Cognition and Language • A Series in Psycholinguistics

Time, Will, and Mental Process

Jason W. Brown

Time, Will, and Mental Process

COGNITION AND LANGUAGE

A Series in Psycholinguistics • Series Editor: R. W. RIEBER

Recent Volumes in this Series:

AMERICAN AND CHINESE PERCEPTIONS AND BELIEF SYSTEMS:
A People's Republic of China–Taiwanese Comparison
Lorand B. Szalay, Jean B. Strohl, Liu Fu, and Pen-Shui Lao

THE COLLECTED WORKS OF L. S. VYGOTSKY
Volume 1: Problems of General Psychology
Volume 2: The Fundamentals of Defectology (Abnormal Psychology and Learning Disabilites)
Edited by Robert W. Reiber and Aaron S. Carton

THE DISCURSIVE SOCIAL PSYCHOLOGY OF EVIDENCE
Symbolic Construction of Reality
Salomon Rettig

EXPERIMENTAL SLIPS AND HUMAN ERROR:
Exploring the Architecture of Volition
Edited by Bernard J. Baars

MEANING AND CONTEXT: An Introduction to the Psychology of Language
Hans Hörmann
Edited and with an Introduction by Robert E. Innis

THE PSYCHOPATHOLOGY OF LANGUAGE AND COGNITION
Harold J. Vetter and Robert W. Rieber

TIME, WILL, AND MENTAL PROCESS
Jason W. Brown

UNDERSTANDING MEXICANS AND AMERICANS:
Cultural Perspectives in Conflict
Rogelio Diaz-Guerrero and Lorand B. Szalay

VYGOTSKY'S SOCIOHISTORICAL PSYCHOLOGY AND ITS CONTEMPORARY APPLICATIONS
Carl Ratner

Time, Will, and Mental Process

Jason W. Brown

New York University Medical Center
New York, New York

Plenum Press • New York and London

Library of Congress Cataloging-in-Publication Data

On file

Chapter 1, "Introduction," is from J. W. Brown (1994). "Implications of Microgenesis for a Science and Philosophy of Mind." *Psychoscience, 1* (2): 87-99.

Chapter 10, "Action," is from J. W. Brown (1995). "Essay on Action." In K. K. Sinha and P. Chandra (eds.), *Advances in Clinical Neurosciences 1995,* No. 5. Ranchi, India: Neuro C.M.E. Reprinted by permission of the editors.

Chapter 11, "Morphogenesis and Mental Process," is from J. W. Brown (1994). "Morphogenesis and Mental Process." *Development and Psychopathology, 6*: 551-563. Reprinted by permission of Cambridge University Press.

Chapter 13, "Process and Creation," is from J. W. Brown (1996). "Process and Creation." In N-E. Sahlin and A. Andersson (Eds.), *Cognition and Creativity*. New York: Kluwer. Reprinted by permission.

ISBN 0-306-45231-6

A Division of Plenum Publishing Corporation
233 Spring Street, New York, N. Y. 10013

10 9 8 7 6 5 4 3 2 1

Printed in the United States of America

Philosophy should not be a ferocious debate between irritable professors, but a survey of possibilities.

ALFRED NORTH WHITEHEAD

Preface

Five years have passed since the publication of my last book, *Self and Process*, the central theme of which was the nature of change in relation to a microgenetic theory of the mental state. The book did not refer to the many changes that were occurring in my own life during the period when it was written but these changes were reflected in a more abstract way, and the conclusions of the book were pertinent to my life and my science. The book asked the question, is change something that objects undergo, or are objects the snapshots that change leaves behind as the mind's record of its passage? This might translate to the question, do I have the capacity to induce a change or am I the outcome of the occurrent state and its causal history? And, this question is *the* fundamental problem of free will.

A year after the book was finished, I was dining at the home of my good friends Detlef and Ingeborg Linke in Bonn. Detlef asked about my next project. I shrugged that I had no idea what to do and he volunteered to collaborate on a book about freedom. The topic had an immediate appeal both for personal reasons and as a natural continuation of the previous line of inquiry. The book would no doubt have been stronger for Detlef's contribution, but he decided to go off to China to write haiku instead (*Kaum gedacht, bist du zersprungen*, Janus, 1992).

Later Detlef would ask, "Jason, why are you so determined to write this book on freedom?" nicely capturing the essence of the problem. Since all roads lead to determinism, the challenge is not to repeat the arguments from causation but to find an unexpected opening for free will. The pivotal section of the work attempts to achieve this goal in the explication of agent autonomy—the causal independence of the self—through the virtual duration of the phenomenal present.

This book, as with each I have written, is a further stage in the growth of a theory of the mind-brain grounded in clinical case study, though by now its concerns have grown quite far from the clinical symptoms with which it started. What is at stake is no longer an extension of clinical theory into the domain of philosophical psychology but a discovery of those very principles on which the clinical phenomena, as well as the objects of philosophy, depend. In the words of the poet, the clinical material has progressed to a sense "of something far more deeply interfused."

For example, the topics of duration and succession that were central in the previous book are revisited and more deeply explored in the initial sections of this work. The mind-dependence of our experience of time and its relation to change in the becoming of the mental state influence the way we understand the past, the present, and the future. This, in turn, affects our concept of causation, which is a relation between sequential moments in the "flow" of time. We infer a past cause from a present effect, or we predict a future effect from a present cause. In object causation, the causal relation is between objects. In agent causation, it is between the self (will) or its intentional states and an object. A theory of time awareness is critical to the analysis of causation and agency.

We want to know what exactly is an action. Our knowledge of points in the brain where actions are modified by stimulation and excision is a good place to begin, but there is still no overall theory of an action in either neuroscience or psychology. The philosophical literature is replete with discussions of agents, acts, and intentional states that avoid an accounting of their psychological, much less neurological, underpinnings. A description of the properties of the self, its acts, and intentional states is a way of writing about them without specifying what they are. It is necessary to dissect the internal structure, the becoming, of the state rather than define the properties which distinguish that state from others. Specifically, what do the labels *agent*, *action*, or *mental state* refer to that would count as a theory of these entities?

If the self is interpreted in Humean terms as a bundle of perceptions or as an assemblage of modules or homunculi, and if the identity of the self is deemed to be illusory, it is the responsibility of the theorist to account for the basis of this illusion, why we have it, and how it differs from other illusions that are perceptually discoverable. The "illusion" of free will, or of the unity of the self, is too powerful to be dismissed as illusory simply because it is inconvenient to theory.

