the belief in necessary connection, the belief in the continued and distinct existence of the external world, and the belief in personal identity. It is fairly clear that Hume is not worried about the anatomical (physiological) concerns that some people, like Pears, have raised. Hume does not neglect the body; he thinks of it as degrading to philosophy to concern itself over the physiological aspects of the self. Furthermore, his rejection of Locke's memory criterion of personal identity is a clear indication of the fact that, in his view, the body cannot be considered as the criterion of personal identity. However, it is true that Hume's main concern in his second-thought is the problem of the criterion of personal identity.

Secondly, the other problem with this interpretation is that there is no indication in the Appendix or in the main text of the *Treatise* that Hume is attracted to any ideas of the necessary ownership of perceptions. He continues to stress in the Appendix that "all perceptions are distinct...[and] are, therefore, distinguishable, and separable, and may be conceived as separately existent, and may exist separately, without any contradiction or absurdity" (THN, 634). Lastly, Hume's distinction between the "vulgar" view about the continued and distinct existence of objects and the philosopher's view indicates that Hume was aware of the problem of the ownership of perceptions. Yet, he thought that the philosophical analysis of mental phenomena can curb that problem. According to him, the vulgar conceives of their perceptions as existing outside of the mind, at least for parts of their durations. In particular, the vulgar do conceive some of their impressions of sensation as such insofar as they cannot distinguish their impressions from bodies. As a result, they ascribe "continued and distinct" existence to their impressions. Nevertheless, Hume's philosophical analysis (the philosopher's view) disallows the idea of perception's

existence outside of the mind. Hume is aware of the problem of the ownership of perceptions and his analysis of mental phenomena is to curb that problem. We must note that the critical point is that Hume's philosophical analysis clearly indicates that Hume is aware of the problem of the ownership of perceptions and tried to account for it. This shows that the inconsistency that Hume raised in the Appendix cannot be about the necessary ownership of perceptions. Hume must have had something else in mind.

Separability versus Unity

Fogelin's interpretation concerns the *unity* and the *separability* of the bundle (three explanations and this is the one he prefers). He states that Hume's position demands a genuine/active self rather than a fictitious one. However, he is quick to point out that his reason for holding this position is different from scholars like Robison and Passmore. He claims that the initial appearance of Hume's theory that the mind as a heap or collection of perceptions, which takes place in the section of "Of Scepticism with Regard to the Senses," is invoked to help him out of a desperate situation. Fogelin thinks that Hume's doctrine of the self must be interpreted as a reply to Berkeley's claim that it is self-contradictory to assume that a perception can exist unperceived. More specifically, Hume must show that the vulgar's empirical mistake—that their perceptions sometimes exist without being perceived—does not involve any contradiction. Thus, he writes:

Hume's argument against Berkeley [for the possibility of unperceived perceptions] depends upon the notion of an *individual* mind from which a perception may be separated, but Hume provides no principle for individuating heaps of perceptions into minds. Strictly speaking, each perception is itself an individual substance and, again strictly speaking, a collection or heap of individual substances is not an individual substance. More remarkably, on Hume's principles, each perception is an individual *mind*, and a collection of minds is not itself a mind. Less strictly, for Hume, perceptions must be connected together loosely enough to allow separation, while at the same time they must be connected together closely enough to constitute a mind from which things can be separated. Hume's radical atomism guarantees the first result, but precludes the second. Without both features (separable perceptions and a unified mind), Hume's theory of perceptions no longer contains a response to

Berkeley's claim that it is self-contradictory to suppose that a perception can exist unperceived. (12)

In a nutshell, Fogelin argues that, for Hume, each perception is an individual mind, but the collection itself is not a mind. Therefore, the bundle lacks unity. This explanation has some merits of its own. Fogelin is quite right to insist that Hume needs both the unity or connectedness and the logical separability of the mind's perceptions in order to account

for his belief that some perceptions can exist unperceived. It is also true to say that Hume's atomism guarantees the logical possibility that some perceptions exist without being related to other perceptions. Moreover, it guarantees the logical possibility that some perceptions can exist unperceived ⁽¹³⁾. Finally, he is correct in saying that Hume's difficulty derives from his inability to explain the unity or connectedness of our perceptions.

Fogelin's proposal also encounters difficulties. One can argue that Hume is still unable to explain the 'real connexion' among our perceptions in the Appendix, even though he was satisfied about his earlier explanation. Garrett, for example, suggests two possibilities: (a) Hume does not provide any principle for individuating heaps of perceptions into minds; and (b) it is that "each perception itself is an individual mind (substance) and a collection or heap of individual substances is not itself a mind (substance)?" (THN, 233). In respect to (a), I may point out that Hume does have such principles, namely, the associative relations of causation and resemblance. In response to (b), it must be noted that, first, Hume does not use the notion of "substance" here in the restricted sense required by his general theory of abstract ideas. And second, it is not clear that a single perception is itself a mind. It is commonly understood that Hume characterizes the mind as a bundle or collection of different perceptions united together by certain relations. In fact, Hume's conception of the mind (self) is inherently a *relational* one. It is caused by the mere existence of different perceptions (bundle of perceptions) associated by the principles of causation and resemblance. To say that each single perception is an individual mind is to undermine this Humean principle. Hence, this raises serious doubts about Fogelin's claim that a single perception entirely unrelated to other perceptions constitutes an individual mind. This, in turn, undermines the whole explanation. In conclusion, although he is right to say that Hume needs both the unity and the separability of the mind's perceptions, it is not clear why Hume became dissatisfied with his earlier account of the unity of perceptions.

Real Connections and the Representation of the Self

This explanation belongs to Don Garrett. He claims that the inconsistency arises as a result of Hume's assent to the following jointly inconsistent propositions:

- A) All of our distinct perceptions are distinct existences.
- B) The Mind never perceives any real connection among distinct existences.

- C) The ownership of any perception is determined either by its causal relations and its relations of resemblance or by its perceived real connection to (i.e., inherence in) a distinct self.
- D) The causal roles of qualitatively identical objects (distinct existences) can differ only in virtue of differences in spatial or temporal location, unless the mind perceives a real connection between at least one of them and some other distinct object.
- E) Many kinds of perceptions are 'no where.'
- F) It is possible that two numerically distinct but qualitatively identical perceptions of any kind, including perceptions that are 'no where,' should occur in different minds at the same time (14).

Garrett argues that all these propositions are entailed by Hume's philosophy. He claims, in particular, that (A) and (B) are entailed by the Separability Principle and (C) is entailed by Hume's view that his theory of personal identity provides the only correct alternative to the theory that ascribes personal identity by relating our perceptions to a substantial self. Moreover, he argues that (D) follows from Hume's two definitions of "cause." In particular, Garrett claims that it certainly follows from Hume's first definition of cause, which implies that qualitatively identical perceptions can differ in their causal roles only in virtue of differences in their spatial or temporal location. Whereas, (E) is entailed by Hume's bundle dualism, and finally, (F) is a basic commonsense belief about the nature of persons—a belief that we expect Hume to maintain its truth. Garrett relates the truth of (F) to the question of other minds. According to him, in Hume's cognitive psychology, the truth of (F) derives from the fact that "an idea of a mind other than one's own is, presumably, structurally similar to one's idea of one's own" (15). Indeed, he says, "[T]o deny (F) would, after all, be to allow a "real" causal connection of sorts, because the existence of one perception would then be sufficient to entail the simultaneous *nonexistence* of any other of the same kind" (16).

The problem that these propositions pose is this: (A) and (B) are in clear conflict with (F). In other words, the Separability Principle is in conflict with our beliefs in other minds. Let's conceive of two qualitatively identical but distinct perceptions, P₁ and P₂, each in a different mind, A and B, which resemble each other and might also be temporally contiguous. It is clear that resemblance and temporal contiguity cannot secure the assignment of P₁ to A, and P₂ to B, because we have assumed that they are qualitatively identical. Therefore, the only other alternative is causation, which has to step in and secure the assignment of P₁ to A, and P₂ to B. It must be noted that what is essential to causation in the physical world is contiguity in space. That is to say, what is required here is a notion of "direct causation" which enables us to link separately the

cause of P₁ in A and P₂ in B. Yet (E) says that many such perceptions have no spatial location—they are “no where”—and they cannot differ in that respect either. The upshot of this is that P₁ and P₂ cannot be distinguished from each other by their relations of resemblance or causation. Logically speaking, causation without contiguity in space can link a single cause in A’s mind both with P₁ in A, and P₂ in B. It follows that the perceptions that are collected in Hume’s bundle by the associative principles cannot be secured into separate bundles as separate minds are secured in real life. So, the problem is that causation and resemblance are inadequate to bind sufficiently our perceptions in the way that is required by a “true idea” of the mind. Hume explicitly discusses the inadequacy of resemblance and causation as the source of the problem in the Appendix. Therefore, there must be some other principle to do the job for Hume. However, since Hume had already rejected the notion of spatial location, he has no choice but to see “inherence in mental substance and real necessary connections among perceptions as the only apparent ways out of his dilemma” (17). Otherwise he would not be able to explain personal identity in a way that will distinguish the contents of one mind from the contents of *other* minds having identical perceptions. As a consequence, a contradiction arises and Hume’s uneasiness in the Appendix is a belated recognition of this fact.

On Garrett’s explanation, Hume’s fundamental problem is the failure of his cognitive psychology. He claims that the “representational resources of [Hume’s] cognitive psychology are ultimately inadequate to the self-referential task of explaining the representation of one’s own mind as an entity potentially distinct from others” (18). This is because we cannot successfully “represent to ourselves the existence of qualitatively identical but numerically distinct perceptions existing in the minds of different individuals at the same time” (19). This results in *incoherence* in the bundles. Garrett briefly points out that Hume’s problem does not lie “in the fact that the *actual* existence of discrete causal bundles is simply causally inexplicable but in the fact, on Humean principles, the *possible* existence of such bundles becomes incoherent” (20). This is a brief exposition of Garrett’s interpretation.

This interpretation faces two major objections that are raised by Fogelin. First, it is said that Garrett fails to cite any textual evidence showing that Hume was even remotely worried about the question of *other* minds. Secondly, it is pointed out that the

kind of conceptual problems that Garrett has posed were quite alien to seventeenth and eighteenth centuries' philosophy, especially to Hume's philosophy. In particular, Fogelin argues that "Hume was uncritically wedded to the way of ideas. He thought that he had immediate access to his own ideas and he simply took it for granted that these ideas were *his*" (21). To the first, Garrett replies that no plausible account of Hume's difficulty can be anticipated in the main body of the *Treatise*. Furthermore, he thinks that his "explanation fits everything that Hume does say about the cause of his dismay in the Appendix: (i) that it renders his former opinions inconsistent; (ii) that it is an inability "to explain the principles, that unite our successive perceptions in our thought or consciousness"; and (iii) that it would be resolved "if our perceptions either inhered in something simple or individual, or the mind perceived some real connexion among them" (22). To the second, Garrett replies that though it is true that Hume thinks that he has immediate access to his ideas, Hume finds it problematic to determine why certain past ideas, which are no longer accessible to us in an immediate way, are still counted as "us." Moreover, he points out that Hume also finds it problematic to determine what the "he" who has this immediate access, actually *is* and how it is conceived in the imagination (23). After all, Hume's original account of personal identity is an attempt to explain the nature of that thing that we call an "I."

It is true that Hume never directly addresses the question of how our beliefs in the existence of other minds are to be explained. Nevertheless, he takes it for granted throughout of the *Treatise* that it is possible for us to be aware of others' thoughts and feelings. In fact, his account of *sympathy* shows that we can actually share the mental states of those we are sympathetic towards. Hume's descriptive account of sympathy provides the underlying cognitive structure for the formation of our beliefs in the existence of other minds. Recent scholars have convincingly shown that Hume does not appear to recognise any epistemological problem about the existence of other minds. It is said that there may be some concerns about the nature of our responses to the mental states of others; but there is no puzzle concerning the possibility of our awareness of what these mental states are. It is said that, on Hume's view, the character of our mental states in many cases reflects our awareness of other minds (24).

What we can gather from these interpretations is that an agreement is emerging

among Hume's scholars about his second thoughts on personal identity. They agree that Hume's difficulty arises as a result of his failure to recognize that the associative principles—resemblance, causation, and contiguity—are inadequate to unite our successive perceptions in our thought or consciousness. This cluster of explanations locates the source of Hume's inconsistency, one way or the other, in his failure to recognize the significance of “real connexion” and the simplicity of the mind. Of course, these explanations are supported by the text showing that Hume has explicitly expressed concerns over the connection between his original analysis of the self and the fact that such an analysis requires either real connections among perceptions or real simplicity and identity of the mind. Most notably, they refer to Hume's famous passage in which he claims that “[d]id our perceptions either inhere in something simple and individual, or did the mind perceive real connexion among them, there wou'd be no difficulty in the case” (THN, 636). Although it is not entirely clear what are Hume's real intentions here, it is pretty obvious that, in Hume's view, the mere existence of real connection among perceptions or perceptions inhering in something simple and individual could have saved him the inconsistency. Roughly, this constitutes the emerging agreement among Hume's scholars concerning his second thoughts about personal identity.

My preferred explanation concurs with the main theme of these interpretations. I maintain that the inconsistency emerges as a result of Hume's failure to recognize a “real connexion” among perceptions or that our perceptions inhere in something simple or individual. Therefore, we all agree that without “real connexion” among perceptions or simplicity and identity of the mind there is no way to bind perceptions into a single mind. Yet, my explanation diverges significantly from these interpretations with regard to the question of *what* caused Hume to recognize the fact that “real connexion” among perceptions or the simplicity and identity of the mind can bind perceptions into one single mind. I think that the existence of qualitatively identical but numerically distinct perceptions in *other minds* can cause incoherence in the bundles. Garrett has also located the source of Hume's inconsistency in his theory of belief in other minds, but he does not explain what particularly caused Hume to realize the inconsistency. In contrast, I suggest that the source of inconsistency lies in Hume's account of sympathy and the psychological mechanism determining our belief in other minds. Specifically, I will argue

that constancy and coherence fail to provide necessary and sufficient conditions for the specification of the criterion of personal identity, and as a result, we are left with no other mechanism to explain a coherent picture of the mind. In fact, constancy and coherence may explain a unified mind as far as the internal unity is concerned, but they fail to do so when the concern is on the unity across the bundles—what holds a bundle apart from other bundles.

My interpretation is comprised of two steps: First, I show that Hume has an account of the criterion (condition) of personal identity. I contend that Hume thinks that constancy and coherence provides necessary and sufficient conditions for the specification of the criterion of personal identity. Second, I argue that Hume's account of the criterion of personal identity fails because it cannot account for our belief in other minds. In the next section, I set out Hume's account of constancy and coherence as the condition (criterion) of personal identity. In the final section, I will show how constancy and coherence fail to account for our belief in other minds.

V.2: Hume's Criterion of Personal Identity: Constancy and Coherence

Contemporary analytic philosophy is divided into two main theories concerning the problem of the criterion of personal identity: the psychological theory and the physical theory. The psychological theory maintains that psychological continuity is required for personal identity, though it differs on the kind of causal mechanism that explains the continuity or persistence of the self over time. In contrast, the physical theory claims that personal identity must consist in the persistence of something physical, like an organism, or a brain. However, the two camps agree that the debate about personal identity is centred on a fundamental question: Under what possible conditions is a person existing at t_1 identical with a person existing at t_2 ? Or what are the logically *necessary* and *sufficient* conditions for a past or future person to be me? This question is called the Persistence Question. An answer to the persistence question is often understood to be the specification of the criterion of personal identity. A criterion is a set of logically necessary and sufficient conditions that determine whether temporally distinct stages of a person's life are stages of one and the same continuant person (25). To say that A and B are logically necessary and sufficient conditions for something, C, is to say that if A and B are present, it follows that C must be present as well, and if A and B are absent, then C must be

absent as well. It is said that any account of personal identity over time necessarily consists in specifying the criterion of personal identity over time. It is worth noting that the criterion of personal identity as such is concerned with what constitutes the identity of persons over time rather than the evidential element of personal identity. The constitutive element of personal identity is essentially metaphysical-cum-semantic, whereas the evidential element is merely empirical. Hume discusses both elements, but the current concern here is with the constitutive element—metaphysical-cum-semantic—alone.

Historically speaking, Hume never directly addresses the question of the criterion of personal identity in his philosophy. Neither does Locke, who gave the problem of personal identity its first identifiable formulation in the history of modern philosophy. The notion of the criterion of personal identity is a contemporary one. Contemporary philosophers have traced the notion of the criterion of personal identity to Locke's theory of identity. In fact, contemporary *psychological continuity* criterion of personal identity, which has dominated the field for quite some times now, has its roots in Locke's theory of identity. There are many contemporary philosophers who identify themselves as 'Lockean' or 'New-Lockean,' and subscribe to some form of the psychological continuity criterion of personal identity. Nonetheless, there is no much of a debate about Hume's criterion of personal identity. Some philosophers have even claimed that Hume is not concerned about the problem of the criterion of personal identity at all. For example, Noonan has claimed that Hume's account of personal identity is not concerned about giving an account of what *constitutes* personal identity over time. He thinks that Hume's account of personal identity does not involve the criterion of personal identity. On the contrary, he argues that Hume's problem is to explain the origin, or *genesis*, of the psychological causes of our universal but mistaken belief in the existence of enduring selves. "For, according to Hume, personal identity is a fiction," Noonan argues, "the ascription of identity over time to persons is a mistake. It is an explicable mistake, and one we all necessarily make, but none the less, it is a mistake. For persons just do not endure self-identically over time. Consequently since there is no such thing as personal identity over time, nor is there any problem of the metaphysical-cum-semantically variety presented by the question: in what does personal identity over time consist?" (26). He concludes that the only problem that exists is the *genetic* one of specifying the

psychological causes of the universal but mistaken belief in the existence of enduring selves, and this is the problem that Hume is concerned in his discussion of personal identity.