When the self that is lost in a collection of elements is regained by a memory system that revives past elements in a simultaneous present, or by a scanning device or binding mechanism that weaves the elements

together, these are ad hoc solutions to an artificial problem that is created by a psychology in which time has been eliminated. The unity of the self is linked to whole–part relations, and these can only be understood if time plays a central role in psychological theory. A duration is not constructed out of instants, a continuity is not imposed on mere succession. Time is not a "psychic addition" to a cognizing machine.

The use of extrinsic props to account for the change and unity lacking in theories of inert constituents is a clue to the decrepitude into which the theory has fallen. Goethe described the situation of present-day cognitive science very well when he wrote that "a great danger for the analytic thinker arises when he applies his method where there is no underlying synthesis . . . [and that] all his observations will only prove more and more an obstruction as their number increases." The unity is there from the beginning; it is the progression from unity to diversity that needs to be explained, not the other way around.

The process through which diverse facts actualize out of unity and potential takes us to the final sections which propose a continuum from developmental to mental process in a cascade of specification of parts out of wholes. The whole to part relation is not reducible to the analysis of a sum to its contents. In the analytic process, parts are *created*, not uncovered. The whole to part transition is the basis of becoming, and becoming is the image of process in the world. If the relations between objects in the human mind—which are the only objects we know—apply to events in the material world, the "laws" of the human mind, or the relations that govern this transformation, are ingredient in physical passage and fundamental in both mind and nature.

On sunny days, every step
Of thought blankets the moment,
Vanishing in the slow passage of time,
And I miss the company of men.

On the first breath of winter
The geese leave the pond.
Ice bends the willows,
Life sleeps in the soil.

You know better than I,
Names are like pebbles
Thrown out of the living core
On the dead crust of the earth.

The spirit of the world lies
In the bed of time,
Waiting to be heard
By those who listen.

Help me, friends, to know
The quiet in the wood,
What passes between men,
The pause between words.

J.W. B.

Contents

CHAPTER 1

Introduction

How exquisitely the individual Mind
(And the progressive powers perhaps no less
Of the whole species) to the external World
Is fitted—and how exquisitely, too . . .
The external World is fitted to the Mind
WILLIAM WORDSWORTH, Recluse, 63

I

This book is not so much an argument for the positions taken on the topics of its theme as an exploration of what those positions would be given the theme that it takes. The approach differs from the usual pick and choose applications of neuropsychological data to philosophical issues which insert clinical symptoms into philosophical arguments simply to illustrate or reinforce their claims. A symptom or deficit is snatched from its context as a fragment of defective behavior and used to strengthen an argument in which it has no authentic share. This is a fraudulent use of the clinical material.

A thoroughly clinical approach directly explores the pathological in search of a richer and more naturalistic conception of mind than either philosophical study or normal psychology alone can afford. The material itself is the ground of the theory. The symptom and its spatio-temporal context *evoke* an organic philosophy of the "abnormal," which is to say, they describe the infrastructure of the normal. In such a description, ideally, every symptom is coherent with the entire body of clinical ob-

servation. The coherence then becomes a philosophy of mind revealed by the pattern of its disruption.

I am aware of only two theories of the pathological that are consistent across their domains of observation—psychoanalytic theory, which is an account of disorders of the self and personality, and microgenetic theory, which is a new, largely untested account of the effects on cognition of brain pathology. As with psychoanalysis, which had its beginnings in the study of aphasia,[1] microgenesis owes much to observations in the neurological clinic. Microgenesis is a theory of mind and brain *process* based on patterns of symptom formation in patients with disturbances of language, action, and perception. A model of the organization of cognition in the normal brain is inferred from the symptoms of brain damage, their change over time, and their relation to pathology in specific brain areas. The theory stems from a dynamic approach to symptom formation and the change in symptoms over time, and a process-based approach to anatomy in relation to evolutionary growth trends. The mapping of symptoms to distributed brain systems and the ability to relate pathology to normal cognitive function, provided the basis for a process theory of the mind–brain state.

The picture that emerged[2] was of an intrinsic core that actualizes over phases from the past to the present and from depth to surface in mind and evolutionary brain structure. In this process, content fractionates from a unitary base into the different modalities. The content undergoes progressive articulation and terminates in the world of object perception. The mind–brain state is a process of becoming reiterated through life. Pathological data aid us in recovering early phases in cognition. The process reconstructed from these data is a progression from potential to actual, not from the primitive to the developed. Whitehead wrote: "clearness in consciousness is no evidence for primitiveness in the genetic process; the opposite doctrine is more nearly true."

The onset of the genetic process forecasts the object. The object does not develop out of sheer multiplicity; the goal is contingent at every phase. Damage to the brain at successive points in becoming exposes such phases or "levels" in the mental state. These phases are, in turn, mediated by brain systems that incorporate the responsible (lesioned) area. A symptom is a piece of preliminary mentation that stands for a phase in the mind of an observer-patient.

The philosophy of mind that derives from microgenetic theory departs from the strongly computational and linguistic approaches of current philosophy. There is a relation to the work of Henri Bergson, William James, and, especially, Alfred North Whitehead, whose controversial metaphysics may, I believe, be one day vindicated by microgenetic studies, but the theory is not linked to a particular school of philosophi-

cal thought. The philosophy originates in the symptom as a subjective datum. The symptom, the subject, and the subjective point of view are primary. This, I believe, is a strength of the theory.

Like evolutionary or, for that matter, developmental or psychoanalytic theories, microgenesis is a retrospective model. It describes how a present state or object came to be what it is. In certain respects, however, microgenesis is more fundamental. Phyloontogeny describes individual or group patterns that seem to be extended in time. From a microgenetic standpoint, these patterns reflect the reiteration of a single instance of becoming over different (evolutionary, lifespan) durations. Put differently, every organism is in a constant process of becoming that reinstantiates itself in some duration. Phylo-ontogeny is the pattern of reinstantiations *over time*. Microgenesis is the *time-creating* pattern of a single instantiation.

This distinction was obscured in the past by a comparison of the time frames of what was presumed to be a common process, i.e., eons in evolution, decades in maturation, milliseconds in microgenesis. Gradually, it became clear that microgenesis does not collapse phylo-ontogeny into milliseconds. The duration of, say, 100 msec. over which mental events unfold, is not a duration into which becoming deposits. This is not the (subjective) time it takes for the process to occur. Events are seriated in the fractionation to objects and this seriation creates temporal facts. The time for events is an ordered succession that is generated out of a simultaneous (timeless) core. In other words, the time of a becoming is the time the becoming creates. Becoming is succession without temporal incrementation. Every actualization is a whole unit of psychological time.

II

Unlike other "genetic" theories, microgenesis does not lend itself to teleological interpretation. It is firmly set in the present. Past, present, and future have their origins in the present state. The goals and motivations of psychoanalytic theory, the purposes and ends that have been imputed to developmental or even evolutionary theories, are not beacons in the future toward which current states are directed, but patterns of recurrence that observation extracts from constraints on emergent form. The "goal" of becoming is an actual (occurrent or present) object. The present does not move toward the future. The future is the next actuality. This next actuality is the ensuing present already developing over the residue of the occurrent state. Indeed, if mental states overlap, as is likely, the onset of the next state, i.e., the immediate past of that state,

is part of the structure of the occurrent state. Every present is re-created in a traversal from the distant to the more recent experiential past. Conversely, the personal past is the ground in which every new present is conceived.

For microgenesis, an explanation for an occurrent event lies in the personal past of the event. This personal past is buried in the present state. The personal past is not a causal sequence of events leading to the present but an experiential past revived in the depth or core of the present as the foundational segment of an irreducible, i.e., nontemporal, span of becoming. The aim of microgenetic analysis is to describe those phases in the revival from past to present that constitute the sum of mind for that moment.

This mode of explanation would seem to have the disadvantage of any historically based theory. We value prediction more than retrodiction. We value the future more than the past. Retrospective theories tend to be labeled as descriptive or hermeneutic and are distinguished from so-called *scientific* theories which are causal and predictive. The concept of time is deeply entwined with the aims of science. For example, prediction in microgenetic theory is an estimation of the effects of constraints on the next actualization, not the outcome of change an occurrent object undergoes. In this view, the immediate prehistory of objects is the source of change in the world, not as in causal science, the simultaneous effects of a present object on another object or its effects in the immediate future. To hold that a direction to the immediate future is scientific and an orientation to the immediate past is nonscientific is to decide on the most basic question of change in the world before its foundations are elucidated.

In causation, there is a fixed sequence leading to the present. This sequence continues toward the future. The causal relation is future-directed: a *present* cause leads to a *future* effect. What is assumed for physical causation is assumed for psychological causation. The imputation of a causal relation between psychological events, e.g., desires causing actions, beliefs causing desires, or the assumption that cognitive "solids" interact, are essential postulates of cognitive "science." Indeed, this is the basis for its claim to be scientific.

In causal theory, the present is an outcome of a chain of events in the past. The chain continues from the present moment toward the future. In principle, physical causation and causal change in relation to time are isotropic or reversible. In microgenetic theory, change is unidirectional and occurs in the process of object recurrence, i.e., in the becoming of an actual (occurrent) object, not in the transition from one (present) object to another (future) object. An account of objects entails an account of their becoming *from past to present*, the future being

imaginary. Objects do not lead to future objects. Objects do not change, they perish[3] and are replaced by near replications. Authentic change is not in the replication, i.e., the replacement, but in the process through which the replication occurs. An iterated replication of worlds does not account for the change between replicates. In microgenesis change is in the becoming of each world.

The topic of change is at the heart of a theory of causation and determinism and is ingredient in the debate on free will. A theory of time is fundamental to this debate, since time and change are inseparable. A theory of change in the world is a theory on the creation of time. Time and change have the same directional properties. From the observer's perspective, time is anisotropic. An action goes from the present to the future, so an account of how change creates the direction from past to present to future—"time's arrow"—is a necessary component of any theory of will and action.[4]

Free will is a voracious problem that touches every aspect of philosophy. The definition of will, agency, and the self; the nature and efficacy of mental contents (e.g., beliefs, desires, and intentions), morals and responsibility; and the brain basis of it all are a few of the topics that need to be covered. The relation of mind to brain determines whether the mental can be reduced to the physical. The concept of change establishes whether a reduction is possible[5] and if so, whether it is satisfied by a causal explanation. The nature of beliefs, desires, or intentions are problems for any theory regardless of whether these "folk psychological" concepts disappear in a reduction of mind to brain. Common sense may not provide the categories for a theory of cognition, but the categories it does provide need to be accounted for.

Ultimately, the question is, what is process in the world? If process is causal, is there room for freedom in a mind–brain reduction? If process is emergent or probabilistic, is this sufficient to account for acts of freedom? Probability is potentially reversible but novelty is not. Were a prior state to be *exactly* reproduced, the reproduction would violate a "law" of universal novelty. How do novelty and probability relate to causation? Probability, but not novelty, is consistent with causal theory. Is the relation between novelty and probability comparable to that between freedom and determinism?

Some of these issues have been discussed in my last book, *Self and Process*, especially that of the self-concept in relation to the hierarchy of the mental life and the continuity of intra- and extrapersonal entities over this hierarchy. Primarily, the book was a study of the nature of duration and the phenomenal present and an effort to extend the theory of time awareness to the problem of voluntary action. The psychology of time was a central part of its argument and remains the cornerstone

of the present book which continues the same line of thought more deeply into change, process, and will.

III

What sort of world is imagined when we think of a world as process? The world of such a thought—really, the thought of such a world—is first of all a mental picture that fills a certain duration. The picture is made possible by the duration. A duration is the beginning of a category or concept and a concept is the nucleus of an object. The innate capacity for duration and category formation is so fundamental to mind, it is part of the definition of what a mind is.

Durations, categories, and concepts span moments in the passage of nature. A mental picture of a stabilized world is a world of thought in which the images of everyday objects sequester in a world object that creates itself through the mind of an observer. This world is carved by sensation into entities in an observer's mind. Conversely, the mind of an observer is pruned by the environment to model the constraints "out there" in the physical world. The observer is one of the many objects in the world that he creates. But an observer is more than just another object, an observer is an object that is penetrated by subjectivity. However, before there is a subject, there is an object, which is a basic entity. What, then, is an object?

From the standpoint of process, an object is a local density that recurs the next moment more or less precisely. It is a configuration that persists across change as a combination of many such recurrences. This combination—the "summation" across moments in the life of the object—is the key to its persistence. The observer has an essential role to play if objects are to endure. This is because the series of near replications that constitute an object, in order to endure, requires the mind of an observer.

Whitehead remarked that the material world "labours under the defect that it can never be perceived." Yet it is reasonable to assume that objects would continue to exist if all possible observers were eliminated. This does not mean we have nothing to learn from mental objects. Mind is an expression of nature. The "laws" of mental process are, I believe, a species of the "laws" of physical nature. If the actualization of an object in the mind is an instance of actualization in the world, the becoming of mental objects provides an account of the becoming of material objects. The theory of microgenesis is driven by this belief. Whatever the fate of the theory, however, we have no choice but to examine the con-

tents of our own mind to know the objects of the world. Objects endure through observers. What, then, is more basic, an object or a subject?