Noonan's genetic description of Hume's account of personal identity captures its fundamental feature, but it looks simplistic. He holds that Hume's project is simply to describe the psychological mechanism which induces our imagination to ascribe a fictitious identity to personas and objects, such as plants, animal bodies, and artefacts. This psychological mechanism by which the fiction of personal identity is caused is ubiquitous. It is responsible for our belief in the identity of such changeable objects as plants, animal bodies, and artefacts, and also for our belief in the continued and distinct existence of the external world. For Noonan, Hume's project ends here. He reduces Hume's general account of identity into the mere analysis of the psychological mechanism that explains the genesis of our beliefs in the identity of persons and objects. Nonetheless, I want to argue that Hume's programme goes beyond the mere genetic explanation of the psychological causes of our beliefs in an enduring self. I think that the psychological belief-determining mechanism that produces the idea of the self (Chapter III), particularly the principles of constancy and coherence supply Hume with the kind of identity conditions (criterion) that he needs for his theory of personal identity. John P. Write, for instance, remarked that "Hume seems to think of constancy *and* coherence as supplying individually sufficient and disjunctively necessary conditions of the impressions to which we ascribe an outer existence" (27).

Noonan's reading of Hume's account of personal identity is not a charitable one. It is true that Hume is strongly committed to a genetic explanation of the psychological causes of our belief in the existence of enduring selves. In fact, one can argue that his naturalism implies it. However, it seems that Hume's project is also aimed at giving a set of necessary and sufficient conditions which determine the identity of persons and objects over time. Hume's general account of identity is twofold: (a) he provides a naturalistic analysis of the psychological mechanism that explains the genesis of our belief in the "fictitious" or "improper" identity of persons and objects over time; and (b) he attempts to set out a set of necessary and sufficient conditions that constitute the identity of persons and objects over time. In this latter sense, Hume's project is aimed at specifying

the criterion of personal identity over time. It may turn out that Hume ultimately fails to provide a satisfactory account of the necessary and sufficient conditions of personal identity, but it does not rule out that fact that he intends to give such an account of the criterion of personal identity. My main goal here is to show that Hume's project offers more than what Noonan, for instance, has been arguing for.

I think that the justificatory status of the question of Hume's criterion of personal identity is determined by the status of his 'scepticism' versus 'realism.' Noonan seems to be espousing, more or less, a sceptical reading of Hume's philosophy in which we find a destructive project aiming at tracing back some of our most fundamental beliefs, or ideas, about the world to its origins in order to show that these beliefs have no proper foundations in our senses or reason. Hume is understood to be offering a detailed account of the psychological mechanism producing the 'fictional' ideas of, for instance, necessary connection, object and personal identity. The basic idea is that since these ideas have no foundations in the senses or reason, it shows that we must be in error to ascribe identity to persons and objects over time. And since these ideas neither are derived from sense impressions nor from demonstrative reasoning, we are left with the sceptical conclusion that there is no reason to believe that these ideas are *real* after all. There is no doubt that Hume is interested in this sceptical conclusion, but as Stroud has pointed out he is not primarily concerned to establish or defend the conclusion that we do not know or have any reason to believe that there is an external world, or that there is a necessary connection between cause and effect as well as a persistence self over time (28). Hume thinks that our beliefs in bodies, persons, and necessary connection are natural human phenomena, and the science of human nature is in charge of explaining them. So, in his view, there is an intelligible way to explain the origins of our beliefs in the existence of enduring selves and bodies as well as necessary connection. It is this positive aspect of his teaching that may ultimately explain his commitment to a criterion of personal identity. I want to explain this.

Recently, Loeb has forcefully argued that Hume's theory of belief is justified with reference to a stability-based theory of justification. On his account, "Hume prefers to determine the justificatory status of a belief with reference to generic properties of belief-forming mechanisms" (29). More directly, Loeb says that "beliefs are assessed

epistemically with reference to the underlying intellectual character or disposition that produces them. The justificatory status of a belief derives from the justificatory status of the mechanism that produces it... his assessments of justification depend upon the properties of kinds of belief-forming mechanisms, rather than features of individual beliefs” (30). Hence, for a belief to be justified it means that it is generated by a psychological mechanism, which puts us in a stable doxastic state, and justification, so understood, depends on those features of the mechanism that produces stability in belief. It seems that those features of the mechanism, particularly constancy and coherence, which put us in a stable doxastic state, would serve as necessary and sufficient conditions for justification. It is important to note that, on Loeb’s interpretation, Hume’s stability-based theory of justification is essentially a *normative* one. The normativity-claim seeks to establish “a set of *substantive* conditions that specify when a belief is justified” (31). In other words, when we say that a belief is justified we mean that there is a set of necessary and sufficient conditions that specifies its truth-conditions.

I find Loeb’s stability-based theory of justification enormously important. I want to suggest that it provides extremely helpful insight for us in order to define Hume’s theory of the self and its condition of identity over time. In particular, I claim that in the same way that the justificatory status of our beliefs derives from the justificatory status of the psychological mechanism that produces them, the justificatory status of the criterion of personal identity is derived from the justificatory status of the psychological belief-determining mechanism that produces it. What constitutes personal identity is, indeed, what constitutes the psychological mechanism causing the universal but mistaken (fictional) belief in the existence of persisting selves. Subsequently, what constitutes that psychological mechanism is how it justifies epistemically our fictional belief in personal identity with reference to the underlying intellectual character or disposition that produces it. And I think that Hume’s principles of constancy and coherence constitute that underlying intellectual character or disposition of the mind that produces our belief in personal identity. In this respect, the psychological belief-forming mechanism, particularly the principles of constancy and coherence, must explain Hume’s account of the criterion of (what constitutes) personal identity.

On Loeb’s interpretation, Hume has a two-stage theory of justification: a

constructive stage (roughly Hume's Naturalism) and a destructive stage (roughly Hume's skepticism). In his constructive epistemological project, Hume tries to sustain his pre-reflective beliefs—causal inference, the continued and distinct existence of external objects and the self—with reference to a stability-based theory of justification. His constructive project, however, gives rise to a destructive result. He realises that these pre-reflective epistemological commitments cannot be sustained because his theory of justification does not adequately explain them. Nevertheless, this does not mean that Hume abandons his theory of justification. Rather, he just wants to abandon the epistemic distinction that he has made between his pretheoretical beliefs in causal inference or the idea of the self, for instance, and his naturalistic (post-empirical) conclusion that these ideas are the results of a psychological belief-forming mechanism. In other words, Hume comes to the conclusion that his theory of justification, which is a normative theory, fails to justify such beliefs only for the reflective person, namely, the philosopher. But the vulgar, “all the unthinking, and unphilosophical part of mankind, that is, all of us, at one time or the other” (THN, 205), still holds on to those pre-reflective ideas, though with no epistemic commitment. Nonetheless, these pre-reflective beliefs somehow must be justified, and Hume seems to be involved in giving certain set of conditions in light of which we may justify such beliefs. Loeb claims that these pre-reflective beliefs could be justified by the specification of certain nonepistemic psychological properties that are basic to Hume's theory of justification, namely, nonepistemic doxastic stability (32)

Similarly, Hume's own answer to the question of what constitutes personal identity has a negative phase and a positive phase. According to the negative phase, personal identity is neither constituted by the identity of a simple material or immaterial substance nor it is constituted by the identity of a simple impression. On the other hand, the positive phase tells us that personal identity is constituted by some form of psychological-functional continuity. We may identify the negative phase of Hume's account of what constitutes personal identity with, what Noonan calls, Hume's genetic specification of the psychological causes of our beliefs in enduring selves. However, Hume's project goes beyond a genetic explanation of the idea of personal identity; he assigns a normative role to the psychological mechanism that produces nonepistemic doxastic stability. Stability as such is an outgrowth of his theory of natural belief. For

Hume, stability is both necessary and sufficient to justify beliefs. Loeb remarks that “we need to locate a property that is *necessary* for the states produced by a psychological mechanism to constitute beliefs, such that to establish that the states are beliefs and thus have this property is also *sufficient* to establish that the beliefs are justified” (33). Stability is said to be that property that constitutes a necessary condition for beliefs, on the one hand, and establishes a sufficient condition to justify beliefs, on the other.

Stability requires a framework that can accommodate it as a property of belief. It seems that mental dispositions or natural propensities of the mind provide such a framework in which stability and steadiness is construed as a property of belief (34). There is no doubt that an essential part of such natural propensity or disposition is those features of our perceptions, namely, constancy and coherence, which produces stability in belief. In short, the stability-based theory of justification claims that the states produced by a psychological mechanism are beliefs, and insofar as they are infixed, it *sufficiently* establishes the fact that beliefs are stable, and thus justified. The stability-based theory provides a general schema for the justification of belief, and an explanatory pattern for the justification of beliefs in objects and persons in particular. Such a pattern of psychological explanation appeals to a natural propensity, or disposition, to ascribe identity to a succession of related perceptions in order to explain the beliefs in the continued and distinct existence of external objects and the self insofar as those beliefs arise from constancy and coherence of the series of perceptions.

There is ample evidence that, for Hume, *steady dispositions* of the mind give rise to beliefs. He quite often talks about ‘vivacity,’ ‘intensity,’ and ‘liveliness’ of our ideas and that they generate ‘firm,’ ‘steady,’ ‘fast,’ ‘firm,’ ‘settled,’ ‘solid,’ and ‘steady’ beliefs (THN, 97, 106, 108, 116, 624, 625, 626, 627, 629, 631). Hume’s conception of stability and steadiness is closely related to his conception of the way that ideas are *infix*ed. On a number of occasions, Hume says that ideas are infixed with force, and vigor (THN, 86, 109, 225), and that enlivened ideas as such produce beliefs. Hume has famously said that belief is a lively or a forceful idea (THN, 96). Moreover, he claims, “we may mingle, and unite, and separate, and confound, and vary our ideas in a hundred different ways; but ‘till there appears some principle, which fixes one of these different situations, we have in reality no opinion” (THN, 96). Hence, for Hume, infixing ideas is the necessary condition

for having an idea. Hume specifies a mechanism that shows how infixing ideas produces steadiness. The most important element involved in infixing steady beliefs is custom and habit. He claims that “custom and habit having...the same influence on the mind as nature, and infixing the idea with equal force and vigour” (THN, 86). ‘Custom’ and ‘habit’ is identified with “every thing...which proceeds from a past repetition, without any new reasoning or conclusion” (THN, 102). Hume observes elsewhere that a “principle [that] has establish’d itself by a sufficient custom, it bestows an evidence and firmness on any opinion, to which it can be apply’d” (THN, 105). Furthermore, the constant conjunction between cause and effect as well as the senses and memory give rise to firmness or steadiness in belief (THN, 109, 86). We may conclude that repetition, particularly custom and habit, generates stability and steadiness in belief. Overall, infixing an idea is essential for having a belief. Custom or habit is a process that results in steadiness, and steadiness gives rise to belief.

To sum up, Hume’s idea of the self and the idea of external objects are produced by the operations of the imagination. In particular, he thinks that certain characteristics of the series of our impressions, namely, constancy and coherence, can account for these ideas. The central point is that constancy and coherence form a complex psychological pattern according to which these ideas are produced. This pattern, which is derived from the features of our perceptions, is a method of rendering our loose and separate perceptions into a unified entity in which we get a coherent picture of the reality. In fact, constancy and coherence make sure that we get that picture; that is to say, constancy and coherence operate as the condition (criterion) we need to have such a unified coherent picture. Therefore, Hume thinks that constancy and coherence provide necessary and sufficient condition for personal identity. However, I think that constancy and coherence fail to provide necessary and sufficient condition for personal identity. I turn to this now.

V.3: The Failure of Coherence and Constancy and the Belief in Other Minds

The inconsistency arises as a result of the failure of *coherence* and *constancy* and the belief in *other minds*. As explained above, coherence and constancy, as certain characteristics of the series of impressions, in conjunction with the imagination, *function* as the basic elements of a psychological belief-determining mechanism responsible for generating the beliefs in the continued and distinct existence of external objects and the

belief in the self. According to Hume's psychological explanation, the ability to think of ourselves as unified mental subjects persisting through time and our belief in the endurance of objects are caused by a *feeling* of sliding along a train of successive related perceptions. Constancy and coherence are directly responsible for producing such a feeling and the belief that it gives rise to. It is an essential part of Hume's theory of mind, however, that the generation of the belief in objects and persons occurring as a result of the functioning of coherence and constancy happens *within* a discreet mind—that is, a closed mental system excluding other minds. To put it differently, it is dependent upon the physical setting that makes its realization possible. It is my contention that Hume's principles of constancy and coherence can successfully account for the arising of belief in the continued and distinct existence of external objects and persons within a discreet mind, but fails to explain the belief in other minds (selves).

Hume claims that we are aware of the contents of other minds, though he never discusses explicitly the belief in other minds. He puts forward an account of sympathy in Book II of the *Treatise*, claiming that we come to know the mental states of others. Accordingly, when one sympathizes with others one receives, by communication, the original passions, inclinations, and sentiments of other minds. Hume characterizes sympathy as “that propensity we have to sympathize with others, and to receive by communication their inclinations and sentiments, however different from or even contrary to our own” (THN, 316). Hume's reference to communication here means that sympathy is a process by which the contents of other minds are somehow transferred to us. There is no indication in Hume's text that he treats the question of other minds, or sympathy for that matter, as a case of causal inference. It is safe to suggest that, for Hume, the cause (s) of the belief in other minds does not lie in reason or senses; rather, it is a complex process of the associative relations of causation and resemblance that leads us to acquire the sentiments and inclinations of others. It is the imagination that makes that transition possible. Hume's reference to the imagination is explicit in the text. In his discussion about our esteem for the rich and powerful and our contempt for meanness and poverty, he states that “the principle of *sympathy*, by which we enter into the sentiments of the rich and poor, and partake of their pleasure and uneasiness” (THN, 362) is the only principle that gives us an esteem for power and riches, and a contempt for poverty. Then,

he goes on to say that “riches give satisfaction to their possessor; and this satisfaction is convey’d to the beholder by the *imagination*, which produces an idea resembling the original impression in force and vivacity. This agreeable idea or impression is connected with love, which is an agreeable passion. It proceeds from a thinking conscious being, which is the very object of love. From this relation of impressions, and identity of ideas, the passion arises, according to my hypothesis” (THN, 362; emphasis added). It appears, then, that the imagination plays a central role in Hume’s account of sympathy and the belief in other minds, and the text clearly indicates it.

Hume’s account of sympathy shows that the communication of the mental states of one person to another is explained in terms of a psychological mechanism. According to such mechanism, one acquires an idea of another person’s mental states through the effects of her appearances and behaviours, and then the idea that is thus acquired is converted into an impression. Such an idea triggers an existing resembling impression of reflection in one’s mind, and it is from the enlivening, forceful, and vivid effects of the resembling impression produced in such a manner in one’s mind that one becomes aware of the mental states of another person, or comes to have belief in other mind. This belief is produced more or less by the same psychological belief-determining mechanism, which produced the beliefs in the existence of objects and persons. In particular, resemblance, mental causation, and the imagination are all operative in both cases. To our surprise, coherence and constancy do not play any role in the production of the belief in other minds. The task of explaining the belief in other minds is given exclusively to resemblance, mental causation and the imagination. I think that this move is problematic.

Hume describes constancy and coherence as the key elements of his novel psychological belief-determining mechanism, which explains the beliefs in the continued and distinct existence of external object and the self. Hume conceives coherence and constancy as the criteria that provide necessary and sufficient condition for the identity of objects and persons. If we take into consideration the fact that Hume’s account of sympathy in Book II shows that he views the belief in other minds as important as the belief in external objects and persons, it follows that the same psychological mechanism, particularly constancy and coherence, should account for one’s belief in other minds. Nevertheless, coherence and constancy do not play any roles in the explanation of our

belief in other minds. The absence of constancy and coherence in the explanation of the belief in other minds shows that not only the belief in other minds is not explained adequately by Hume's psychological belief-forming mechanism, but also the absence of constancy and coherence undermines even their roles as the necessary and sufficient condition for our beliefs in objects and personal identity—beliefs that are generated by the same psychological belief-determining mechanism. In a nutshell, in Hume's psychological explanation of belief, different elements including constancy and coherence are operative, and that Hume may have realized later that constancy and coherence cannot provide the necessary and sufficient conditions to explain those beliefs. This problem becomes more pressing, if we note that, on Hume's account, coherence and constancy and the way that they affect the imagination are understood to be the cause of most of our fundamental philosophical beliefs. In fact, the picture of mind that Hume presents in Book I of the *Treatise* is incompatible with his account of mind in Book II.

The central claim is that when we conceive of two discrete minds as having qualitatively identical but numerically distinct perceptions we lack a proper mechanism, like the one operating in the cases of the belief in objects and the self, to distinguish properly the contents of one's mind from those of others. On this view, no discrete bundle could exist. Quite possibly, the qualitatively identical perceptions could be flouting freely in various distinct bundles while Hume is unable to come up with a viable explanation to account for the unity of the bundles. This seems to be a serious blow to Hume's bundle theory of the self. The way out of this dilemma is either to grant spatial location for perceptions or to assign 'real connexion' among perceptions or attribute simplicity and identity for the bundles in order to individuate the qualitatively identical perceptions existing in different minds. Hume, however, rejects the notion of spatial locations for the minds, the real connection among perception, and simplicity and identity of the mind altogether. The outcome of Hume's rejection of spatial location, real connexion, and simplicity of mind is that he finds himself unable to secure the discreteness of different minds. He cannot explain how the contents of one's mind are distinct from the contents of other minds.

Hume's account of sympathy creates this problem. Accordingly, we become aware of the mental states of others in such a way that "the sentiments of others can never

affect us, but by becoming, in some measure, our own” (THN, 593). Hume makes it clear that the kind of transition by which we are carried from our own mental states to the contents of other minds is a genuine awareness of their mental states, and not a mere recognition of their thoughts, passions, and feelings. An indication for his claim is that he sets out his account of sympathy in terms of his theory of association of ideas. He shows how an idea, which one receives by observing the behaviours of another person, is converted into an impression, and then by the lively and forceful effects of that impression one comes to share the other person’s mental states. This suggests that Hume has a psychological mechanism, like the one operating in the cases of objects and personal identity, which accounts for the transition of the mental states of others to our own. However, such psychological mechanism, particularly constancy and coherence, fail beyond the scope of a discreet mind. This claim calls for further investigation.

The mind properly understood, consists of a temporal succession of perceptions connected by the natural relations of resemblance, contiguity, and causation. Each perception in that temporal succession of perceptions exists distinctly, and therefore, is separable from the rest. Moreover, Hume holds that there is no real connection among these successive perceptions. Hence, there is nothing to secure unity of the mind; and the idea of the self is merely a ‘fictitious’ one produced by the mere operation of a psychological belief-determining mechanism. As explained earlier, Hume conceives both mental and external objects as fragmentary, episodic, interrupted, and discontinuous entities. The mind’s key function is to generate a continuous, uninterrupted, and unified whole out of such fragmentary and episodic objects. Recall the formality I used in Chapter III to explain the temporal and episodic nature of Hume’s succession of related perceptions:

When I fix my gaze on an object, say, a tree, my perceptual experience goes like this:

1) A

When I close my eyes or turn my head for a second there is only an interval:

2) □

When I fix my gaze upon the same object after the interval, my perceptual experience goes as follows:

2) Á

My real perceptual experience of the three different stages is like this:

4) $A \square \hat{A}$

In (1) there is an uninterrupted and invariable perception. The object of my perception, the tree, and the perception itself is in a state of “perfect identity.” There is no interruption and the object is continuous; the object is the same with itself. It is one episode where “[w]e have a distinct idea of an object, that remains invariable and uninterrupted thro’ a suppos’d variation of time; and this idea we call that of *identity* and *sameness*” (THN, 253). This is entailed by Hume’s *principium individuationis*, or principle of identity. In (2) we are not having any perceptual experience. When I shut my eyes or turn my head the object of my perception, the tree, along with the perception itself, vanishes altogether. In the interval an interruption occurs, which results in the episodic disappearance of the objects. The interruption indicates the fact that the object of my perception in the interval does not exist. It is an episode of the fragmentation of the object of my perception. In (3) I fix my gaze upon the same object after the interval and observe it again. But the object of my perception, $\hat{A}$, is not strictly identical with A due to the occurrence of interruption in (2). Here, I experience a new episode or fragment of the previous object, though a resembling one. In (4) we have the entire perceptual experience of the object. (4) shows the fragmentary, episodic, and discontinuous nature of external objects and our perceptions of those objects; it exhibits constancy in its full force. It shows that constancy among the series of impressions (perceptions) generates the feeling of continuity in the series out of the interrupted and discontinuous successive related perceptions. Hume draws the conclusion that there is neither “perfect identity” or “simplicity” in objects nor in the mind. Nevertheless, Hume argues that there is a natural propensity in us that tends to ascribe identity to objects as well as persons. In fact, Hume contends that if a perceiver is presented with episode (4), s/he would experience it like episode (1). There is a natural universal propensity or inclination in the mind that tends to ‘confuse’ or ‘mistake’ episode (4) with (1); that is, to regard a temporal succession of related objects (perceptions) as a single and continued object. The identity of objects and persons are mere fictional inventions of certain natural features of the series of impressions—constancy and coherence—and the way that they affect the imagination. Hume offers an original and novel psychological explanation of our propensity to confuse

or mistake the interruption for identity.

As I have explained in Chapter III, Hume distinguished four different successive stages that the mind went through to produce our beliefs in objects and persons. These stages were the following:

(1) When we observe constancy in successions of related perceptions that are interrupted, it triggers a psychological reaction according to which we ascribe identity to the successive related perceptions due to a ‘propensity’ or ‘inclination’ that mistakes or confuses the resembling but successive series of perceptions with one constant and uninterrupted perception.

(2) Yet, the palpable interruptions in the successive related perceptions induce us to consider constant but interrupted perceptions as diverse and different (THN, 199, 205, 220, 254). We take constant but interrupted perceptions as different because identity in the strict sense requires uninterruptedness and invariableness of perceptions through a supposed variation of time.

(3) The fact that we take constant but interrupted perceptions both as identical and diverse, we find ourselves at a loss, or involved in a kind of contradiction (THN, 199, 205). To reconcile this contradiction, we come to believe that objects continue to exist even when they are not perceived. In fact, we disguise the interruption or “rather remove it entirely, by *supposing* that these interrupted perceptions are connected by a real existence, of which we are insensible” (THN, 199, 205, 254). We postulate an ‘unperceived perception’ filling the gap between the series of interrupted perceptions. So, we take an interrupted series of successive impressions as an uninterrupted series.

(4) When we unite the interrupted and broken perceptions by a continued existence we come to believe that objects continue to exist when they are not perceived.

In (4) above, when the mind is experiencing $\hat{A}$, it is in a similar disposition, or state, as when it is in A . However, Hume explicitly argues that when the mind is the state of the interval, $\square$, it “makes us regard the first impressions, A , as annihilated, and the second, $\hat{A}$, as newly created” (THN, 199). Of course, there is no impression when the interval, $\square$, occurs. As explained above, constancy gives rise to the idea and the feeling of a continuous series of impressions in (4). Now, the question is from where is this lively feeling of continuity derived? Or from where does it get its force and vivacity, if the first

impression is entirely annihilated and the second impression is newly created? We must note that it is the force, liveliness, and vivacity of the impressions that generate the belief in its continuity. Hume argues that the force, vivacity and liveliness that exist when we are having the impression of A is smoothly transmitted or transferred (Hume uses both verbs) to the state of interval, $\square$. For Hume, memory plays a crucial role in making this transition possible: “This supposition, or idea of continu’d existence acquires a force and vivacity from the memory of these broken impressions, and from that propensity, which they give us, to suppose them the same” (THN, 199). The memory of the first impression, $\ddot{A}$, which is ontologically distinct and separate from the impression itself, A, triggers the idea of the first impression conveying its force and vivacity to the state of interval. According to Hume, ideas that are faint images of impressions lose their force and vivacity, but certain ideas when triggered bring about the original force and vivacity of the first impressions. It seems that this is a case in which memory triggers the idea of the first impression and, by doing so, it brings about its original force and vivacity and, then, transmits that force and vivacity to the state of the interval. That force and vivacity gives rise to what Hume calls an ‘unperceived perception’, or an ‘insensible object.’ The reason that such transition takes place is that when the mind experiences $\square$, the disposition that it is *in* is almost the same as when the mind is in the state of A; at least the perceiver feels it that way. This feeling produces the belief in the continuity, and thus, the identity, of temporal succession of perceptions.

Hume thinks that this four-stage psychological belief-determining mechanism or system (to use Hume’s own language), as outlined above, provides the most profound explanation of our belief in the existence of objects and persons. Most notably, such a psychological explanation reveals the ‘fictitious’ or ‘constructed’ nature of our beliefs in the identity of objects and persons. In this sense, the material world and the self are equally fictions; they are the imaginative constructions of the mind. According to Hume, constancy fills the gap among interrupted individual perceptions by interpolating an unperceived perception within the series of resembling perceptions, and thus joins them together. Similarly, coherence fills the gap between two different series, or more, of perceptions. As I argued in Chapter III, Hume holds that the principles of constancy and coherence are specified to explain certain kinds of regularities within the series of

perceptions. Roughly, constancy accounts for the identity of objects that undergo gradual but constant changes, whereas coherence explains the identity of those objects that their position and qualities are changed in such a manner that they have become hardly knowable (THN, 195). Coherence is supposed to account for those cases that the objects have undergone considerable or total alterations, but we still identify them as the same objects. For instance, Hume claims that “[w]hen I return to my chamber after an hour’s absence, I find not my fire in the same situation, in which I left it: But then I am accustom’d in other instances to see a like alteration produc’d in a like time, whether I am present or absent, or near or remote” (THN, 195). The fire that Hume observes upon his return to his chamber is not the same as when he first observed it. The fire has changed considerably in terms of qualities and spatial location. The other example that Hume gives in the main text of the *Treatise* sheds more light on his account of coherence. There he talks about the noise of a door turning upon its hinges, and a little later seeing a porter who brings a letter from his friend, two hundred leagues distant. In this example, we have different series of impressions that are qualitatively distinct but somehow exhibit a pattern of regularity among dissimilar perceptions. Obviously, when he hears the noise of a door turning upon its hinges and then sees a porter a little later, he assumes that the door has existed all along and that it has now been opened. He also comes to believe that the stairs that the porter used to get upstairs are still in being. Moreover, when he sees his friend’s hand-written letter, he spreads out before his mind the whole sea and continent between them, and supposes the effects and continued existence of posts and ferries (THN, 196). These different series of perceptions are qualitatively distinct but somehow connected. In short, for Hume, constancy is intended to account for the existing gap in the series of impressions that consists of qualitatively identical but numerically distinct perceptions, while coherence is supposed to explain the gap between different series of perceptions that are not identical in terms of qualities and spatial locations.

Hume’s psychological explanation of our belief in the existence of objects and persons enjoys a degree of success. However, it seems to fail on a crucial moment. Hume is relatively successful in explaining the belief in objects and persons, but his resources are inadequate for explaining the belief in *other minds*. Constancy and coherence are not apt to account for one’s belief in other minds. Strictly speaking, they are specified to fill

the gaps among the individual series, or different series, of impressions that subsist within a discrete mind. Supposedly, if there were any individual series of impressions, or different series of impressions, between two discrete minds, constancy and coherence seem unable to explain the regularity pattern existing among them. In fact, when Hume discusses the belief in other minds through his account of sympathy, he suggests that the imagination makes possible the transition of one's mind to those of others. Hume's account of sympathy is stated within the context of his theory of the association of ideas. Particularly, he is explicit in distinguishing between the idea belonging to other minds and the related impression in one's mind. It appears that this is a case of a succession of related ideas and impressions, and that this succession of related perception in turn gives rise to belief in other minds. It is quite clear that all the elements involving the generation of the belief in the existence of objects and persons are present here as well: the succession of related perceptions, the lively feeling and the imagination. Yet, constancy and coherence, which play the central role in the explanation of the belief in objects and persons, are entirely omitted from Hume's psychological explanation of belief in other minds. This can pose a severe problem for Hume's general theory of belief, which in turn accounts for the inconsistency.

Hume strongly believes that we can have belief in other minds (through sympathy). Hume has a rich theory of belief in other minds, though he does not explicitly address it. Hume has a theory of belief in other minds that parallels with his own theory of belief in the continued and distinct existence of objects and the self (35). In the same way that our belief in the existence of external objects and the self is reflected in our responses to our environment, the belief in other minds is also implicit in our responses to other minds around us. The fact that Hume attempts to account for the belief in other minds, through the effects of sympathy, indicates that he considers the belief in other minds as significant as the beliefs in the existence of objects and persons.

In Hume's view, our belief in other minds is neither derived from the senses nor from reason. It is not derived from the senses because other minds are different from observing different human beings. It is largely agreed that Hume is committed to a form of property dualism. He makes a categorical distinction between the physical kind and the mental kind. The mind as a collection of different but related perceptions is to be

distinguished from the body upon which the mind is causally dependent. As I have shown (Chapter I), Hume holds that some mental objects lack extension and spatiality; a moral reflection or a smell or sound cannot be placed or said to be circular or square; it would be absurd to ascribe thickness, length, and breadth to our desires. They are incompatible with extension and local conjunction with matter and body. The implication of this for one's belief that there are other minds or selves is that such beliefs cannot be derived from the observation of the bodies that other minds are causally dependent upon. For Hume, the mind is ontologically distinct from the body. Therefore, my belief in the existence of other minds cannot be derived from observing other human beings, the biological kind that they are. In other words, our senses cannot tell us anything about the cause (s) of our belief in other minds.

Likewise, reason cannot explain the cause (s) of our belief in other minds as well. Hume distinguishes between two different categories of reason: demonstrative and probable. He makes his crucial distinction between knowledge and probability by differentiating between seven different philosophical relations. For Hume, demonstrative reasoning is described as a set of logical relations that depend upon the ideas that we compare together. On the contrary, probable reason is said to be a set of relations that is factual and contingent in the sense that they are affected by the manner or order in which ideas appear in the mind (THN, 69-70). When I say that reason cannot explain the cause (s) of the belief in other minds, I mean that neither probable reasoning nor demonstrative reasoning can account for our belief in other minds. It is obvious that demonstrative reasoning cannot produce the belief in other minds because it essentially involves the so-called *relations of ideas*, which is the rational foundation of "the sciences of Geometry, Algebra, and Arithmetic; and in short, every affirmation which is either intuitively or demonstratively certain" (EHU, 25). The belief in other minds, however, cannot be affirmed either by intuitive or demonstrative reason. On the contrary, it is quite possible that the belief in other minds involve *matters of fact*. In other words, the belief in other minds might be the product of probable reasoning insofar as it depends upon some form of causal inference. However, it seems that the belief in other minds cannot be the product of causal inference.

According to Hume, causal inference is the product of constant conjunction of

two different objects, or events, in which we infer the occurrence of an unobserved object or event from the occurrence of an observed object. Now, if we claim that the belief in other minds is the product of causal inference, it means that such inference is made available by the experience of the customary constant conjunction of cause and effect. However, that does not seem to be the case here because Hume never mentions that the belief in other minds is the product of causation; rather, he claims that it results from resemblance. So, I conclude that reason cannot produce the belief in other minds. In particular, the customary constant conjunction of cause and effect (causal inference) cannot account for our belief in other minds.

Alternatively, I ascribe the production of this belief to the imagination. The belief in other minds must be generated by some other principles; for instance, it could be generated by the imagination or by the same psychological belief-determining mechanism responsible for generating our belief in objects and persons. Hume, however, never mentions that the belief in other minds is produced by this mechanism. Instead, he tries to accommodate the deficiency of his account by introducing the notion of sympathy as a medium through which we come to know that not only there are other minds, but also that we can have knowledge of their contents and functionings. This implies two important things: First, our belief in the existence of other selves is as significant as the belief in the existence of objects and persons. Hume must have an explanation for the belief in other minds in order to avoid a Cartesian solipsistic conception of the mind. Secondly, Hume's failure to provide an account for his belief in other minds in terms of his psychological belief-determining mechanism results in the failure of such explanatory mechanism altogether. For example, Pitson remarks that "one could simply regard Hume's position here as a mark of his failure to recognize that the other minds belief needs to be accounted for, just as much as the other beliefs with which he is concerned in Book I of the *Treatise*. Perhaps, the fact that he does not attempt any such account might even be ascribed to an unspoken recognition that it would inevitably end in failure" (36). It is just this kind of failure that I have been attributing to Hume. If I am right, Hume's account of sympathy is inadequate to explain our belief in other minds.

Hume's theory of belief in other minds is contained in his account of sympathy. He characterises sympathy as the communication of others inclinations and sentiments, a

kind of process through which the mental states of others are somehow transmitted to us (THN, 316). “The sentiments of others can never affect us,” Hume argues, “but by becoming, in some measure, our own” (THN, 593). It is said that “the minds of men are mirrors to one another” (THN, 365). Hume lays out his account of the psychological mechanism of sympathy as follows:

When any affection is infus’d by sympathy, it is at first known only by its effects, and by those external signs in the countenance and conversation, which convey an idea of it. This idea is presently converted into an impression, and acquires such a degree or force and vivacity, as to become the very passion itself, and to produce an equal emotion, as any original affection. (THN, 317)

When we become aware of someone’s sentiments, passions, and inclinations by the effects of his gesture, appearance, or behaviour, the mind immediately passes from these effects to its causes and in this way forms a lively idea of the person’s state of mind (THN, 575). The idea acquired in this way gives rise to an impression. This impression, which is naturally a vivid and lively perception, generates a ‘feeling’ that eventually gives rise to the belief in other minds. The belief that arises in this manner in the existence of other minds belongs, like the belief in the existence of objects and persons, to the category of ‘natural belief.’ Hume’s notion of sympathy is not a cognitive process of causally inferring the mental states of others from their behaviours and utterances ⁽³⁷⁾. Instead, sympathy that triggers “a complex process of association tends to occur as a result of which one is led almost irresistibly to acquire sentiments corresponding to those experienced by these others” ⁽³⁸⁾. Hume clearly states that the transition by which the mind is carried from its own perceptions to those of others is one that is made by the imagination. Hume formulates his account of sympathy “in the language of his theory of association and, in particular, the notion that an idea may be converted into an impression through the enlivening effects of a related impression” ⁽³⁹⁾. Whether or not this is a plausible account of the belief in other minds, it clearly suggests that it would arise from the association of ideas. Indeed, here we have the same kind of naturalistic relation among perceptions that we had in the case of the belief in objects and persons. Consequently, the same psychological belief-determining mechanism must be at work here as well. However, such mechanism, particularly constancy and coherence are not included in Hume’s naturalistic account of sympathy. The pressing question now is this: why does Hume not include constancy and coherence in his account of sympathy?