The world is made up of objects and events. Objects are moments of change where some stability is achieved. Events are the vehicles of change in stable objects. I would say an event is a spatiotemporal discontinuity in object replication. If a replication is more or less exact, the chair remains a chair. If there is a discontinuity, the chair burns to ashes or is shattered or flung across the room. This is an event in the history of the chair. If objects are reinstantiations of more or less identical spatiotemporal configurations, and if events are discontinuities in such reinstantiations, is the difference between an object and an event the degree to which a configuration is approximated in each replication; i.e., is it a function of the distribution of spatiotemporal change?

The distinction of objects and events depends on a theory of change, specifically how objects endure, since events are created when objects cease to exist or when they change to other objects or affect them. The *distinction* between objects and events is mind-dependent since the endurance (duration) of an object, thus its stability or change, requires a segment of changeless persistence that is *thought up* in the imagination. Duration is not *in* the world. There is change in the world, there is a before and after, but there is no duration. It was Henri Bergson's genius to recognize the significance of this fact. A block present in the world may be a God's eye present that is everywhere, but without an observer it is still a durationless edge of passage.

For process theory, the concept of change is fundamental. There is change from moment to moment in brain activity and in the expression of this activity in normal and pathological states, as well as the interior flux of images and feelings. To know change, moreover, is to know its opposite, stability. For a theory such as process theory, which is based in change, the problem is stability. For a theory of stable objects the problem is how they change. A theory that is founded on stability requires change as an addition to objects. The addition of change is needed to explain the nature of an event. A theory that is founded on change requires duration for object stability.

We are accustomed to think of events as interactions or concatenations of objects. Similarly, we think of objects as the conscious or nonconscious "agents" of events, or in cognitive theory, of objects as contents or representations in the mind, and events as the effects of operations that are applied to them. We think that change is extrinsic to an object. The idea of a world as process is *mind-boggling*. Everything is in constant transformation. The world seems, to borrow Dante's phrase from another context, like "a ship without a pilot on a stormy sea."

How change is overcome as moments are "chunked" into entities that appear to interact is a complex topic. It is so complex that it is easier to assume that solid objects interact in the first place. This is one reason that objects are assumed to exist *ab origo*. Yet there is another, deeper reason for the belief in stable objects. Mind has evolved to stabilize change. The world is perceived through a distorting prism. The distortion is necessary in order to perceive the world. Objects are the observer error in the brain's instrument of observation. This observer error is built into the brain by evolution. The duration that leads to object categories becomes a belief in the existence of external objects. This belief requires a mind in opposition to a world. After a mind fills the world with objects, the next step is to populate that mind with objectlike contents.

An object-centered approach turns everything into an object. A thought is an object, a noun is an object. Once a thing is named or thought of, the thought or name of the thing endorses its separateness. Language and perception taxonomize the inner and the outer world. Process theory has to deal with this state of affairs, i.e., things and names as anchors for the objects of thought and perception. The fixation of change in solid entities undermines our ability to understand the entities themselves. The understanding is not to be found in the object but in the process of becoming through which the object appears. Objects are inert actualities. They are perceived at the expense of an awareness of the recurrence that is the basis of their existence. We live in an *actual* world, the sources and vitality of which are hidden from us. Objects are distractions from the underlying connectedness of all things, both in the mind and, presumably, in the world as well.

IV

The concept of time and change dictates the concept of objects and events and thus the implicit role of causal theory as an account of object relations. Object causation, say between two billiard balls, is not, as it seems, an event in the material world, but an event between mental objects since all known objects are mental. Physical causation is an inference about a material world that is the presumed basis for mental objects. The belief in object causation may be derived from the feeling of "mental causation" which concerns the interaction between concepts or, at the earliest stage in development, between an agent and an action. Indeed, the concept of causation in the world of physics is an extension to science of the discourse of human agency.[6] If object and mental cau-

sation explain occurrences in mind and perceptible nature, their difference is the role of agency and freedom in a causal world.

In contrast, a subjectivist account of agency entails an *inner* story about agents, actions, beliefs, desires, choices, reasons, ends, and so on. The story would detail the nature of these contents, where they begin, where they leave off, their derivations and boundaries and their relations to other mental contents. Take the example of an intention. It is not enough to say that an intention is a state in which there is a direction toward an object. The direction might be an epiphenomenon that accompanies the object as it becomes clearer. That is, the direction becomes apparent when a content reaches a certain phase in its actualization. We speak of acts and actions but what exactly is an action? Are the plans that precede voluntary actions part of the action, are they thoughts, actions proper, or perceptual contents? One cannot decide what is an action without an account of the microtemporal features that constitute an act in its entirety. The nature of acts, agents, and intentions will determine whether there is interaction between them, and, if so, whether it is causal.

There is a difference between a philosophical approach to these problems and that of microgenesis. A philosophical question concerning, say, belief, might deal with its penetration by language or the nature and truth value of statements or propositions, while a microgenetic approach asks what is its momentary history, its before and after, and the correlates of belief mentation with brain process. Philosophy slices mind for its static architecture. Microgenesis takes the continuum of mental process as the fundamental reality.

This approach can lead to new insights on the nature of freedom, but this depends on what sort of freedom one is talking about. Generally, since Hobbes, freedom has been defined by the effects of constraints on action, i.e., in limiting one's options. This definition of freedom is centered in an object theory of the agent. It has nothing to do with interior events in the generation of a volition. The interior story can claim that an action is not necessary to decide if the will is free. The willing in a freely willed act is similar to the willing in a freely willed thought or image. Even if an action is compelled, there may be freedom in its timing. Suppose I am instructed to kill someone against my will but am told I can decide when and where the killing will occur. It seems fruitless to debate what elements of an act are free and what elements are not, or the degree of freedom in a given act. Extrinsic constraints are mere impediments. We want to know whether *any* act is free. The freedom in free will is not in the exercise of freedom—not in the action that follows a decision—but the ability to decide independent of whether the chosen act or any act occurs.

Freedom is also defined by certain properties. When such properties are found in a behavior, it is said to be free. This property approach, however, does not get at the interior events. A somnambulist might act in a manner consistent with a volition except for the lack of later recall. If the descriptive properties of a freely willed act include its later recall, say, as a feature of conscious as opposed to automatic or trance state cognition, and if lacking such recall the act is judged to be unfree, what about the same behavior in a person with forgetfulness who has full consciousness of an action, and acts in a way that would be termed free, but later does not recall it? To an observer, a freely willed act and its simulacrum may be difficult if not impossible to distinguish. Moreover, the attribution of freedom to an agent on the basis of an outward description of his act is highly circular. This definition of freedom follows from everyday experience. We have an experience that we label as free, we use the properties of that experience to define freedom, and then term a behavior free when it exhibits those properties.

A theory of time, change, and the mental state determines what concept of freedom, if any, can be tolerated within the confines of the theory, which establishes the limits of agency, thus the limits and meanings of free will. From the standpoint of microgenetic theory, the freedom in free will is the ability of an agent, i.e., the self in a state of agentive awareness, to choose among options or to decide to act or not to act. This definition holds even if the conscious self as an agent is not fully autonomous or "free standing," but is a chosen self, emerging out of the deep self as one of many potential selves that might have developed. If the conscious self is a product, and if concepts generated by the self prefigure its actions, an action is biased by the dispositions of the chosen self to generate the options that it does. The options are almost irrelevant. The self and its choices are delivered into a decisional state.

The problem of free will is as much a part of a planned or deliberative[7] action of the finger as the choice of a summer vacation. Complexity augments the feeling of agency through incremental discharge but is not essential for it. More important is autonomy. If the conscious self is a product of an unconscious core self, does the conscious self function independent of its preliminary phases? Can the self initiate, guide, or veto an action? If so, is dualism necessary for control? From the standpoint of the agent, these questions are independent of extrinsic constraints or coercions. The self, as Kant argued, is not driven by external conditions but by its own internal state.

Moreover, even internal constraints such as drug addiction, habits and hypnosis are inessential. For the libertarian, extrinsic constraints block or inhibit the freedom of action, while intrinsic constraints are

not determining. For microgenetic theory, extrinsic constraints merely amplify or disinhibit intrinsic ones. The effect of the intrinsic is always primary. Every constraint limits one's options,[8] a gun to one's head or one's skill in karate no less than one's height, weight, mortality, and the absence of wings, but the number of options is irrelevant to whether any option, in this construal of freedom, is freely chosen. Otherwise, we are fated to incessantly quibble over the degrees of restrictions, coercions, lack of opportunity, genetic predisposition, failed upbringing, and so on. We want to know if free will exists, its criteria and parameters, not the quantity of freedom in a given circumstance.

Change evokes stability which in turn is a brake on change. Freedom demands responsibility which in turn is a constraint on action. The topic of responsibility is central to any discussion of free will. There is a political, if not theoretical, mandate to presume that an individual is responsible for his or her acts. All human intercourse depends on this assumption. Indeed, the effort to assume responsibility for one's volitions regardless of their freedom, e.g., arguments for individual responsibility in compatibilist accounts of will, pervades much of the literature on this topic. What, then, is responsibility?

One can ask, is there a primitive moral sense[9] that generates a feeling of responsibility? Certainly, it seems likely we have evolved with some social instincts or patterns of behavior derived from biological dispositions. For example, such a pattern might correspond with a type of loyalty in which self-interest is sacrificed for, or realized through, the group benefit. This could originate in instincts that preserve the integrity of a family or a pack, parental instincts, or the hierarchic relations among members. The sense of responsibility to oneself for acts that are in conflict with the group or independent of its values, could originate in the evolution of instincts of social cohesion to cultural valuations, which gradually internalize to accompany the growth of the self-concept. In the shift from core disposition to social cohesion, i.e., from wantonness to responsibility or from selfishness to submission or compassion, etc., the developing self-concept appropriates cultural attitudes. These attitudes infiltrate the drives and create a personal valuation that is deemed constitutive or defining by the individual.

The fact that evolution gives us "moral" dispositions, however, cannot anchor the "truth" or certainty of a given set of values. Evolutionary dispositions influence behavior because they have survival value, not because they are ***values***. Microgenesis is consistent with value relativism in that values are learned adaptations of inherited dispositions. But the range of what can be learned within the confines of the dispositions is so wide that one cannot justify what is learned by what *can be* learned or, to put it differently, by the way that learning fractionates the drives.

Thus, the core self is carved up by social values through learning. Certain values are central (constitutive) to a definition of the individual, others are peripheral to the self or define the society. But all values, regardless of their positive or negative valence, have their locus in the self-concept. When responsibility to one's self, whether to altruism or self-gratification, supersedes that to another person or to a group, one set of values, those perceived as constitutive, is privileged over others, and the privileged set biases the options and choices of the "free" agent, regardless of how the sets are conceived, i.e., in terms of strengths, hierarchies, and so on.

Responsibility, therefore, is the feeling of allegiance to values that are apprehended as personal, i.e., constitutive, or cultural, assimilated to the self-concept, thus constitutive, or in conflict with it, thus enforced in learning, censure, and punishment. Values internalize as the feeling of responsibility to one's self and others independent of reasons which are in my view not causes but justifications. Core values parent beliefs and desires and shape individual and group action. The importance of responsibility to freedom reduces to the nature of values, their relation to action and to the self-concept. Ultimately, the self's feeling of responsibility or the judgment by others of one's behavior as responsible, is the resolution of the inevitable conflict that inheres in the self's own valuations.

V

This, then, is an outline and preview, perhaps for some a warning, of the perspective that is brought to bear on the main topics of this work, namely, the relation of time and change to the mind–brain state, the relation of duration and category formation to objects and acts, specifically to the "structure" of voluntary and automatic action, and agency, the role of value, belief and desire in action generation, intentionality, and mental process. My goal has been to search out a meaning of agency that conforms with a theory of mental process, i.e., microgenesis, that is coherent with the clinical data and its implications for a concept of subjective time.

At rock bottom, the problem of free will depends on the boundedness of mind to brain process, and whether mind and/or brain events are causal or emergent. If mental events are tightly bound (identical or epiphenomenal) to physical brain states, i.e., if mental events are irrelevant to a physical description or eliminated altogether, and if the physical brain states are instances of causal change in the world, there is no *intrinsic* freedom. If the "laws" of mental activity are a species of natural

"law," free will requires either a loose construal of causation, which is probably something other than causation, or a different account of change, i.e., as noncausal.[10] I am inclined to the noncausal option. On this view, free will obtains as a variant of the universal change exhibited by complex systems in the actualization of wholes to parts. The alignment of change in the mental state with change in the material world, and the implications of this alignment for free will, entail a self or agent that exists to serve, sustain, and voice, perhaps even in some sense exploit, but never veer from the silent will of nature.

Surely, there will be readers who would have wished for more neuroscience from a neurologist or, perhaps, less or better philosophy from a poacher in a neighbor's field. They may also interpret these failings as a lack of explicitness on topics of personal concern. On my part, there is a preference for intuition over exegesis and for the creativity in potential, in context and allusion, over the dry bones of the actual. But there is more to it than this.