Noticeably, Hume never raises this concern in the main text of the *Treatise* or elsewhere. Although he classifies our belief in other minds as natural belief, it turns out that such belief is not strictly generated by constancy and coherence. Constancy and coherence cannot postulate the existence of an unperceived perception to fill the gap between the interrupted series of successive perceptions that exists between one's mind and other minds. It seems that constancy and coherence are constrained to operate within a discreet mind. Strictly speaking, they are specified to postulate the existence of an insensible object operating upon the contents of one's mind. Let's look at the formal rendering. The standard rendering of an interrupted series of perceptions within a mind was like this:

1) $A \square \acute{A}$

The pattern of rendering successive related perceptions between two distinct minds is as follows:

2) $M_1 \square M_2$

As explained above, in (1) the vivacity and liveliness that exist when we receive the impression of A is transmitted to the state of interval, $\square$. Memory played a crucial role in making this transition possible. It was suggested that memory triggers the idea of the first impression and by doing so brings about its original force and vivacity and, then, transmits that force and vivacity to the state of the interval. That force and vivacity gives rise to an 'insensible object.' The reason that such transition took place was that, when the mind experienced $\square$, the mind's disposition was almost the same as when it was in the state of A. That feeling was responsible for generating the belief in the continuity or identity of temporal succession of perceptions. It is very important to note that it was the memory of the first impression, A, that was responsible for transmitting the force, vivacity and liveliness of the state of A to the state of $\square$. On the contrary, when the mind is in the state of $\acute{A}$, it cannot transmit its force and liveliness to the state of $\square$. The reason for this is not clear because Hume does not discuss it. Perhaps the reason is that it does not involve memory. However, in (2), when M_1 (mind₁) somehow triggers an idea in M_2 (mind₂) through the mechanism of sympathy, that idea converts (Hume's own word) into an impression. This impression lies at the origin of the lively feeling and the subsequent belief in the existence of other minds. Note that, the impression in M_2 , unlike (1), in

which the impression responsible for the production of the belief in objects and persons was triggered by the memory of the first impression, A, is triggered by the qualitatively identical idea in M₂. Here, unlike (1), where the gap was filled by the memory of the first impression, the gap between M₁ and M₂ is filled strictly by the conversion of ideas, originated in M₂, into impressions within M₂ itself. Apparently, the conversion of ideas into impressions in M₂ is made possible by its memory. But it is strictly constrained to the contents of M₂ alone. But what is the implication of this?

It seems that situation (2), like situation (1), is constituted by a series of resembling impressions, but situation (2), unlike (1) which is a series of resembling impressions existing in one discrete mind, is composed of a series of resembling impressions existing in two distinct minds. There is no doubt that, for Hume, the occurrence of a qualitatively identical impression in M₁ is the origin of the belief in other minds in M₂. When M₂ observes certain behaviours or utterances of M₁ it acquires an idea of M₁'s state of mind, and then the idea which thus acquired triggers related (resembling) impression of reflection in M₂. The key point is that there is a series of resembling impressions constituted by qualitatively identical impressions existing in two distinct minds. In other words, there is an impression in M₁, which produces behaviour or utterance, and that behaviour or utterance, in turn, triggers an idea in M₂, but what generates the belief in other minds is not the idea that it receives, but the mere existence of two qualitatively identical impressions in two distinct minds. I suggest that, for Hume, these qualitatively identical impressions in two different minds create a series of resembling impressions and this series in turn produces the belief in other minds. I also suggest that this series of resembling impressions functions quite similar to the typical series of impressions that produces the idea of the self.

However, it is important to note that the idea and the converted impression in M₂, which is the origin of the belief in other minds, are exclusively produced from the contents of M₂, and that there is no way that the contents of M₁ can be *literally* transferred or transmitted to M₂. The reason that such literal transmission is impossible is that there is no *spatial* contiguity between M₁ and M₂. Contrastingly, in the case of the transmission of mental contents within a series of resembling impressions existing in a discrete mind, i.e., of the form of A □ Á, Hume seems to *assume* the literal transmission

of mental contents within a series of resembling perceptions. On the one hand, he seems to assume the existence of spatial contiguity in the transmission of mental contents within a discrete mind, and on the other, he denies the literal transmission of the qualitatively identical impression in M₁ to M₂ due to lack of spatial contiguity between them. The fact that the belief in other minds is exclusively produced from the contents of M₂ shows that Hume rejects the literal transmission of mental contents between two distinct minds. Hume, however, may allow temporal contiguity to facilitate the transition of the mind along the succession of resembling perceptions. As a matter of fact, Hume's unwavering conviction that some of our perceptions lack spatial location, and "*that an object may exist, and yet be no where*" (THN, 235) reinforces this claim.

There are several reasons for the literal transmission of mental contents in a series of resembling perceptions within a discrete mind. First, Hume's rejection of the literal transmission of mental contents between two distinct minds *per se* indicates that the contents of a bundle is inhibited in, or confined to, a physical setting (like a brain). Secondly, in the *Treatise*, Hume talks about 'an imaginary dissection of the brain' in which he discusses how "the animal spirits are run into all the contiguous traces, and rouze up the other ideas, that are related to it" (THN, 60). Whenever these spirits are dispatched into different regions of the brain, they first excite the proper idea "when they run precisely into the proper traces, and rummage that cell, which belongs to the idea" (THN, 61). However, since the motion of the animal spirits is "seldom direct, and naturally turns a little to the one side or the other; for this reason the animal spirits, falling into the contiguous traces, present other related ideas in lieu of that, which the mind desir'd at first to survey. This change we are not always sensible of; but continuing still the same train of thought, make use of the related idea, which is presented to us, and employ it in our reasoning, as if it were the same with what we demanded. This is the cause of many mistakes and sophisms in philosophy; as will naturally be imagin'd, and as it would be easy to show, if there was occasion" (THN, 61). In this passage, Hume makes two fundamental claims: First, he claims that the psychology of identity ascription—the mistake that we make in ascribing continuity to a train of related but interrupted ideas—is ultimately explained in terms of certain physical underpinning. When the animal spirits are dispatched to a region of brain they not only trigger the proper idea, but they also

trigger the ideas that are contiguous with the former. Secondly, and more importantly, he claims that the relation of contiguity, like the relations of causation and resemblance, plays a crucial role in the production of the fictional idea of identity. In other words, it says that spatial contiguity is a necessary condition for connecting the series of resembling impressions together, and that such necessity implies literal transmission of mental contents within a discrete mind.

To return to my discussion about the existence of the series of impressions between two discrete minds, since there is no literal transmission of the mental contents between M_1 and M_2 it implies that there is no *physical causal relation* between the contents of the two minds. The ability for Hume to draw a causal inference is bound to the experience of the constant conjunction of relevant kinds of cause and effect. And since there is no one single case in which we could perceive another's inner mental contents as preceding another's bodily movements and behaviours, as their causes. This tells us that we fundamentally lack direct access (awareness) to the mental contents of other minds. In other words, there is no physical causal link between the contents of one's mind and the contents of other minds. Of course, there is a kind of causal relation between the bodily movements and behaviours of one's mind and other minds, but this does not amount to a physical causal relation between the mental contents of two discrete minds. This suggests that the contents of the two minds cannot be causally related. For Hume, causal relation is founded on the constant conjunction of cause and effect, and since there is no constant conjunction between cause and effect, it would follow that there is no constancy between the contents of M_1 and M_2 . That is to say, constancy fails to fill the gap in the series of perceptions existing between $M_1 \square M_2$. In short, Hume identifies events as contiguous if they are beside each other in time or in space, and this means that contiguity in space is a necessary condition for any relation of causation. Since there is no spatial contiguity between M_1 and M_2 it follows that there is no causal relation between the two minds. Therefore, lack of spatial contiguity in the series involving M_1 and M_2 entails lack of constancy within the series.

At this point, I consider an objection. On the one hand, Hume holds that most of our perceptions are indivisible and, therefore, cannot be located in space; and on the other, he leaves no doubt that all perceptions, including the perceptions that are "no

where,” are causally related—the three natural relations of resemblance, contiguity and causation hold among all perceptions. Apparently, what this means is that spatial contiguity is not a necessary condition for holding between causes and their effects. If spatial contiguity is no longer a necessary condition of causal relation it follows that a perception in M_1 can cause another perception in M_2 . Basically, Hume thinks that contiguity “has little or no influence” in the successive relation among perceptions constituting a mind or thinking person (THN, 260). This conclusion may undermine my explanation of Hume’s inconsistency. To respond, for Hume, not all perceptions are divisible. The perceptions of sight and touch are said to be divisible and spatially located, and the rests are all indivisible (THN, 236). This, at least, shows that there is a ‘weaker’ spatial contiguity among our perceptions. Furthermore, although it is true that Hume believes that contiguity plays little or no role in the causal relations among perceptions, he argues elsewhere that “every thing is in common between soul and body. The organs of the one are all of them the organs of the other. The existence therefore of the one must be dependent on that of the other” (EST, 166). Consequently, “[w]here any two objects are so closely connected [body and mind], that all alterations, which we have ever seen in the one, are attended with proportionable alterations in the other; we ought to conclude, by all rules of analogy, that, when there are still greater alterations produced in the former, and it is totally dissolved, there follows a total dissolution of the latter” (EST, 165). Hume argues that not only the soul, or mind, is mortal, but also that there is a close relation between the mind and the body. Accordingly, there is a kind of spatial contiguity among our perceptions. Each mind is embedded in its own body. That is to say, the causal relations among the series of impressions are to be held within its own discrete body; the causal relations as such are constrained to its physical settings.

The reason why Hume thinks that contiguity has little or no influence at all in causal relations is that he intended to identify the contents of the mind as ‘properties,’ or ‘qualities,’ lacking physical description. In fact, on Hume’s account, there are certain qualities, or properties, which are irreducible to the physical *simpliciter*. Most of our perceptions, except perceptions of sight and touch, are strictly qualities that are irreducible to the physical. These mental properties do not occupy an independent and distinct ontology in the physical realm; they cannot be, and shall not be, categorized and

explained in physical language either. These mental states resist capture in the nomological net of physical theory. Perhaps Hume's intention for arguing the latter is that he wants to rebut the prevailing materialist philosophy of his time, most notably that of Hobbes and Spinoza. Hence, most of our perceptions are certain properties that are not connected by the relation of contiguity in the sense that materialism requires. However, perceptions are not distinct substances in the sense that Descartes' substance dualism required either. There is an intimate causal interaction between body and mind. The mind's existence is causally dependent upon the body's existence, and dissolves when the body dissolves. The soul, or mind, does not survive the body. In conclusion, for Hume, the discreteness of the mind is secured by the physical setting that it is causally dependent. This means that there is a 'weaker' kind of spatial contiguity among the perceptions. This in turn means that spatial contiguity is a necessary, though not sufficient, condition for holding between cause and effect.

We have seen that the inconsistency arises as a result of a tension in Hume's theory of belief. I have shown that constancy and coherence, which play central roles in determining our belief in objects and persons, fail to play the same role in determining our belief in other minds. Hume sets out his theory of belief in other minds through his account of sympathy. He relies on certain explanatory elements of his psychological belief-determining mechanism responsible for generating our belief in objects and persons. For instance, it is said that the force, vivacity and liveliness of impressions gives rise to a feeling and this in turn generates the belief in other minds. Moreover, Hume holds that the imagination is responsible for the other minds belief. In fact, the imagination makes possible the transmission of mental states between two distinct minds. There is a temporal succession of perceptions in the process of coming to believe in other minds. These elements are part of Hume's psychological belief-determining mechanism, which he initially used to explain the belief in objects and persons. Nevertheless, constancy and coherence as the core element of such psychological explanatory mechanisms do not play any role in his explanation of the belief in other minds. The upshot of this is that Hume's psychological belief-determining mechanism is unable to account for all categories of natural beliefs. In Book I of the *Treatise*, Hume lays out an original theory of belief, which explains the origins of our beliefs in objects and persons.

It is also intended to account for our belief in the existence of necessary connection among causes and effects. However, that theory seems to be failing when he arrives at Book II, where he indirectly sets out his theory of belief in other minds through his account of sympathy. He maintains that the belief in other minds is an essential part of his philosophical system, but the explanatory mechanism producing such belief presented in Book II is quite different from the original explanatory mechanism defended in Book I. This shows that there is a tension in Hume's general theory of belief.

This tension in Hume's theory of belief has a significant bearing on his theory of the self; it implies that constancy and coherence no longer provide necessary and sufficient condition for the specification of the criterion of personal identity. As I argued in V.2, Hume understands coherence and constancy as the necessary and sufficient conditions under which we come to believe in the continued and distinct existence of external objects and the existence of a person over time. He argues that the identities that we ascribe to external objects such as plants, animals, and artifacts depend upon the coherence and the constancy of certain sets of perceptions and the way that they affect our imagination. Moreover, Hume explicitly discusses the role of constancy and coherence in producing the identity of a person over time (THN, 253). However, he realizes later that constancy and coherence are inadequate to provide logically necessary and sufficient conditions for the specification of the criterion of identity. In fact, the failure of constancy and coherence in explaining our belief in other minds not only suggests that Hume's psychological belief-determining mechanism faces a serious problem, but also that constancy and coherence can no longer function as the necessary and sufficient conditions for the specification of our belief in the existence of objects and persons. The direct outcome of this is that it can greatly undermine Hume's psychological belief-determining mechanism, an original contribution to philosophical enterprise. I think that this tension in Hume's theory of belief may have caused him to think twice.

Conclusion

To conclude, my preferred explanation concurs with some of the above interpretations. It holds that the inconsistency emerges as a result of Hume's failure to recognize some "real connexion" among perceptions. Nevertheless, my explanation diverges from them in respect to the question of *what* caused Hume to realize that "real connexion" among

perceptions or the simplicity of the mind can save him from the inconsistency.

Most notably, Garrett has claimed that since spatial contiguity plays no role in Hume's account of mental processes, he is unable to account for the idea of personal identity in a way that distinguishes the contents of one mind from *other* minds when those contents are resembling and simultaneous. The result is that we can always conceive of two momentary but discrete bundles as having qualitatively identical perceptions that contiguously occur temporally, as in the case of two individuals watching the same movie to the exclusion of their other mental contents. As a result, Garrett suggests that Hume's problem is not about the causal explicability of the *actual* existence of discrete bundles; rather, the *possible* existence of such bundles becomes incoherent. The heart of the problem is that Hume's original theory of personal identity explained in terms of the relations of causation and resemblance is inadequate and insufficient to produce an idea of the self as strong as the actual idea that we have in real life. In other words, the actual idea of the self in real life is quite coherent, but Hume's account of the self, which clearly allows for the existence of qualitatively identical but numerically distinct perceptions existing in the minds of different individuals at the same time, presents an idea of the self that is incoherent.

My account concurs with certain themes of Garrett's interpretation, but departs from it in some important respects. There is general agreement that Hume's main concern is the unity of the bundles; that is, how to secure the contents of one mind from that of another. Furthermore, we agree that the existence of qualitatively identical but numerically distinct perceptions in different minds can cause incoherence in the bundles. In particular, Garrett has located the source of inconsistency in Hume's theory of belief in other minds. Nevertheless, he does not explain what caused Hume to recognize the inconsistency. In contrast, I have argued that the source of inconsistency lies in Hume's account of sympathy and the psychological mechanism determining our belief in other minds. I argued that Hume's account of constancy and coherence, as laid out in T.I.iv.2, establishes that the principles of constancy and coherence unite our series of successive perceptions in the mind by a process of imaginative postulation, or supplementation, of an unperceived perception, or an unsensed sensibillum. I contend that the principles of constancy and coherence may successfully account for the arising belief in the idea of the

continued and distinct existence of external objects and the idea of personal identity, but they fail to explain our belief in other minds (selves). In fact, Hume's account of *sympathy* in Book II of the *Treatise* establishes that we *actually* share the mental states of other minds around us. This shows that Hume treats the belief in other selves, like the beliefs in external objects and persons, as an important part of his psychology of identity ascriptions. Therefore, he must use the same psychological belief-determining mechanism, particularly the principles of constancy and coherence, to account for our belief in other minds. Nevertheless, coherence and constancy do not play any role whatsoever in the explanation of the arising belief in other minds. Furthermore, the absence of constancy and coherence in the explanation of the belief in other minds shows that not only the belief in other minds is not explained adequately by Hume's psychological belief-forming mechanism, but also the absence of constancy and coherence undermines even their roles as the necessary and sufficient conditions for the beliefs in objects and personal identity. This problem becomes even greater, if we note that Hume's theory of mind as defended in Book I of the *Treatise* is incompatible with his account of the self in Book II.

Notes and References

Introduction

- 1) Martin R, and Barresi J., *The Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century*, New York: Routledge 2000, p. 1.
- 2) See: Kenneth Winkler "All is Revolution in Us: Personal Identity in Shaftesbury and Hume," *Hume Studies*, Vol. 24. No. 1 (April, 2000), pp. 3-40.
- 3) See: Locke, *Essays*, 2. 27; Shaftesbury, *Characteristics*, 2: 100-6; Berkeley, *Alciphron*, 7.8; Butler, "Of Personal Identity," *Spectator* 578; Chambers, *Cyclopedia*, 'Identity'; Watts, *Philosophical Essays*, 12. 7; Perronet, *Second Vindication*, Clarke, *Letter* and three defences thereof; Collins, four works: *A Letter*, *A Reply*, *Reflections*, and *An Answer*.
- 4) Kemp Smith, Norman, "The Naturalism of David Hume (I.)," *Mind*, New Series, Vol. 14, No. 54, (Apr., 1905), p. 153.
- 5) Hougeland, John, *Artificial Intelligence: The Very Idea*, Cambridge: The MIT Press 1984, p. 41.
- 6) Fodor, Jerry, *Hume Variations*, Oxford: Clarendon Press 2003, p. 135.
- 7) Biro, John, "Hume and Cognitive Science," *History of Philosophy Quarterly*, Vol. 2, No. 3, (Jul., 1985), p. 134.
- 8) See: Daniel Dennett, "Artificial Intelligence as Philosophy and as Psychology," in *Brainstorms: Philosophical Essays on Mind and Psychology*, Cambridge: The MIT Press 1978, pp. 122-26. Dennett views Hume's role in the history of functionalism pivotal, but he thinks that Hume's functionalism suffers from the problem of undischarged homunculi; Don Garrett, "Hume's Naturalistic Theory of Representation," *Sentese*, Vol. 152, No. 3, (Oct., 2006), pp. 317-18; Tony Pitson, *Hume's Philosophy of The Self*, New York: Routledge 2002, p. 23; Simon Blackburn, *How to Read Hume*, London: Granta 2008, pp. 52-54. Blackburn draws our attention to Kant's view on the nature of the self as a logical or formal construction.
- 9) Block, Ned, "Trouble with Functionalism," in *The Nature of Mind*, ed. D. Rosenthal, Oxford: Oxford University Press 1991, p. 223.
- 10) Fodor, Jerry, *Concepts: Where Cognitive Science Went Wrong*, Oxford: Clarendon Press 1998, p. 10.
- 11) Shoemaker, Sydney, "Self and Substance," *Nous*, Vol. 31, Supplement: Philosophical Perspectives, 11, (1997), pp. 287-90.
- 12) Johnson, W. E., *Logic, Part III: The Logical Foundations of Science*, New York: Dover 1964, Chapter IX.
- 13) Olson, Eric T., "What does Functionalism Tell Us about Personal Identity?" *Nous*, Vol. 36, No. 4 (2002), p. 693.
- 14) *Ibid.*, p. 696.
- 15) *Ibid.*, p. 697.
- 16) Shoemaker, S., Swinburne, R., *Personal Identity*, Oxford: Blackwell 1984, p. 93.