If change is realized in the becoming of whole to part, there is a surround of indefiniteness that is ingredient in the description of every part. The part is a part of some field and the groundlike quality of the field, i.e., its potential to give rise to the part, is the warrant for its indefiniteness. Every object has this "complementary" character. Indeed, there have been times in the writing of this book when ambiguity seemed the model of clarity. This feeling is not unknown to those working in purely physical science.[11] It is not, therefore, in the spirit of mysticism but of scientific inquiry to suppose that every elucidation conceals a deeper layer of uncertainty. To seize or at the least glimpse the mystery of that is for me and, I hope, the reader, a goal worth pursuing.

NOTES

1. M. Solms and M. Saling, "On Psychoanalysis and Neuroscience: Freud's Attitude to the Localizationist Tradition," *International Journal of Psycho-Analysis* 67 (1986):397–400.
2. J. Brown, *Life of the Mind* (Hillsdale, NJ: Erlbaum, 1988); J. Brown. *Self and Process* (New York: Springer-Verlag, 1991).
3. After A. N. Whitehead, *Process and Reality* (New York: Macmillan, 1929).
4. The connection between time and agency was most forcefully argued in H. Bergson, *Time and Free Will*, trans. F. L. Pogson (1889; reprint, London: Swan, Sonnenschein, 1923).
5. *Self and Process*, 171–173.
6. B. van Fraassen, "Discussion," *Philosophy and Phenomenological Research* 53(1993): 442.
7. Frankfurt's witty discussion of de-liberation as the loss of liberty of the drives in the gaining of the autonomy of reason leads him to conclude, as do I, that "reason de-

pends on will." H. Frankfurt, *The Importance of What We Care About* (Cambridge: U.K. Cambridge University Press, 1988).

8. On this point, see R. Kane, *Free Will and Values* (Albany, SUNY Press, 1985).
9. J. Wilson, *The Moral Sense* (New York: Free Press, 1993).
10. The position is that free will exists but is incompatible with universal causation. Some philosophical arguments for this view can be found in P. Van Inwagen, *An Essay on Free Will* (Oxford: Oxford University Press, 1983). Van Inwagen bases his argument for freedom on the reality of moral responsibility. I would consider moral responsibility an *outcome* of a proof of the existence of free will, not an argument for it.
11. For example, David Bohm notes that an essential contribution of Niels Bohr to quantum theory was to bring the "ambiguity of *meaning* (which we ordinarily associate only with the mind) into a crucial role in the understanding of the behavior of matter." David Bohm, "Soma-Significance: A New Notion of the Relationship between the Physical and the Mental," *Psychoscience* 1(1994):6–27.

    From a somewhat different perspective, Kane, *Free Will and Values*, has written that, paradoxically, compatibilist and deterministic accounts of behavior have grown in influence during the same period that support for determinism has eroded in the physical sciences.

CHAPTER 2

Change

ARGUMENT: *An object is a process of becoming actual that gives way to another object. An actualization creates temporal facts out of simultaneity or timelessness. Authentic change occurs in the becoming of the object in a mind that perceives the world. Apparent change seems to occur between existing objects in the world. Authentic change is novel or emergent. The idea of causation is inferred from apparent change as a theory on the succession of objects in the course of their replacement.*

> All my life, whether in poetry or research,
> I had alternated between a synthetic approach
> and an analytic one—to me these were the
> systole and the diastole of the human mind.
>
> GOETHE

CHANGE AND MOTION

Since Aristotle, the problem of change in relation to time has been closely linked to the problem of relative motion; is motion equivalent to change and if so, what is the relation of motion and change to time. Movement is a type of change but what type of change is it, and in what sense is change movement? Movement requires something that moves. Do objects move? This seems an odd question. Movement, being change in position over time, depends on a theory of space and time. What happens to movement if change gives rise to objects that are con-

tinuously disappearing? When would a movement occur? Can an object move if it vanishes as soon as it appears?

Movement is a change in position, not just a change. To ascribe every change to a change in position (though not the reverse) is to spatialize change. When a thing moves it is not the same entity in a different location. Every movement is a change, but does every change involve a movement? One might think an object could remain unmoved while the setting around the object shifted. The object would change its position without moving. This is a problem of relative motion between part-objects. Every change in a part is also a change in a whole. The world cannot move and displace a motionless object.

The microgenetic concept of movement is that movement is not the change an object undergoes but the discrepancy between two instances of the "same" object, or an "altered" object, when the initial object (state) is usurped by another in a different space, in a different time, in another momentary world. The motion of an object from one position to another is, in truth, the actualization of successive worlds in which different positions of the object are embedded.

In prerelativity, the fact that objects moved or changed at different rates, suggested a standard (Newtonian) absolute space and time to which they could, in principle, be compared. This also conforms with—indeed, was generated by—the commonsense experience of personal time. The apprehension of duration and time passing, the cycle of circadian rhythms, vary so slightly from a presumed standard (clock) time that time seems less an average of the variance than an external flow to which subjective experience is an approximation. The impression of common sense is that the private experience of time is closely subordinate to an external or objective time that is a property of the physical world. Certainly, there is no widespread intuition that time is a subjective or mind-dependent phenomenon that is imposed on the physical world or that objective time, even if it exists, is unknowable.

The Nature of Relative Motion

In the park where I am sitting writing this chapter, everything is moving at a different rate. Someone lying on the grass not moving does not have a different time to my perspective than someone chasing a frisbee. In my perception, the sun and the grass, the people, the dogs, all are acting at the same time. Since each object has a different rate of motion, and presumably a different rate of change, my perceptual experience persuades me I am participating with these objects in a universal time

in relation to which all local change, including my own, can be measured.

With the advent of relativity theory, time became the measure of motion in relation to a reference in a system of space–time coordinates. Lovejoy has written that in early relativity, "no one reference system was any more real or true than another, and apparently a reference system could be determined only by an arbitrary setting-up of coordinates, i.e., by an act of mind".[1] For Einstein, objectivity was the invariance across observers. Still, as Lovejoy notes, to the extent the world varies with perspectives, it is not objective.

In relativity theory, the speed of light is a reference for the distance traveled, or the change in position, in a system of coordinates. As a dimension of space–time, time is eliminated as a distinct phenomenon. What then happens to change? The position of an object is changed in every change, and a change in position is a change in space. It is impossible to examine change independent of space. But this does not mean that change (or time) is equivalent to change in position or relative motion in a given system of space–time coordinates.

If the observer is the reference for a change, or the comparison of a pair of changes, the rate of change will depend on the relative movement of one object to another in a triadic relation to the reference. In this way, movement is brought into relation with change. But every object in my perception is simultaneously in motion. A figural element isolates in a field and the relation between the field and the figure is part of my judgment of relative motion. Since I am the reference for the whole field of this perception, indeed, since the whole field *is my perception*, every change in the perception is generated by the potential from which I as an observer develop. Objects and perceiver develop in the same perception (mental state). The motion on which time seems to depend is the changed world that deposits in every mental state. Motion or relative motion is not time. Motion is the filling-in between changed objects in the duration of the subjective present.

Apparent Motion

One might suppose that if change were motion, motion should occur in all change. Yet a line drawing of a (Necker) cube does not move as it flips from one orientation to another, though the figure changes in the illusory reversal. An object that alternates from a right to left orientation occupies time but does not actually (in the ordinary sense) move. In the phi phenomenon, illusory movement occurs between adjacent alternating light sources. The "filling-in" of apparent motion is often con-

trasted with the perception of real motion, just as the mental is contrasted with the physical. However, the difference between motion that is "apparent" and motion that is "real" is that in the former the subjective contribution is more emphatic. The filling-in of apparent motion is a clue to the basis of the perception of veridical motion. The filling-in is a completion across actual objects that bridges the (timeless) gap from one mental state to the next.

We speak of real and illusory space but on a theory of a continuity from the intra- to the extrapersonal, an illusory change in spatial position is linked to perceived movement in the world. An apparent movement can seem real, while a real movement can take on illusory features. The apparent motion of a real object can seem real or imaginary. So can the apparent motion of an apparent object; e.g., the movement of an hallucination. What is the nature of real motion in an apparent object? What, then, is real motion in a real object but a collusion of the senses in establishing its reality or inattention to the subjectivity from which the object springs.

It is not clear whether the apparent (imaginary) movement of a real object occurs in a real or a mental space. Wittgenstein wrote, "when I write in the air, am I writing in my mind or in the air?" The distinction between private and public space is fuzzy in everyday life, not just in pathological cases. The terms illusory and veridical tend to freeze, to amplify and polarize this distinction. That is why it will not do to say the cube or the light source is real and the reversal or apparent motion an illusion. Which orientation of the cube is the "real" one? In what space does the apparent movement occur?

In affections of the brain or disorders of the middle ear, "real" objects undergo illusory changes. For example, an object can become nearer or farther, contract or expand, rotate, invert[2], etc. What is real and what is apparent is not always obvious. If apparent movement is a change without a movement, how does apparent movement differ from other movements or other changes in relation to time? In fact, the 2 to 3 second cycle for the reversal of a cube illusion has been associated with the duration of the present.[3] Does this imply, as I believe it may, that the change that occurs in illusory movement is more closely related to the awareness of time, and to authentic change, e.g., in the example of completion across proximate states, or their basis in subjectivity, than the apparent change that occurs in the career of an external object?

An object moves in perceptual space. An apparent movement is a movement in an apparent space. A motionless object, say a hypothetical electron fixed in orbit, would still undergo a change in position by virtue of the change in the objects around it. Objects in the mind, and physiological processes in the brain, are constantly changing their position,

in my perspective and in the physical universe. A change in position is a condition of motion. But a change in position is nothing if not a comparison between a prior and a present actuality. Motion is assumed to be the vehicle of this change. But the essential fact is the *comparison* on which the judgment of motion depends, and a comparison requires a subjective duration. Change in position and change in motion are not direct perceptions but different interpretations of the discrepancy which the comparison reveals. A comparison over instances in a determination of objective motion requires the persistence of the initial position. This entails a duration. The relation is reciprocal. The duration enables the comparison, while an implicit comparison, e.g., across the limits of the present, generates the duration. The interpretation that issues from the comparison, such as that motion or change occurs when object A becomes object B, is a commonsense account of the connectedness between successive actualities.

Successive objects or object fields *have* to be different for there to be a duration. Without different objects, or changes in the "same" object, there would be no object or event to anchor the duration. A lack of change over a duration is no less inconceivable than a certain duration of timelessness. Time is in the *comparison* of changed events or objects or positions, not the change itself that occurs from one event to another, even if the judgment of a duration cannot occur without a change. It follows that if time is not, strictly, a change in position but a comparison of changed positions, and if there is no absolute reference for local change, then:

1. Change is necessary for time but not identical with it, assuming there is not an infinite number of times. This follows unless everything that changes has its own time, and/or
2. Time is illusory. This implies the absence of time independent of mind, or the subjectivity of all time, and/or
3. Change and time do not differ across objects, but rate of change and relative motion are mind-dependent. Physical passage is uniform from one "instant" to another while one object (state, position) that changes more rapidly than another is a comparison across subjective moments.

DURATION

Duration[4] requires the suspension of a before and an after. Yet a before that is retained until the after appears is instantaneous with it. How could a past event "held" by the present assume its pastness? Reflex or

feedback systems do not generate a duration in mental experience. In thermostats or conditioned reflexes a preceding event alters an ensuing event. The prior event "persists" as an alteration of the present, not as a suspension of the past to the present in a duration of time. One cannot add instants together to construct a duration. Duration is fundamental.

In scientific writings, objects (instants, mental contents, etc.) are treated as durationless. Change is introduced as an extrinsic element, e.g., an output, an occurrence *between* two instants, or an event at the forward interface of a (solid) object. Whitehead wrote that for science, change is "the importation of the past and of the future into the immediate fact embodied in the durationless present instant." He went on to say, "the ultimate fact for observational knowledge is perception through a duration; namely, that the content of a specious present, and not that of a durationless instant, is an ultimate datum for science."[5]

Change occurs within a duration but changing events within that duration are simultaneous with respect to each other. We see a world that is changeless in relation to events in the world but we are struck by the simultaneity of the whole field of events as it is changing. The change of the whole field is its reconfiguration; every element that changes, changes the whole field. The forward movement of one object is the reverse movement of another. An illusion such as the apparent motion induced on a stationary train when another train rolls by, illustrates the dependency of one motion on another. This dependency underlies (is buried in) the experience of veridical motion.

The simultaneity obtains because it constitutes a temporal "slab" of perceptible nature. The capacity to form such a slab is the capacity of duration to stabilize change in a simultaneity of all change within the slab. This is how nature is witnessed. Whitehead wrote that we recognize (interpret) nature as atomic, i.e., as a sequence of durationless instants, but apprehend (intuit) it as continuous. The problem of duration is the problem of how atomic entities are created from the continuum of process.