Chapter I

Hume's Functionalism: A First Approximation

- 1) Hall, Roland, "Hume's Use of Locke on Identity," *Locke Newsletter*, Vol. 5 (1974), p. 56. This paper has become a classic and authoritative historical investigation of Hume's connections to Locke on the subject of personal identity.
- 2) The first philosophical formulation of the question of personal identity took place in the second

edition of Locke's *Essay* 2. 27 in response to William Molyneux's request in 1694. The following works appeared after Locke's first treatment of the subject. Shaftesbury, *Characteristics* 2: 100-6, 275-6; Clarke, *Letter* and the three subsequent defence of it; Collins, *A Letter, A Reply, Reflections, and An Answer*; Berkeley, *Alciphron* 7. 8; Butler, "Of Personal Identity," *Spectator* 578; Chambers, "Identity," *Cyclopaedia*; Watts, *Philosophical Essays* 12. 7; Perronet, *Second Vindication*. Hume wrote his section of "Of Personal Identity" almost fifty years after Locke's theory of identity in 1739. The most notable work on the subject is Reid's "Of Identity," which appeared 1785. In this period, eighteenth century, the question of personal identity was hotly debated in English Philosophy.

3) Allison, Henry, "Locke's Theory of Personal Identity: A Re-Examination," *Journal of the History of Ideas*, Vol. 27, No. 1 (Jan. -Mar., 1966), p. 56.

4) Martin R, and Barresi J., *The Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century*, p. 17.

5) *Ibid.*, p. 11.

6) Contemporary philosophy treats the problem of the criterion of identity as the problem of specifying logically necessary and sufficient conditions for an entity identified at t_1 to be the same entity at t_2 . (See: Noonan, Harold, *Personal Identity*, New York: Routledge 1989, p. 2.)

7) Philosophers conceive Locke's general theory of identity and his theory of personal identity as fundamentally concerned with offering a certain set of necessary and sufficient conditions (criterion) of identity. For instance, E. J. Lowe, argues that although the credit for introducing the notion of a criterion of identity into philosophical discourse should be given to Frege, Locke was the first modern philosopher "who grasped the key point that that in which identity 'consists' for things of one sort (say, parcels of matter) may not be the same as that in which it 'consists' for things of another sort (say, living organisms)" (See: Lowe, E. J., "What is a Criterion of Identity?" *The Philosophical Quarterly*, Vol. 39, No. 154 (Jan., 1989), pp. 1-21). Or most notably, Michael Ayers in explaining Locke's metaphysical *principium individuationis* states that "what matters with respect to questions of identity is that some 'kind' is instantiated, some general predicate satisfied, at determinate times and places. The meaning of the predicate, the idea with which it is associated, determines the criterion of identity of the particulars" (See: Ayers, Michael, *Locke: Epistemology & Ontology II*, London: Routledge 1991, p.210). Noonan also claims that "Locke's aim is clearly enough to establish a necessary and sufficient condition of identity: x is identical with y if and only if x and y are of the same kind and have the same beginning of existence" (See: Noonan, Harold, *Personal Identity*, p. 35).
8) McCann, Edwin, "Locke on Identity: Matter, Life, and Consciousness," in *The Empiricist: Critical Essays on Locke, Berkeley, and Hume*, ed. Atherton, M., Maryland: Rowman and Littlefield Publishers, Inc. 1999, p. 63.

9) This is evident, first, in Locke's correspondence with Molyneux where he explicitly characterizes the principle of individuation as a logical and metaphysical issue. See: Letter 1592, 20 January 1693, in *The Correspondence of John Locke*, ed. E. S. De Beer, Vol. IV, Oxford: Clarendon Press, 1979. Locke in response to Molyneux says: "since you mention logik and metaphysiks in relation to my book [*Essay*], whether either of those sciences may suggest to you any new heads fit to be inserted into my *Essay* in a second edition." Molyneux replied: "unless you think it may not do well to insist more particularly and at large on Aeternae Veritates and the Principium Individuationis." (Letter 1609, 2 March 1693).

10) McCann, Edwin, "Locke on Identity: Matter, Life, and Consciousness," in *The Empiricist: Critical Essays on Locke, Berkeley, and Hume*, p. 64.

11) See: King, Peter, "The Problem of Individuation in the Middle Ages," *Theoria*, Vol. 66 (2000). pp. 165-73.

- 12) Atherton, Margaret, "Corpuscles, Mechanism, and Essentialism in Locke," *Journal of the History of Philosophy*, Vol. 29 (1991), P. 33. For further reading see: (Maurice Mandelbaum, "Locke's Realism," in *Philosophy, Science, and Sense Perception*, Baltimore: Johns Hopkins Press 1964. This paper is considered to be the first work that has linked Locke's philosophy to 'mechanism,'); Edwin McCann, "Lockean Mechanism," in *Philosophy: Its History and Historiography*, ed. A. Holland, Dordrecht: D. Reidel 1985; G. A. J. Rogers, "The Intellectual Setting and Aims of the *Essay*" in *The Cambridge Companion to Locke's Essay Concerning Human Understanding*, New York: Cambridge University Press 2007, pp. 7-33.; Peter Alexander, *Ideas, Qualities, and Corpuscles: Locke and Boyle on the External World*, New York: Cambridge University Press 1985. For general background see (: Marie Boas (Hall), "The Establishment of the Mechanical Philosophy," *Osiris* Vol. 10 (1952). pp. 412-541; R. Kargon, *Atomism in England from Harriot to Newton*, Oxford: Clarendon Press 1966, esp. chaps. 1, 8 and 9; and Robert Westfall, *The Construction of Modern Science*, New York: John Wiley & Sons 1971, esp. chaps. 2 and 4).
- 13) Alexander, Peter, *Ideas, Qualities, and Corpuscles: Locke and Boyle on the External World*, New York: Cambridge University Press 1985, p. 9.
- 14) Noonan, Harold, "Locke on Personal Identity," *Philosophy*, Vol. 53, No. 205 (Jul., 1978), p. 344.
- 15) McCann, Edwin, "Locke on Identity: Matter, Life, and Consciousness," in *The Empiricist: Critical Essays on Locke, Berkeley, and Hume*, p. 65.
- 16) Jorge Gracia distinguishes between six fundamental issues in the metaphysics of individuality, viz. the intension of individuality, the extension of individuality, the ontological status of individuality, the principle of individuation, the discernibility of individuals, and the reference to individuals. These issues posit different aspects of the problem of individuation. It is said that the first issue is essentially a logical one, whereas the second, third, and fourth are metaphysical in nature. The fifth is epistemological and the sixth is concerned with semantics. The central concern in all of these philosophical issues concerning individuation involves the determination of the necessary and sufficient conditions for individuality. However, it is crucial to note that each of these philosophical issues determines, or specifies, the necessary and sufficient conditions of individuality on its own terms. For instance, the question concerning the intension of individuality refers to certain features or characteristics that are involved in determining the individuality of things. To determine the intension of individuality, for instance, is nothing other than to determine the necessary and sufficient conditions of individuality. The intension of "human being," to use an Aristotelian example, is what characterizes it as "rational animal," that is, a being with the capacity to reason. In this case, the capacity to reason is both logically necessary and a sufficient condition in its Aristotelian contexts, to determine the individuality of human beings. Roughly, it involves the traditional conceptual clarification and defining the notion of individual. Gracia identifies five features, or characteristics, that are traditionally understood as logically necessary and sufficient conditions for individuality, namely, indivisibility, distinction or difference, division, identity, and impredicability. The issue concerning the "principle of individuation" also deals with the specification of necessary and sufficient conditions of individuality. However, unlike the former which is strictly concerned with the logical conditions of individuality, the latter is mainly concerned with the metaphysical conditions to bring about individuality. Gracia describes the distinction between these two as involving "the determination of the necessary and sufficient conditions for something to be individual" as opposed to the determination of "necessary and sufficient conditions for something to become individual". For example, the necessary and sufficient conditions for something to be human are, traditionally understood, the capacity to reason and the condition of animality, whereas the necessary and sufficient conditions for something to become human, say Socrates, certainly requires more than what is

required by its logical conditions. Indeed, the capacity to reason and animality per se do not make a human Socrates. It certainly needs “principles” or “causes” to bring about Socrates distinct from the conditions that give him the features of reason and animality. Likewise, the issue concerning the discernibility of individuals, too, “seeks the necessary and sufficient conditions on the bases of which minds may know something as individual”. Nevertheless, the concern here is an epistemological one. It concerns the cause or principle that allows us to know something as individual; it involves the conditions under which something is known as individual qua individual. How does one know, for example, that Socrates is an individual and in that respect is distinguished from Plato? The discernibility of individuals as an epistemic issue highlights certain significant problems in the question of individuation. In fact, it draws our attention to the relation between a knower and the object of his knowledge, an entirely new approach to the problem of individuation. Accordingly, it suggests that any solution to the problem of individuation has to depend to a certain extent on the nature of knower and the object of his knowledge. The semantic, or the linguistic analysis, account, unlike the epistemic account that talks about the conditions according to which we know something as individual, is concerned about the conditions under which we could talk about something as individual. Again here philosophers are concerned with the specification of the necessary and sufficient conditions of individuation (See: Gracia, Jorge, *Individuality: An Essay on the Foundations of Metaphysics*, New York: State University of New York Press 1988, pp. 8-21).

17) Gracia, Jorge, *Individuality: An Essay on the Foundations of Metaphysics*, p. 180.

18) Locke himself has used the word “consists” in the *Essay*. See. *Essay*, II, xxvii, 5.

19) Martin R, and Barresi J., *The Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century*, p.16.

20) Locke adopted the self-reflexive nature of consciousness directly from Descartes, but then he wedded that element to his own memory criterion of personal identity. For Descartes, not only is the mind always active and thinking, but we are also aware of this thinking at each and every moment. The notion of self-reflexive consciousness consists of two central principles, namely, (a) we are always self-intimately aware of thinking, (b) we cannot go wrong about our awareness of thinking. In the *Second Meditation* he argues that the Demon cannot deceive him that he is sitting in front of a hot fire, and having the visual perceptions of it or having the feeling of the heat. This is called the doctrine of the infallibility of consciousness. The combination of these two principles—self-intimation and the infallibility of consciousness—results in the doctrine of the perfectly transparent mind. That is, the mind is present or transparent to itself at each moment. Descartes in response to Bourdin’s objection to the reflexive character of thought says, “the initial thought by means of which we become aware of something does not differ from the second thought by means of which we become aware that we were aware of it, anymore than this second thought differs from a third thought by which we become aware that we were aware that we were aware” (CSM II, 382). For Descartes, each thought not only encapsulates infinite reflexive mental acts but also necessarily incorporates a self-reference (See: Martin R, and Barresi J., *The Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century*, p.9). Locke adopts Descartes belief in the reflexivity of consciousness, and uses it to account for the continuity of person over time and at a time. Consciousness unifies a person both synchronically and diachronically.

21) Allison, Henry, “Locke’s Theory of Personal Identity: A Re-Examination,” *Journal of the History of Ideas*, Vol. 27, No. 1 (Jan. - Mar., 1966), p.58.

22) McCann, Edwin, “Locke on Identity: Matter, Life, and Consciousness,” in, *The Empiricist: Critical Essays on Locke, Berkeley, and Hume*, p. 76.

23) Allison, Henry, “Locke’s Theory of Personal Identity: A Re-Examination,” *Journal of the History*

of Ideas, p. 43.

24) Locke differentiated between the Real Essence and the Nominal Essence of things. The real essence of things is the internal constitution of any particular thing, which is unchangeable and unknown to us, whereas the nominal essence of a thing refers to the abstract ideas of the species, or relates to the *sort* of the things (*Essay*, 417). With respect to the organism, its real essence is its mechanical/internal constitution, as the causal basis of its powers and qualities. But it is the thing's nominal essence that determines its life or functional organization. The thing that enables us to know something, if any, about the real nature of things, is its nominal essence, that is, the abstract ideas that the perceiving mind gives rise to. The implication of this is that what determines the internal constitution or functional organization of organisms, say, an oak tree, are our abstract ideas or nominal essences of an oak tree. The life of an organism, or its functional organization, which is determined by the nominal essence of organisms, is not something over and above the constituent matter of an organism; it is, indeed, a kind of artifact that is constructed primarily out of our abstract ideas.

25) This terminology is due to Louis Loeb. See: Louis, E. Loeb, *Stability and Justification in Hume's Treatise*, Oxford: Oxford University Press 2002, pp. 139.

26) The main parameters of the scheme that I have used here is due to Don Garrett (See: Don Garrett, *Cognition and Commitment in Hume's Philosophy*, Oxford: Oxford University Press 1997, p. 167.

27) Hume, David, "Letter to a Physician," in *The Letters of David Hume*, ed. J. Y. T. Greig, Vol. 1, Oxford: Clarendon Press, 1932, p. 12.

28) A. E. Pitson argues that Hume "has a view of consciousness which would suggest that a subject stands in relation to his perceptions in something like the same way that an observer of those perceptions would do...there is a further, crucial dimension to consciousness on Hume's account. For I am conscious not merely of each of my perceptions individually as they occur; I am also conscious of myself as a *succession* of such perceptions" (Pitson, A. E., *Hume's Philosophy of the Self*, p. 34). He seems to be suggesting that Hume holds a version of reflexive theory of consciousness. But I am not convinced that Hume thinks of consciousness in terms of a self-reflexive activity of the mind as over and beyond itself.

29) Some commentators like Galen Strawson seem to be suggesting this line of thought. See: (Strawson, Galen, *The Evident Connexion: Hume on Personal Identity*, Oxford: Oxford University Press 2011, pp. 68-69.

30) Hobbes, Thomas, "De Corpore," in *The English Works* Vol. I. Trans. W. Molesworth. London: Scientia Aalen 1655/1839, 9.7.

31) Pitson, A. E., *Hume's Philosophy of the Self*, p. 23.

32) Pears, David, *Hume's System: An Examination of the First Book of His Treatise*, Oxford: Oxford University Press 1990, pp. 129-31. Eric Olson also believes that functionalism does not imply diachronic identity. I will discuss this issue in Chapter IV of this work. (See: Eric, Olson, "What does Functionalism Tell Us about Personal Identity?," *Nous*, Vol. 36, No. 4 (2002), pp. 682-98.

33) Stroud, Barry, *Hume*, London: Routledge and Kegan Paul 1977, p. 17.

34) Fodor, Jerry, *Hume Variations*, p. 27.

35) Armstrong, David, *The Mind-Body Problem: An Opinionated Introduction*, Oxford: Westview Press 1999, p. 27.

36) Flage, Daniel, "Hume's Dualism," *Nous*, Vol. 16, No. 4, (Nov., 1982), p. 539.

37) Anderson, Robert, "The Location, Extension, Shape, and Size of Hume's Perceptions" in D. W. Livingston & J. T. Kind (eds.), *Hume: a Re-evaluation*, New York: Fordham University Press 1976, p. 165.

- 38) *Ibid.*, p. 166.
- 39) *Ibid.*, p. 166.
- 40) Fodor, Jerry, *Hume Variations*, p.37. Fodor has also defended the same point elsewhere. See: Jerry Fodor, "The Revenge of the Given," in B. P. McLaughlin & J. Cohen (eds.), *Contemporary Debates in Philosophy of Mind*, UK: Basil Blackwell 2007, pp. 105–16.
- 41) Garrett, Don, *Cognition and Commitment in Hume's Philosophy*, p. 39.
- 42) See: John Bricke, *Mind and Morality: An Examination of Hume's Moral Psychology*, Oxford: Clarendon Press 1996.
- 43) Flage, Daniel, "Hume's Dualism," *Nous*, Vol. 16, No. 4, (Nov., 1982), p. 539.
- 44) Pitson, A. E., *Hume's Philosophy of the Self*, p. 61.
- 45) *Ibid.*, p. 61.
- 46) See: Phillip, Cummins, "Hume as Dualist and Anti-Dualist," *Hume Studies*, Vol. 21, No. 1 (Apr., 1995), p. 50. See also: Daniel, Flage, *David Hume's Philosophy of Mind*, London: Routledge and Kegan Paul 1990, p. 121.
- 47) Pitson, A. E., *Hume's Philosophy of the Self*, p. 62. See also: John, Yolton, *Perceptual Acquaintance: From Descartes to Reid*, Oxford: Blackwell 1984, pp. 199-200.
- 48) *Ibid.*, p. 62.
- 49) *Ibid.*, p. 62.
- 50) *Ibid.*, p. 63
- 51) *Ibid.*, p. 63.
- 52) *Ibid.*, p. 63.
- 53) Anderson, Robert, "The Location, Extension, Shape, and Size of Hume's Perceptions," p. 170.
- 54) *Ibid.*, p. 170.
- 55) See: Ned Block, "What is Functionalism," in N. Block, D. Langendean, & J. Katz (eds.), *Readings in Philosophy of Psychology Volume I*, Cambridge, Mass: Harvard University Press 1980, p. 176.
- 56) *Ibid.*, p. 175
- 57) *Ibid.*, p. 175.
- 58) *Ibid.*, p. 176.
- 59) *Ibid.*, p. 176.
- 60) The Latin translation runs as it follows: *totum in toto & totum in qualibet parte*.
- 61) Pitson, A. E., *Hume's Philosophy of the Self*, p. 64.
- 62) *Ibid.*, p. 56.
- 63) *Ibid.*, p. 59.
- 64) Block, N, "What is Functionalism," in N. Block, D. Langendean, & J. Katz (eds.), *Readings in Philosophy of Psychology Volume I*, p. 181. This is a common view among the functionalists.
- 65) Pitson, A. E., *Hume's Philosophy of the Self*, p. 65.

Chapter II

Hume's Computational Theory of Mind

- 1) Stroud, Barry, *Hume*, p. 1.
- 2) Biro, John, "Hume's New Science of the Mind," in *The Cambridge Companion to Hume*, ed. D. F. Norton, Cambridge, Lon.: Cambridge University Press 1993, p. 43.
- 3) *Ibid.*, p. 47.
- 4) *Ibid.*, p. 45.
- 5) Kemp Smith, Norman, *The Philosophy of David Hume*, London: Macmillan & Ltd 1963, pp. 459-63.

- 6) Garrett, Don, *Cognition and Commitment in Hume's Philosophy*, p. 14. Garrett claims that Hume's "distinction of representational faculties into memory and imagination... constitutes a rejection of the Cartesian ideal of a higher and radically nonimagistic representational faculty of intellect" (Garrett, p. 39). And his distinction between reason and imagination has to do with his idea that "reasoning is an operation that occurs on the representations of the imagination" (Garrett, 27), and since there are a large number of other operations or aspects of the imagination in addition to demonstrative and probable reasoning, Hume makes the distinction. And finally, Hume opposes imagination to the understanding because the understanding is an aspect of the imagination. The understanding could be conceived either as opposing memory or as opposing reason (Garrett, p. 29).
- 7) Streminger, Gerhard, "Hume's Theory of Imagination," *Hume Studies*, Vol. 6, No. 2, pp. 93-94.
- 8) *Ibid.*, p. 95.
- 9) *Ibid.*, p. 96.
- 10) Norton, F. David, "Editor's Introduction," in *A Treatise of Human Nature*, ed. D. F. Norton, M. J. Norton, Oxford: Oxford University Press 2000, p. I22.
- 11) Hume describes an experiment which allows us to see a minima visual impression. He say, "Put a spot of ink upon paper, fix your eye upon that spot, and retire to such a distance, that at last you lose sight of it; 'tis plain, that the moment before it vanish'd the image or impression was perfectly indivisible" (THN, 27). When we see the spot of ink from a close distance, the idea, or image, of it in the mind has many parts, but as we moves away from it causes fewer ideas or images in the mind. We get to a point that there is only one single image or idea, just before the spot is too far away to cause any more images. Here we have an idea, or image, with no parts; this is a minimal perception in the mind.
- 12) Streminger, Gerhard, "Hume's Theory of Imagination," p. 108.
- 13) Streminger thinks that Hume makes a distinction between natural (instinctive), rational (justified and reasonable), and irrational (unjustified and unreasonable) beliefs in the *Treatise*, but in the *Enquiry* he talks only about natural and rational beliefs. See: Gerhard, Streminger, "Hume's Theory of Imagination," p. 109.
- 14) *Ibid.*, pp. 96-99.
- 15) Hume's distinction between the "philosophical relation" and the "natural relation" plays a vital role in his philosophy. He classifies the relations of 'resemblance,' 'contiguity,' and 'causation' as natural relations, and in contrast he lists other seven relations as the philosophical relations, that are: resemblance, identity, relations in space and time, quantity, quality, contrariety, and causation. Obviously, some of these relations overlaps with the natural relations.
- 16) Norton, D. Fate, "Editor's Introduction," in *A Treatise of Human Nature*, ed. D. F. Norton, M. J. Norton, Oxford: Oxford University Press 2000, p. I25.
- 17) Stroud, Barry, *Hume*, p. 54.
- 18) Biro, John, "Hume's New Science of the Mind," in *The Cambridge Companion to Hume*, p. 52.
- 19) Stroud, Barry, *Hume*, p. 76.
- 20) *Ibid.*, p. 76.
- 21) Fodor, Jerry, *The Modularity of Mind: An Essay on Faculty Psychology*, Cambridge (Mass.): The MIT Press 1983, p. 31.
- 22) *Ibid.*, p. 31.
- 23) Biro, John, "Hume's New Science of the Mind," in *The Cambridge Companion to Hume*, p. 43n4.
- 24) Fodor, Jerry, *Hume Variations*, p. 134.
- 25) Biro, John, "Hume's New Science of the Mind," in *The Cambridge Companion to Hume*, pp. 54-55.

- 26) *Ibid.*, p. 55.
- 27) Dennett, Daniel, *Brainstorms: Philosophical Essays on Mind and Psychology*, Cambridge (Mass.): The MIT Press 1981, p. 101.
- 28) *Ibid.*, p. 101.
- 29) *Ibid.*, pp. 122-23.
- 30) Brentano, Franz, *Psychology from an Empirical Standpoint*, trans. by A.C. Rancurello, D.B. Terrell, and L. McAlister, London: Routledge 1973. Appendix.
- 31) Biro, John, "Hume and Cognitive Science," in *Historical Foundation of Cognitive Science*, ed. J-C Smith, Dordrecht: Kluwer Academic Publishers 1991, p. 133.
- 32) Biro, John, "Hume's New Science of the Mind," in *The Cambridge Companion to Hume*, p. 66.
- 33) Biro, John, "Hume and Cognitive Science," in *Historical Foundation of Cognitive Science*, p. 133.
- 34) Dennett, Daniel, *The Intentional Stance*, Cambridge (Mass.): The MIT Press 1989, p. 17.
- 35) Dennett, Daniel, *Brainstorms: Philosophical Essays on Mind and Psychology*, p. 30.
- 36) Reichenbach, Hans, *Experience and Prediction*, Chicago: University of Chicago Press 1938, pp. 211-12.
- 37) Stich, Stephen, "Dennett on Intentional System," in *Collected Papers: Mind and Language*, Vol. 1, Oxford: Oxford University Press 2011, p. 74.
- 38) Fodor, J, and Lepore, E, *Holism: A Shopper's guide*, Cambridge (Mass.): Blackwell 1992, p. 154.
- 39) See: Jerry Fodor, "Fodor's Guide to Mental Representation," in *A Theory of Content: And Other Essays*, Cambridge (Mass.): The MIT Press 1990, pp. 3-29.
- 40) See: Christopher Viger, "Where do Dennett's Stances Stand? Explaining our Kind of Mind," in *Dennett's Philosophy: A Comprehensive Assessment*, eds. Ross, D, Brook, A, Thompson, D, Cambridge (Mass.): The MIT Press 2000, pp.131-47.
- 41) Dennett, Daniel, *Brainstorms: Philosophical Essays on Mind and Psychology*, pp. 80-81.
- 42) Biro, John, "Hume and Cognitive Science," in *Historical Foundation of Cognitive Science*, p. 131.
- 43) Other people like J. Haugeland have also concluded that Hume fails to articulate a plausible functionalistic mental mechanism. See: J. Haugeland, *Artificial Intelligence: The Very Idea*, Cambridge (Mass.): The MIT Press 1985, p. 44.
- 44) See: D. Dennett, "'Competition in the Brain,'" in http://www.edge.org/q2008/q08_11.html.
- 45) See: Ned Block, "Troubles with Functionalism," pp. 211-29.
- 46) Bower, Kenneth, "Imagery: From Hume to Cognitive Science," *Canadian Journal of Philosophy*, Vol. 14, (1984), p. 223.
- 47) Biro, John, "Hume and Cognitive Science," in *Historical Foundation of Cognitive Science*, p.129.
- 48) *Ibid.*, p. 120.
- 49) Stroud, Barry, *Hume*, p. 15
- 50) Fodor, Jerry, *Hume Variations*, p. 27.
- 51) *Ibid.*, p. 27.
- 52) *Ibid.*, p. 27.
- 53) *Ibid.*, pp. 16-17.
- 54) *Ibid.*, p. 8.
- 55) *Ibid.*, p. 134.
- 56) *Ibid.*, p. 15.
- 57) Fodor enumerates five main theses comprising RTM. See: Jerry Fodor, *Concepts: Where Cognitive Science Went Wrong*, pp. 6-25.
- 58) Fodor, Jerry, *Concepts: Where Cognitive Science Went Wrong*, p. 8.
- 59) *Ibid.*, P. 7.

- 60) Fodor, J. and Pylyshyn, Z., "Connectionism and Cognitive Architecture: A Critical Analysis," *Cognition*, 28(1988), p. 8.
- 61) Fodor, Jerry, *Concepts: Where Cognitive Science Went Wrong*, p. 8.
- 62) *Ibid.*, p.9.
- 63) Fodor, Jerry, *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, Cambridge (Mass.): The MIT Press 1987, p. 17.
- 64) Fodor, J. and Pylyshyn, Z., "Connectionism and Cognitive Architecture: A Critical Analysis," p. 19.
- 65) Fodor, Jerry, *Concepts: Where Cognitive Science Went Wrong*, p. 10.
- 66) *Ibid.*, p. 12.
- 67) Fodor's theory targets the Inferential Role Semantics theory of meaning as well. See Jerry Fodor, *Concepts: Where Cognitive Science Went Wrong*. pp. 12-15.
- 68) Fodor, Jerry, *Hume Variations*, p. 42.
- 69) *Ibid.*, p.43.
- 70) Garrett, Don, "Hume's Naturalistic Theory of Representation," *Synthese*, Vol. 152, No. 3, (Oct., 2006), p. 305.
- 71) Baier, Annette, *A Progress of Sentiments: Reflections on Hume's Treatise*, Cambridge (Mass.): Harvard University Press 1991, p. 160.
- 72) Garrett, Don, "Hume's Naturalistic Theory of Representation," p. 307.
- 73) *Ibid.*, p. 305.
- 74) *Ibid.*, p. 305
- 75) Owen, David and Cohen, Rachel, "Hume on Representation, Reason and Motivation," *Manuscript*, Vol. 20, 1997, pp. 55-56.
- 76) *Ibid.*, p. 56.
- 77) Garrett, Don, "Hume's Naturalistic Theory of Representation," p. 209.
- 78) Raynor, David, "Hume and Berkeley's Three *Dialogues*," in *Studies in the Philosophy of the Scottish Enlightenment*, ed. M. A. Stewart, Oxford: Clarendon Press 1990, p. 236.
- 79) Fodor, Jerry, "The Mind-Body Problem," *Scientific American* Vol. 244, No. 1, (Jan., 1981), p. 122.
- 80) See: Simmons, Alison. "Are Cartesian Sensations Representational?" *Nous*, Vol. 33, No.3 (Sep., 1999), pp. 347-69.
- 81) Garrett, Don, "Hume's Naturalistic Theory of Representation," p. 314. The reason is that intrinsic representation would give rise to relations of representation that could not be understood in terms of any of those seven philosophical relations.
- 82) *Ibid.*, p. 314.
- 83) *Ibid.*, p. 315.
- 84) *Ibid.*, pp. 315-16.
- 85) *Ibid.*, p. 316.
- 86) *Ibid.*, p. 316.
- 87) Garrett, Don, "Hume's Theory of Ideas," in *The Blackwell Companion to Hume*, ed. E. Radcliffe, USA: Blackwell Publishing Ltd 2008, p. 48-49.
- 88) Garrett, Don, "Hume's Naturalistic Theory of Representation," P. 316-17.
- 89) Garrett, Don, "Hume's Theory of Ideas," p. 323.
- 90) *Ibid.*, p. 319.
- 91) Fodor, Jerry, *Hume Variations*, p. 55.
- 92) *Ibid.*, p. 42.

93) *Ibid.*, p. 42.

94) *Ibid.*, p. 40.

95) See: Julesz, Bela, *Foundation of Cyclopean Perception*, Chicago: The University of Chicago Press 1971.

96) Fodor, Jerry, *Hume Variations*, p. 49.

Chapter III

Hume's Psychology of Identity Ascriptions

1) Louis Loeb treats the case of the belief in 'material substrata' that Hume talks about in the section of "Of the Ancient Philosophy" as different from the case of the belief in the external objects that Hume discusses in the section of "Of Scepticism with Regard to the Senses." See: Loeb, Louis, E., *Stability and Justification in Hume's Treatise*, pp. 187-93. John Yolton, Daniel Flage and Harold Noonan have argued along the same line. See: Yolton John *Perceptual Acquaintance from Descartes to Reid*, pp.173-80; Flage, Daniel, *David Hume's Theory of Mind*, pp. 61-65; Noonan, Harold, *Hume on Knowledge*, London: Routledge 1999, pp. 201-203.

2) Waldow, Anik, "Identity of Persons and Objects: Why Hume Considered Both As the Two Sides of the Same Coin," *Journal of Scottish Philosophy*, Vol. 8, No. 2, (Sep., 2010), p. 164.

3) Price, H. H., *Hume's Theory of the External World*, Oxford: The Clarendon Press 1940, P. 19.

4) H. H. Price has said that Hume's distinction between the two principles of Constancy and Coherence is unnecessary because he could have easily reduced the two principles to one. See: Price, H. H., *Hume's Theory of the External World*, P. 37. Louis E. Loeb has also taken the same view. See: Loeb, Louis, E., *Stability and Justification in Hume's Treatise*, pp. 187-93.

5) See: Bennett, Jonathan, *Locke, Berkeley, Hume: Central Themes*, Oxford: The Clarendon Press 1971, pp. 329-30; See also: Collier, Mark, "Filling the Gaps: Hume and Connectionism on the Continued Existence of Unperceived Objects," *Hume Studies* Vol. 25 (1999), p. 159; Laird, John, *Hume's Philosophy of Human Nature*, London: Archon Books 1967, pp. 149-150; and Wright, John, P., *The Sceptical Realism of David Hume*, Minneapolis: University of Minnesota Press 1983, p. 64. For example, Bennett claims (p. 329) that Hume registers his dissatisfaction with the coherence-based explanation by means of a "blatantly unargued assertion." Collier endorses Bennett's reading (p. 159). Write has taken the same stance (p. 64).

6) I have borrowed Loeb's terminology here. See: Loeb, Louis, E., *Stability and Justification in Hume's Treatise*, pp. 187-88.

7) Black, Tim, "The Distinction between Coherence and Constancy in Hume's *Treatise* I. Iv. 2," *The British Journal for the History of Philosophy*, Vol. 15, No. 1 (Feb., 2007), pp.1-25.

8) I have adopted Stroud's outline. See: Stroud, Barry, *Hume*, p. 101.

9) *Ibid.*, p. 101.

10) Price, H. H., *Hume's Theory of the External World*, pp. 33-4.

11) See: H. H. Price, *Hume's Theory of the External World*, p. 50; T. Black, "The Distinction between Coherence and Constancy in Hume's *Treatise* I. Iv. 2," *The British Journal for the History of Philosophy*, p.7.

12) Price, H. H., *Hume's Theory of the External World*, p. 60.

- 13) *Ibid.*, p. 60.
- 14) *Ibid.*, p. 61.
- 15) *Ibid.*, p. 60.
- 16) Black, Tim, "The Distinction between Coherence and Constancy in Hume's *Treatise* I. Iv. 2," p. 7.
- 17) Price, H. H, *Hume's Theory of the External World*, p. 20.
- 18) Among the advocates of this interpretation are, most notably, Jonathan Bennett (1971, p. 323); Stroud (1977, p. 100); Oliver Johnson, *The Mind of David Hume*, University of Illinois Press 1995, pp. 250-52; Louis Loeb (2002, pp. 178-79.
- 19) Black, Tim, "The Distinction between Coherence and Constancy in Hume's *Treatise* I. Iv. 2," p. 8.
- 20) Schnall, Ira, "Constancy, Coherence, and Causality," *Hume Studies*, Vol. 30, No. 1, April (2004). 38; and Gomberg, Paul, "Coherence and Causal Inference in Hume's *Treatise*," *Canadian Journal of Philosophy*, Vol. VI, No. 4, Dec (1976), p. 693.
- 21) See: Louis Loeb, 2002, p. 188; H. H Price, 1940, pp. 37-38; Jonathan, Bennett, 1971, p. 329-30.
- 22) Stroud, Barry, *Hume*, p. 100.
- 23) Price, H. H, *Hume's Theory of the External World*, p. 39.
- 24) The example is due to Loeb. See: Louis Loeb, *Stability and Justification in Hume's Treatise*, p. 178.
- 25) Loeb, Louis, *Stability and Justification in Hume's Treatise*, p. 179.
- 26) *Ibid.*, p. 179.
- 27) Price, H. H, *Hume's Theory of the External World*, p. 67.
- 28) *Ibid.*, p. 68.
- 29) *Ibid.*, p. 60.
- 30) Loeb, Louis, *Stability and Justification in Hume's Treatise*, p. 148.
- 31) *Ibid.*, p. 148.
- 32) See: Louis Loeb, *Stability and Justification in Hume's Treatise*, pp. 152-54.
- 33) *Ibid.*, p. 176.
- 34) *Ibid.*, p. 173.
- 35) Price, H. H, *Hume's Theory of the External World*, p. 50.
- 36) *Ibid.*, p. 50.
- 37) *Ibid.*, pp. 50-51.
- 38) *Ibid.*, pp. 52-53.
- 39) *Ibid.*, p. 53.
- 40) *Ibid.*, p. 53.
- 41) Loeb, Louis, *Stability and Justification in Hume's Treatise*, p. 181.
- 42) Bennett, Jonathan, *Locke, Berkeley, Hume: Central Themes*, p. 324.
- 43) Loeb, Louis, *Stability and Justification in Hume's Treatise*, p. 182.
- 44) *Ibid.*, p. 182.
- 45) *Ibid.*, p. 183.
- 46) *Ibid.*, p. 183.
- 47) *Ibid.*, p. 183.
- 48) This terminology is due to Price. See: H. H. Price, *Hume's Theory of the External World*, p. 55.
- 49) Loeb, Louis, *Stability and Justification in Hume's Treatise*, p. 182.

- 50) *Ibid.*, p. 184.
- 51) Price, H. H, *Hume's Theory of the External World*, p. 57.
- 52) *Ibid.*, p. 58.
- 53) *Ibid.*, p. 82.

Chapter IV

Hume's Functionalism and Personal Identity

- 1) Pitson, A. E. *Hume's Philosophy of the Self*, p. 91.
- 2) Clarke and Butler, for example, had criticized Locke's consciousness account of the self on the ground that without a substantial self that remains identical over time a person would not be responsible for her past actions, nor would she have any reason to be concerned about her future actions and their consequences. The same criticism could be leveled against Hume as well. Collins responded that continuity of consciousness alone is enough to make us responsible for our past actions, but he also said that 'Sympathy and Concern' (in Samuel Clarke, *The Works of Samuel Clarke*, New York, Garland Publishing, 1978) can bound distinct episodes of one's consciousness to each other. This kind of response finds full treatment in Hume's account of the self as it regards passions.
- 3) McIntyre, Jane, "Hume and the Problem of Personal Identity," in *The Cambridge Companion to Hume*, Second Edition, eds. Norton D. F. And Norton, M. J, Cambridge: Cambridge University Press 2009, p. 191.
- 4) See: Pitson, A. E., *Hume's Philosophy of the Self*, pp. 86-88. This categorization is due to him.
- 5) *Ibid.*, p. 91.
- 6) McIntyre, Jane, "Hume and the Problem of Personal Identity," p. 192.
- 7) I have used Norton schematic explanation here. See: David, F. Norton, "Editor's Introduction," in *A Treatise of Human Nature*, p. I52.
- 8) See: David, F. Norton, "Editor's Annotations," in *A Treatise of Human Nature*, p. 493.
- 9) Pitson, A. E., *Hume's Philosophy of the Self*, p. 127.
- 10) McIntyre, Jane, "Hume and the Problem of Personal Identity," p. 193.
- 11) Pitson, A. E., *Hume's Philosophy of the Self*, p. 92.
- 12) *Ibid.*, p. 128.
- 13) Shoemaker, S., Swinburne, R, *Personal Identity*, Oxford: Blackwell 1984, p. 93.
- 14) *Ibid.*, p. 93.
- 15) Fuller, Gary, "Functionalism and Personal Identity," *Studies in Personalist Philosophy*, Vol. 8, Spring 1992, p. 135.
- 16) Shoemaker, S., Swinburne, R, *Personal Identity*, p. 89.
- 17) *Ibid.*, p. 110.
- 18) Fuller, Gary, "Functionalism and Personal Identity," p. 135.
- 19) Shoemaker, S., Swinburne, R, *Personal Identity*, p. 95.
- 20) *Ibid.*, p. 96.
- 21) *Ibid.*, p. 95.
- 22) *Ibid.*, p. 95.
- 23) Quine, Willard Van Orman, *Word and Object*, Cambridge: The MIT Press 1960, p. 193.
- 24) Shoemaker, Sydney, "Self and Substance," *Nous*, Vol. 31, Supplement: Philosophical Perspectives, 11, (1997), p. 287.
- 25) *Ibid.*, p. 284.

- 26) *Ibid.*, P. 301.
- 27) *Ibid.*, p. 285.
- 28) *Ibid.*, p. 285.
- 29) *Ibid.*, p. 291.
- 30) *Ibid.*, p. 288.
- 31) See: W. E Johnson, *Logic, Part III: The Logical Foundation of Science*, New York: Dover 1964.
- 32) Shoemaker, Sydney, "Self and Substance," p. 288.
- 33) *Ibid.*, p. 289.
- 34) *Ibid.*, p. 292.
- 35) *Ibid.*, pp. 292-93.
- 36) *Ibid.*, p. 292.
- 37) *Ibid.*, p. 295.
- 38) *Ibid.*, p. 296.
- 39) Olson, Eric, "What Does Functionalism Tell Us about Personal Identity?" p. 686.
- 40) To understand the distinction here, comparing it with Locke's distinction between personhood and humanhood may help.
- 41) Olson, Eric, "What Does Functionalism Tell Us about Personal Identity?" p. 687.
- 42) Olson, Eric, *What Are WE? A Study in Personal Ontology*, Oxford: Oxford University Press 2007, p. 30.
- 43) *Ibid.*, p. 34.
- 44) See: Shoemaker, S., Swinburne, R, *Personal Identity*, pp. 92-97.
- 45) Olson, Eric, "What Does Functionalism Tell Us about Personal Identity?" p. 683.
- 46) *Ibid.*, p. 689.
- 47) *Ibid.*, p. 689.
- 48) *Ibid.*, p. 689.
- 49) Shoemaker, S., Swinburne, R, *Personal Identity*, pp. 108-11. In the BST procedure the mental contents or informational state of brain A is recorded and then transferred to brain B that it is mental contents are already erased.
- 50) Olson, Eric, "What Does Functionalism Tell Us about Personal Identity?" p. 690.
- 51) *Ibid.*, p. 690.
- 52) *Ibid.*, p. 690.
- 53) *Ibid.*, p. 693.
- 54) *Ibid.*, p. 691.
- 55) *Ibid.*, p. 692.
- 56) *Ibid.*, p. 693.
- 57) *Ibid.*, p. 693.
- 58) Shoemaker, S., Swinburne, R, *Personal Identity*, p. 93.
- 59) Shoemaker, Sydney, "Self, Body, and Coincidence," *Proceeding of the Aristotelian Society*, Vol. 73, (1999), p. 299.
- 60) Olson, Eric, "What Does Functionalism Tell Us about Personal Identity?" p. 696.
- 61) Parfit, Derek, *Reasons and Persons*, Oxford: Oxford University Press 1984, p. 225.
- 62) *Ibid.*, p. 211.
- 63) *Ibid.*, p. 225.
- 64) *Ibid.*, p. 225.

- 65) Shoemaker, Sydney, "Critical Notices. Reasons and Persons. Derek Parfit," *Mind*, Vol. 94 (1985), p.446.
- 66) *Ibid.*, p. 446. See also: Owen, D. and Cohen, R., "Hume on Representation, Reason and Motivation," *Manuscript*, Vol. 20, 1997, pp. 55-56.
- 67) Noonan, Harold, *Personal Identity*, p. 83.
- 68) See: McIntyre, Jane, "Is Hume's Self consistent?" in *McGill Hume Studies*, ed. D. F. Norton, N. Capaldi, W. Robison, San Diego: Austin Hill Press 1979, pp. 79-88.
- 69) Chisholm, Roderick M., *Person and Object*, London: Allen & Unwin, 1976, p. 83.
- 70) Noonan, Harold, *Personal Identity*, p. 83.
- 71) *Ibid.*, p. 84.
- 72) *Ibid.*, p. 84.
- 73) *Ibid.*, p. 85.
- 74) *Ibid.*, p. 85.
- 75) *Ibid.*, p. 86.
- 76) *Ibid.*, p. 88.
- 77) McIntyre, Jane, "Hume and the Problem of Personal Identity," P. 203.
- 78) Clarke, Samuel, *The Works of Samuel Clarke*, 3 vols. New York: Garland Press 1978, p. 787.
- 79) Martin R, and Barresi J., *The Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century*, p. 41.
- 80) Olson, Eric, "What Does Functionalism Tell Us about Personal Identity?" p. 693.

Chapter V

Hume's Second Thoughts about Personal Identity

- 1) For a collection of different explanations of Hume's inconsistency. See: Garrett, Don, *Cognition and Commitment in Hume's Philosophy*, pp. 167-75.
- 2) Patten, S. C., "Hume's Bundles, Self-Consciousness, and Kant," *Hume Studies*, Vol. 2, No. 1 (Apr., 1976), p. 62.
- 3) *Ibid.*, p. 64.
- 4) Stroud, Barry, *Hume*, p. 132.
- 5) *Ibid.*, p. 140.
- 6) Garrett, Don, *Cognition and Commitment in Hume's philosophy*, p. 172.
- 7) *Ibid.*, p. 172.
- 8) Stroud, Barry, *Hume*, p. 126.
- 9) Garrett, Don, "Hume's Naturalistic Theory of Representation," p. 312.
- 10) Pears, David, "Hume on Personal Identity," *Hume Studies*, Vol. 19, No. 2 (Nov., 1993), p. 290.
- 11) *Ibid.*, p. 293.
- 12) Fogelin, Robert, *Hume's Skepticism in the Treatise of Human Nature*, London: Routledge and Kegan Paul 1985, p. 107.
- 13) Garrett, Don, *Cognition and Commitment in Hume's Philosophy*, p. 179.
- 14) *Ibid.*, pp. 180-81.
- 15) *Ibid.*, pp. 181-82.
- 16) *Ibid.*, p. 182.
- 17) *Ibid.*, p. 186.
- 18) *Ibid.*, p. 186.
- 19) *Ibid.*, p. 185.

- 20) *Ibid.*, p. 184.
- 21) Fogelin, Robert, *Hume's Skepticism in the Treatise of Human Nature*, pp.104-105.
- 22) Garrett, Don, *Cognition and Commitment in Hume's Philosophy*, p. 184.
- 23) *Ibid.*, p. 185.
- 24) Pitson, A. E. *Hume's Philosophy of the Self*, p. 158.
- 25) Noonan, Harold, *Personal Identity*, p. 2.
- 26) *Ibid.*, pp.77-88.
- 27) Wright, John P., *The Sceptical Realism of David Hume*, Minneapolis: University of Minnesota Press 1983, p. 64.
- 28) Stroud, Barry, *Hume*, p. 96.
- 29) Loeb, Louis, *Stability and Justification in Hume's Treatise*, p. 13.
- 30) *Ibid.*, p. 13.
- 31) Goldman, Alvin I., "What is Justified Belief," in *Naturalizing Epistemology*, ed. Hilary Kornblith, Cambridge: The MIT Press 1987, p. 91.
- 32) Kemp Smith has also talked about similar nonepistemic psychological property for the specification of the conditions of these ideas, but he views *irresistibility* or *inevitability* of 'natural beliefs' as both necessary and sufficient conditions for the justification of our beliefs. Natural beliefs, so understood, have "Nature's sanctions"; "we have no choice but to accept them; they impose themselves upon the mind." Thus, they are "removed beyond the reach of our sceptical doubts." See: Kemp Smith, Norman, *The Philosophy of David Hume*, pp. 87, 455, 486. The important point in Kemp Smith's interpretation is that Hume's justificatory status of belief depends on the nonepistemic character of certain facts about beliefs or the psychological mechanism that produces such beliefs. Since our natural beliefs in causation, body, and the self are simply irresistible or inevitable, it should follow that the irresistibility and inevitability itself and the underlying psychological mechanism generating it would specify necessary and sufficient conditions for justified beliefs. However, Loeb argues that Kemp Smith's interpretation fails to do so because it "has no explanation of the normative force of epistemic judgments to attribute to Hume." See: Loeb, Louis, *Stability and Justification in Hume's Treatise*, p. 22.
- 33) Loeb, Louis, *Stability and Justification in Hume's Treatise*, p. 65.
- 34) *Ibid.*, p. 68.
- 35) Pitson, A. E., *Hume's Philosophy of the Self*, pp. 142-62; Anik Waldow has written a whole book on the subject. See: Waldow, Anik, *David Hume and the Problem of Other Minds*, New York: Continuum 2009; P. F. Strawson has said that Hume never explicitly discusses the question of other selves. See: Strawson, P. F., *Scepticism and Naturalism*, London: Methuen 1987, p.11.
- 36) Pitson, A. E., *Hume's Philosophy of the Self*, p. 143.
- 37) *Ibid.*, p. 153.
- 38) *Ibid.*, p. 153.
- 39) *Ibid.*, p. 153.

Bibliography

- Alexander, Peter, *Ideas, Qualities, and Corpuscles: Locke and Boyle on the External World*, New York: Cambridge University Press 1985.
- Allison, Henry E., "Locke's Theory of Personal Identity: A Re-examination," *Journal of History of Ideas*, Vol. 27, (1966), pp. 41-58.
- Anderson, Robert, "The Location, Extension, Shape, and Size of Hume's Perceptions" in *Hume: a Re-evaluation*, ed. D. W. Livingston & J. T. Kind, New York: Fordham University Press 1976, pp. 153-71.
- Ashley, L. and Stack, M., "Hume's Theory of the Self and its Identity," *Dialogues*, Vol. 13, (1974), pp. 239-54.
- Atherton, Margaret, "Corpuscles, Mechanism, and Essentialism in Locke," *Journal of the History of Philosophy*, Vol. 29 (1991), pp. 47-67.
- . "Locke on Essences and Classification," in *The Cambridge Companion to Locke's Essay Concerning Human Understanding*, ed. L. Newman, Cambridge: Cambridge University Press 2007, pp. 258-85.
- Ayers, Michael, *Locke: Epistemology & Ontology II*, London: Routledge 1991.
- . "Individuals without Sortals," *Canadian Journal of Philosophy*, Vol. 4 (1974), pp. 125-27.
- . "Mechanism, Supperaddition and the Proof of God's Existence in Locke's *Essay*," *Philosophical Review*, Vol. 40, No. 2, pp. 210-51.
- . "Locke versus Aristotle on Natural Kinds," *The Journal of Philosophy*, Vol. 78, No. 5 (May, 1981), pp. 247-72.
- Armstrong, David, *The Mind-Body Problem: An Opinionated Introduction*, Oxford: Westview Press 1999.
- Baier, Annette, *A Progress of Sentiments: Reflections on Hume's Treatise*, London: Harvard University Press 1991.
- Barber, F. Kenneth, Gracia, J. J. E., *Individuation and Identity in Early Modern Philosophy: Descartes to Kant*, New York: State University of New York Press 1994.
- Basson, A. H., *David Hume*, London: Pelican 1958.
- Baxter, Donald, "A Defence of Hume on Identity Through Time," *Hume Studies*, Vol. 13, No. 2, (1987), pp. 323-42.
- . *Hume's Difficulties: Time and Identity in the TREATISE*, New York: Routledge 2008.
- Beauchamp, Tom L., "Self Inconsistency or Mere Self-Perplexity," *Hume Studies*, Vol. 5, No. 1 (Apr., 1979), pp. 37-44.
- Bennett, Jonathan, *Locke, Berkeley, Hume: Central Themes*, Oxford: The Clarendon Press 1971.
- Berkeley, George, "Alciphron: or, The Minute Philosopher. In Seven Dialogues. Containing an Apology for the Christian Religion, against those whoe are called Free-thinkers," in *The Works of George Berkeley, Bishop of Cloyne*, 9 vols. ed. A. A. Luce and T. E. Jessop, Edinburge: Thomas Nelson 1948-57.
- Biro, John, "Hume's New Science of the Mind," in *The Cambridge Companion to Hume*, ed. D. F. Norton, Cambridge: Cambridge University Press 1993, pp.33-63.

- . “Hume and Cognitive Science,” *History of Philosophy Quarterly*, Vol. 2 (1985), pp. 257-74.
- . “Hume on Memory and Self-identity”, *Review of Metaphysics*, Vol. 30 (1976), pp.19-38.
- . “Hume’s Difficulties with the Self,” *Hume Studies*, Vol. 5 (1979), pp. 45-54.
- Black, Tim, “The Distinction between Coherence and Constancy in Hume’s *Treatise* I. Iv. 2,” *The British Journal for the History of Philosophy*, Vol. 15, No. 1 (Feb., 2007), pp.1-25.
- Blackburn, Simon, “Personal Identity and the Self,” in *How to Read Hume*, London: Granta Books 2008, pp. 45-54.
- Block, Ned, “What is Functionalism,” in *Readings in Philosophy of Psychology Volume I*, ed. N. Block, D. Langendean, & J. Katz, Cambridge, Mass: Harvard University Press 1980, pp. 171-85.
- . “Troubles with Functionalism” in *The Nature of Mind*, ed. D. Rosenthal, Oxford: Oxford University Press 1991, pp. 211-29.
- Boas, Marie, “The Establishment of the Mechanical Philosophy,” *Osiris* Vol. 10 (1952). pp. 412-541.
- Bower, Kenneth, “Imagery: From Hume to Cognitive Science,” *Canadian Journal of Philosophy*, Vol. 14, 1984, pp. 217-34.
- Brentano, Franz, *Psychology from an Empirical Standpoint*, trans. by A.C. Rancurello, D.B. Terrell, and L. McAlister, London: Routledge 1973.
- Bricke, John, *Hume’s Philosophy of Mind*, Edinburgh: Edinburgh University Press 1980.
- . *Mind and Morality: An Examination of Hume’s Moral Psychology*, Oxford: Clarendon Press 1996.
- Brooks, D. H. M., “Strawson, Hume, and the Unity of Consciousness,” *Mind*, Vol. 94, (Oct., 1985), pp. 583-86.
- Butler, Joseph, “The Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature,” in *The Works of Joseph Butler*, 2 vols. ed. D. F. Norton, Oxford: Clarendon Press 1896.
- Capaldi, Nicholas, *David Hume: The Newtonian Philosopher*, Boston: Twayne’s World Leader Series 1975.
- Chambers, Ephraim, *Cyclopaedia; or, An Universal Dictionary of Arts and Sciences; containing an Explication of the Terms, and an Account of the Things Signified thereby, in the several arts, both Liberal and Mechanical; and the several Sciences, Human and Divine*, London, 1742.
- Clack, Jerold, “Chisholm and Hume on Observing the Self,” *Philosophy & Phenomenological Research*, Vol. 33, No.3, (Mar., 1973), pp. 338-48.
- Clarke, Samuel, “A Letter to Mr. Dodwell; wherein All the Arguments in his Epistolary Discourse against the Immortality of the Soul are particularly answered,” in *The Works of Samuel Clarke*, 4 vols. New York: Garland Press 1978.
- Collier, Mark, “Hume and Cognitive Science: The Current Status of the Controversy Over Abstract Ideas,” *Phenomenology and Cognitive Science*, Vol. 4, No. 2, (Jun., 2005), pp. 197-207.
- . “Filling the Gaps: Hume and Connectionism on the Continued Existence of Unperceived Objects,” *Hume Studies*, Vol. 25, 1999, pp. 155-70.

- Collins, Anthony, "An Answer to Mr. Clarke's Third Defense of his Letter to Mr. Dodwell," in *The Works of Samuel Clarke*, Vol. 3. New York: Garland Press 1978.
- Copi, Irving, "Essence and Accident," *The Journal of Philosophy*, Vol. 51, No. 23, American Philosophical Association (Nov., 1954), pp. 706-19.
- Cummins, Phillip, "Hume as Dualist and Anti-Dualist," *Hume Studies*, Vol. 21, No. 1, (Apr., 1995), pp. 47-55.
- Davidson, Donald, "Hume's Cognitive Theory of Pride," *Journal of Philosophy*, Vol. 73, No. 19, (1976), pp. 744-57.
- Davis, Lawrence, "Functionalism and Personal Identity," *Philosophy and Phenomenological Research*, Vol. 58, No. 4, (Dec., 1998), pp. 781-804.
- Dennett, Daniel, "A Cure for the Common Code," in *Brainstorms: Philosophical Essays on Mind and Psychology* Cambridge Mass.: MIT Press 1981, pp. 90-109.
- . "Artificial Intelligence as Philosophy and as Psychology," in *Brainstorms: Philosophical Essays on Mind and Psychology* pp. 109-29.
- . "The Reality of Selves," in *Consciousness Explained*, D. Dennett, Boston: Little, Brown and Company 1990, pp. 412-30.
- . *The Intentional Stance*, Cambridge (Mass.): The MIT Press 1989.
- Doddridge, Philip, *A Course of Lectures on the Principle Subjects in Pneumatology, Ethics, and Divinity: with References to the most Considerable Authors on each Subject*, ed. S. Clark, London: J. Buckland et al 1763.
- Flage, Daniel, *David Hume's Theory of Mind*, London: Routledge 1990.
- . "Hume's Dualism," *Nous*, Vol. 16, No. 4, (Nov., 1982), pp. 527-41.
- Fodor, Jerry, *Hume Variations*, Oxford: Oxford University Press 1997.
- . *Concepts: Where Cognitive Science Went Wrong*, Oxford: Clarendon Press 1998.
- . *The Language of Thought*, New York: Crowell 1975.
- . *The Modularity of Mind*, Cambridge, Mass.: MIT Press 1983.
- . *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, Cambridge (Mass.): The MIT Press 1987.
- . *Representations*, Cambridge, Mass.: Bradford Books/MIT Press 1981.
- . *A Theory of Content and Other Essays*, Cambridge, Mass.: Bradford Books/MIT Press 1990.
- . "The Mind-Body Problem," *Scientific American* Vol. 244, No. 1, (Jan., 1981), pp. 114-25.
- Fodor, J. and Lepore, E., *Holism: A Shopper's guide*, Cambridge (Mass.): Blackwell 1992.
- Fodor, J. and Pylyshyn, Z., "Connectionism and Cognitive Architecture: A Critical Analysis," *Cognition*, 28(1988), pp. 3-71.
- Fogelin, Robert, *Hume's Skepticism in the Treatise of Human Nature*, London: Routledge and Kegan Paul 1985.
- . "Hume and the Missing Shade of Blue," *Philosophy and Phenomenological Research*, Vol. 45, No. 2, (Dec., 1984), pp. 263-71.
- . *Hume's Skeptical Crisis*, Oxford: Oxford University Press 2009.
- Fuller, Gary, "Functionalism and Personal Identity," *Studies in Personalist Philosophy*, Vol. 8, Spring 1992, pp. 133-43.
- Garrett, Don, "Hume's Self-Doubts about Personal Identity," *The Philosophical Review*, Vol. 90, No. 3, (Jul., 1981), pp. 337-58.

- . *Cognition and Commitment in Hume's Philosophy*, Oxford: Oxford University Press 1997.
- . "Hume's Naturalistic Theory of Representation," *Synthese*, Vol. 152, No. 3, (Oct., 2006), pp. 301-19.
- . "Hume's Theory of Ideas," in *The Blackwell Companion to Hume*, ed. E. Radcliffe, USA: Blackwell Publishing Ltd 2008, pp. 41-57.
- Gaskin, J. C. A. *Hume's Philosophy of Religion*, London: Macmillan Press 1978.
- Goldman, Alvin I., "What is Justified Belief," in *Naturalizing Epistemology*, ed. Hilary Kornblith, Cambridge: The MIT Press 1987, pp. 91-113.
- Gomberg, Paul, "Coherence and Causal Inference in Hume's *Treatise*," *Canadian Journal of Philosophy*, Vol. VI, No. 4, (Dec., 1976), p. 693-704.
- Gracia, Jorge, "Introduction," in *Suarez on Individuation: Metaphysical Disputation V: Individual Unity and Its Principle*, Milwaukee: Marquette University Press 1982.
- Hall, Roland, "Hume's Use of Locke on Identity," *Locke Newsletter*, Vol. 5 (1974), pp. 56-75.
- Haugeland, John, *Artificial Intelligence: The Very Idea*, Cambridge: MIT Press 1985.
- Hatfield, Gary, *The Natural and the Normative: Theories of Spatial Perception from Kant to Helmholtz*, Cambridge Mass.: MIT Press 1990.
- Hobbes, Thomas, *Leviathan*, ed. E. Curley, Cambridge: Hackett Publishing Company, Inc. 1994.
- . "The Elements of Law," in *The Elements of Law, Natural and Politic*, ed. J.C.A. Gaskin.) Oxford: Oxford University Press, 1990.
- . "De Corpore," in *The English Works* Vol. I. Trans. W. Molesworth. London: Scientia Aalen 1655/1839, 9.7.
- Hume, David. "Of Personal Identity," in *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, revised by P. H. Nidditch, pp. 251-63. Oxford: Clarendon Press, 1978.
- . "Appendix," in *A Treatise of Human Nature*. Oxford: Clarendon Press, 1978.
- . "Of the Immateriality of the Soul," in *A Treatise of Human Nature*. Oxford: Clarendon Press, 1978.
- . "Of the Immortality of the Soul," in *Of the Standard of Taste and Other Essays*. New York: The Library of Liberal Arts, 1965.
- . *A Treatise of Human Nature*, Oxford: Clarendon Press, 1978.
- . *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, ed. L. A. Selby-Bigge, revised by P. H. Nidditch, Oxford: Clarendon Press, 1975.
- . *Dialogues Concerning Natural Religion*, ed. N. Kemp Smith, New York: The Library of Liberal Arts, 1948.
- . "Letter to a Physician," in *The Letters of David Hume*, ed. J. Y. T. Greig, Vol. 1, Oxford: Clarendon Press, 1932, p. 12-18.
- Julesz, Bela, *Foundation of Cyclopean Perception*, Chicago: The University of Chicago Press 1971.
- Kemp Smith, Norman, "The Naturalism of Hume I.," *Mind*, New Series, Vol. 14, No. 54. (Apr., 1905), pp. 149-73.
- . "The Naturalism of Hume II.," *Mind*, New Series, Vol. 14, No. 55. (Jul., 1905), pp. 335-47.
- . *The Philosophy of David Hume*, London: Macmillan & Ltd 1963.

- King, Peter, "The Problem of Individuation in the Middle Ages," *Theoria*, Vol. 66 (2000), pp. 165-73.
- Kargon, R., *Atomism in England from Hariot to Newton*, Oxford: Clarendon Press 1966.
- Laird, John, *Hume's Philosophy of Human Nature*, London: Methuen 1932.
- Leibniz, Gottfried Wilhelm, *Theodicy: Essays on the Goodness of God, the Freedom of Man and the Origin of Evil*, ed. A. Farrer, trans. E. M. Huggard, London: Routledge 1951.
- Locke, John, "Of Identity and Diversity" in *An Essay Concerning Human Understanding* ed. P. H. Nidditch, Oxford: Clarendon Press 1975, pp. 328-48.
- . *The Correspondence of John Locke*, ed. E. S. De Beer, Vol. IV, Oxford: Clarendon Press, 1979.
- Loeb, Louis, E., *Stability and Justification in Hume's Treatise*, Oxford: Oxford University Press 2002.
- Lowe, E. J., "What is a Criterion of Identity?" *The Philosophical Quarterly*, Vol. 39, No. 154 (Jan., 1989), pp. 1-21.
- Martin R, and Barresi J., *The Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century*, New York: Routledge 2000.
- Mandelbaum, Maurice, "Locke's Realism," in *Philosophy, Science, and Sense Perception*, Baltimore: Johns Hopkins Press 1964, pp. 1-60.
- McCann, Edwin, "Locke on Identity: Matter, Life, and Consciousness," in *The Empiricist: Critical Essays on Locke, Berkeley, and Hume*, ed. Atherton, M., Maryland: Rowman and Littlefield Publishers, Inc 1999, pp. 63-89.
- . "Lockean Mechanism," in *Philosophy: Its History and Historiography*, ed. A. Holland, Dordrecht: D. Reidel 1985, pp. 209-31.
- McIntyre, Jane, "Is Hume's Self consistent?" in *McGill Hume Studies*, ed. D. F. Norton, N. Capaldi, W. Robison San Diego: Austin Hill Press 1979, pp. 79-88.
- . "Personal Identity, and the Passions," *Journal of the History of Philosophy*, Vol. 27, No. 4, (Oct., 1989), pp. 545-57.
- . "Hume and the Problem of Personal Identity," in *The Cambridge Companion to Hume*, Second Edition, eds. Norton D. F. And Norton, M. J, Cambridge: Cambridge University Press 2009, pp. 177-209.
- Nathanson, Stephen, "Hume's Second Thoughts on the Self," *Hume Studies*, Vol. 2, No. 1 (Apr., 1976), pp. 36-46.
- Noonan, Harold, "Locke on Personal Identity," *Philosophy*, Vol. 53, No. 205, (Jul., 1978), pp. 343-51.
- . *Personal Identity*, London: Routledge 1989.
- . *Hume on Knowledge*, London: Routledge 1999.
- Norton, F. David, "Editor's Introduction," in *A Treatise of Human Nature*, ed. D. F. Norton, M. J. Norton, Oxford: Oxford University Press 2000.
- Olson, Eric, "What Does Functionalism Tell Us About Personal Identity?" *Nous*, Vol. 36, No. 4, (2002), pp. 682-98.
- . *What Are WE? A Study in Personal Ontology*, Oxford: Oxford University Press 2007.
- Owen, D. and Cohen, R., "Hume on Representation, Reason and Motivation," *Manuscripta*, Vol. 20, 1997, p. 47-76.
- Patten, S. C., "Hume's Bundles, Self-Consciousness, and Kant," *Hume Studies*, Vol. 2,

- No. 1 (Apr., 1976), pp. 59-64.
- Pears, David, "Hume on Personal Identity," *Hume Studies*, Vol. 19, No. 2 (Nov., 1993), pp. 289-99.
- . *Hume's System: An Examination of the First Book of His Treatise*, Oxford: Oxford University Press 1990.
- Perronet, Vincent, *A Second Vindication of Mr Locke, Wherein his Sentiments relating to Personal Identity are clear'd from some Mistakes of the Rev. Dr. Butler in his Dissertation on the Subject*, London: Bristol Thoemmes Press 1991.
- Pike, Nelson, "Hume's Bundle Theory of the Self: a Limited Defence," *American Philosophical Quarterly*, Vol. 4 (1967), pp. 159-65.
- Pitson, A. E. *Hume's Philosophy of the Self*, New York: Routledge 2002.
- . "Hume and the Mind/Body Relation," *History of Philosophy Quarterly*, Vol. 17, (2000), pp. 277-95.
- Passmore, John, *Hume's Intentions*, Cambridge: Cambridge University Press 1952.
- Price, H. H., *Hume's Theory of the External World*, Oxford: The Clarendon Press 1940.
- Parfit, Derek, "Personal Identity," *Philosophical Review*, Vol. 80, (Jan., 1971), pp. 3-27.
- . *Reasons and Persons*, Oxford: Oxford University Press 1984.
- Perry, John, "Personal Identity, Memory, and the Problem of Circularity," in *Personal Identity*, ed. J. Perry, Berkeley: University of California Press 1975, pp. 135-59.
- . "Personal Identity and the Concept of a Person," in *Identity, Personal Identity, and the Self*, ed. J. Perry, Indianapolis: Hackett Publishing Company, Inc. 2002, pp. 119-45.
- Paton, H. J., "Self-Identity," *Mind*, New Series, Vol. 38, No. 151. (Jul., 1929), pp. 312-29.
- Quine, Willard Van Orman, *Word and Object*, Cambridge: The MIT Press 1960.
- Raynor, David, "Hume and Berkeley's Three Dialogues," in *Studies in the Philosophy of the Scottish Enlightenment*, ed. M. A. Stewart, Oxford: Clarendon Press 1990. pp. 231-51.
- Reichenbach, Hans, *Experience and Prediction*, Chicago: University of Chicago Press 1938.
- Reid, Thomas, "Of Identity," in *Essays on the Intellectual Powers of Man*, ed. D. Brookes, University Park: Pennsylvania State University Press 2002.
- Robison, Wade, "Hume on Personal Identity," *Journal of History of Philosophy*, Vol. 12, No. 2 (Apr., 1974), pp. 181-93.
- Rogers, G. A. J. "The Intellectual Setting and Aims of the *Essay*" in *The Cambridge Companion to Locke's Essay Concerning Human Understanding*, New York: Cambridge University Press 2007, pp. 7-33.
- Rorty, Amelie Oksenburg, "Pride Produces the Idea of Self': Hume on Moral Agency," *The Australasian Journal of Philosophy*, Vol. 68, No. 3 (Sep., 1990), pp. 255-69.
- Roth, Abraham Sesshu, "What was Hume's Problem with Personal Identity," *Philosophy and Phenomenological Research*, Vol. 61, No. 1, (2000), pp. 91-114.
- Schnall, Ira, "Constancy, Coherence, and Causality," *Hume Studies*, Vol. 30, No. 1, April (2004). pp. 33-50.
- Shaftesbury, Anthony Ashley Lord, *Characteristics of Men, Manners, Opinions, Times*, ed. J. M. Robertson, 2 vols. Indianapolis: Bobbs-Merrill 1964.
- Shoemaker, Sidney, "Personal Identity and Memory," in *Personal Identity*, ed. J. Perry

- Berkeley: University of California Press 1975, pp. 119-35.
- . “Self and Substance,” *Nous*, Vol. 31, Supplement: Philosophical Perspectives, 11, (1997), pp. 283-304.
- . Shoemaker, S., Swinburne, R, *Personal Identity*, Oxford: Blackwell 1984.
- . “Functionalism and Personal Identity,” *Nous*, Vol. 38, No. 3, (2004), pp. 525-33.
- . “Critical Notices. Reasons and Persons. Derek Parfit,” *Mind*, Vol. 94 (1985), pp. 443-53.
- . “Self, Body, and Coincidence,” *Proceeding of the Aristotelian Society*, Vol. 73, (1999), pp. 287-306.
- Simmons, Alison. “Are Cartesian Sensations Representational?” *Nous*, Vol. 33, No.3 (Sep., 1999), pp. 347-69.
- Strawson, Galen, *Selves: An Essay in Revisionary Metaphysics*, Oxford: Clarendon Press 2009.
- . *The Secret Connexion: Causation, Realism, and David Hume*, Oxford: Clarendon Press 1989.
- . *The Evident Connexion: Hume on Personal Identity*, Oxford: Oxford University Press 2011.
- Stich, Stephen, “Dennett on Intentional System,” in *Collected Papers: Mind and Language*, Vol. 1, Oxford: Oxford University Press 2011, pp. 71-91.
- Strawson, P. F. “Chisholm on Identity Through Time,” in *Language, Belief and Metaphysics*, ed. H. F. Kiefer, M. K. Munitz, Albany: SUNY Press 1970, pp. 183-86.
- . *Scepticism and Naturalism*, London: Methuen 1987.
- . “Entity and Identity,” in *Contemporary British Philosophy*, Fourth Series, ed. H. D. Lewis, London: Allen & Unwin 1976.
- Spinoza, Baruch, *The Collected Writings of Spinoza* Volume 1, trans. E. Curley, Princeton: Princeton University Press, 1985.
- Stevenson, G. P. “Humean Self-Consciousness Explained,” *Hume Studies*, Vol. 24, No. 1, (1998), pp. 95-129.
- Stern, Cindy Debra., “Hume and the Self at a Moment,” *History of Philosophy Quarterly*, Vol. 4, (1987), pp. 217-33.
- Streminger, Gerhard, “Hume’s Theory of Imagination,” *Hume Studies*, Vol. 6, No. 2, pp. 93-104.
- Stroud, Barry, *Hume*, London: Routledge and Kegan Paul 1977.
- Viger, Christopher, “Where do Dennett’s Stances Stand? Explaining our Kind of Mind,” in *Dennett’s Philosophy: A Comprehensive Assessment*, eds. Ross, D, Brook, A, Thompson, D, Cambridge (Mass.): The MIT Press 2000, pp.131-47.
- Waxman, Wayne, *Hume’s Theory of Consciousness*, Cambridge: Cambridge University Press 1994.
- Westfall, Robert, *The Construction of Modern Science*, New York: John Wiley & Sons 1971.
- Winkler, Kenneth, “All is Revolution in Us: Personal Identity in Shaftesbury and Hume,” *Hume Studies*, Vol. 26, No. 1, (Apr., 2000), pp.3-40.
- Ward, Andrew, “Hume, Demonstratives, and Self-Ascriptions of Identity,” *Hume Studies*, Vol. 11, No. 1, (Apr. 1985), pp. 69-93.
- Wright, John P., *The Sceptical Realism of David Hume*, Minneapolis: University of Minnesota Press 1983.

- Waldow, Anik, "Identity of Persons and Objects: Why Hume Considered Both As the Two Sides of the Same Coin," *Journal of Scottish Philosophy*, Vol. 8, No. 2, (Sep., 2010), pp. 147-67.
- . *David Hume and the Problem of Other Minds*, New York: Continuum 2009.
- . "Hume's Belief in Other Minds," *British Journal for the History of Philosophy*, Vol. 17, No. 1, 2009, pp. 119-32.
- Yandell, Keith, "Continuity, Consciousness, and Identity," *Hume Studies*, Vol. 18, No. 2, (Nov., 1992), pp. 331-47.
- Yolton, John, *Perceptual Acquaintance: From Descartes to Reid*, Oxford: Blackwell 1984.