Change and Duration

Things and events resolve from process in a world of appearance that actualizes out of change. The things and events we perceive are not as we perceive them, while what we are unable to perceive, the process chunked into the appearances, is the fundamental reality. The change we measure is an artifact of a perceived concatenation of events, while authentic change is concealed within the momentary origins of the events themselves.

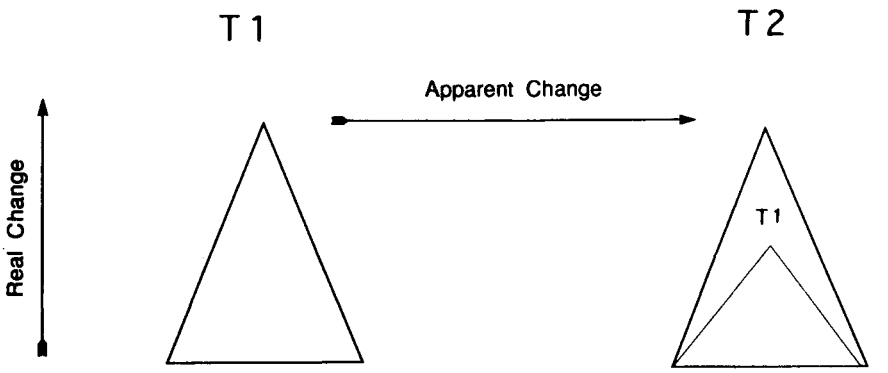

FIGURE 2.1. Intrinsic change is in the "vertical" sequence (becoming) from the base to the surface of the mental state at T1 and T2. Apparent change is the "horizontal" difference in an object across the replacement of mental states at T1 and T2; i.e., the replacement of T1 by T2.

There are, then, two types of change: (1) authentic or intrinsic change over phases in becoming; and (2) apparent change from one mature (actual) object to the next. The first type is a context–item transformation of qualitative forms without incrementation. The second type is an empty (timeless) gap between the disappearance of mature objects (worlds) and their replacement.

In authentic change, the transition is over the momentary lifespan of the object. For Whitehead, this transition proceeds from a conceptual to an actual object. The development is ordered and atomic, i.e., atemporal. The atomic nature of becoming accounts for the atomic (though unbounded) nature of duration, since every duration is created within a becoming. The distinction of successive (before, after) phases in the coming together or "concrescence" of an actual (final) entity through an atemporal sequence is paradoxical.[6]

Succession becomes precedence when becoming terminates since precedence requires the duration the becoming creates. Duration is foundational, thus prior to the existence of an object. Successive phases are collapsed within the thickness of the absolute present with precedence extracted secondarily after duration is established. The precedence in a becoming gives a duration without incrementation, even though the awareness of the precedence depends on the subjective time that the becoming creates. The change over becoming is the invisible undersurface of passage. Wordsworth wrote, the "time in which the pulse of Being everywhere was felt."

The *absolute* present, the duration of becoming in the human mind, is not experienced in consciousness. Time awareness entails a

discrepancy between the complete becoming of an object and the incomplete becoming (decay) of past objects revived in the present one[7]. The actualization of an object is graded to give a *phenomenal* (specious) present that appears to span successive objects. The duration of the phenomenal present (now), i.e., the span across apparent objects, is extracted from disparities in the "stacking" of serial images in the absolute present. In spite of the impression of a (short-term memory) span across objects, the now is elaborated in the becoming of a single object in the occurrent state.

The duration from potential to actual is the absolute present of an object. This momentary present spans the birth, growth, and death of the object. The object is renewed in every actualization. This renewal, the transition over phases in becoming, is the process of authentic change in the world. The substrate of this process, a context–item transform, is related to Whitehead's concept of the concrescence of entities through a many–one process. The becoming is reiterated in overlapping waves, with the object the complex of the reiteration.

What is termed apparent change is really an absence of change between mature objects. Since the absence of change is, arguably, the absence of time, the "interval" or boundary condition between objects is timeless, and therefore nonexistent. The timelessness between objects helps to bind them together in the continuum. The continuity across actual objects, i.e., from one state of the world to the next, is a result of the changeless "interval" across becomings so that successive states are collapsed together with a filling in by illusory change, as in the phi phenomenon, of the "interstices" of world states.

Newton-Smith[8] has argued that a world in which objects appear, disappear (cease to exist) and reappear in a different location is consistent with the hypothesis of time over a changeless interval. The problem with such a world, however, is the inability to extract seriality without anchoring the change to successive appearances. It is not possible to account for the transition from one appearance to the next. When could change take place but in the microprocess of each reappearance. In microgenetic theory, objects appear and disappear but every appearance is an actualization from past to present. This actualization gives a direction to time and links it to the change that occurs in the actualization process. A lack of change between actualities with an appearance of continuity between them, precisely because there is no intervening change or time, is inconsistent with the hypothesis that a period of time could transpire in which no change occurs.

Apparent change (change between object appearances) is vacuous, but the replacement of one object by the next is interpreted by a subject as the very essence of change. The object is perceived (imagined) to

undergo change, while the change that deposits the object is imperceptible. The timeless, changeless interval between final objects is construed as change. The change in this (false) interval is construed as movement; and movement and change together are taken to mediate the transition between objects (instants) in temporal succession.

Now, there is "change" in the sense of a shift across final objects. An object differs from the preceding object. But the change does not occur in the shift. The changed object is a novel object that has actualized in the decay of the old one. The perceiver fills the interval with imaginary change.[9] The interval does not exist "in reality." The completion across this imaginary interval is the underpinning of veridical motion. The duration (interval) over which change or movement is perceived to occur is within the duration of the (specious) present. This duration spans changing appearances, even though the lack of change and the absence of time between actual objects make it all but impossible that the interval actually exists.

What Is an Object?

An object is a complex in a direction of becoming that deposits the world at its terminus. Other worlds achieve actuality as endpoints or facts in the specification of potential. Facts are actual entities. An actual entity is a fact about the world. The world is the factual context in which the entity occurs. If becoming is continuous and can terminate at any point within a given actualization, there is the possibility of an infinite number of mental worlds. In principle, however, the possibilities are finite and few in number. Pathological studies suggest there are only so many mental worlds the (human) mind can generate.

On the microgenetic view, a mental object corresponds to the world not as a copy or an approximation but as a perspective. The perspective is an image of a world. Since any perspective is mental, a sum of all perspectives gives a composite image but still does not give the object. The perspectival, and imaginal, basis of objects is evident in everyday perception. If one alternately closes and opens first the right then the left eye, the monocular object will shift from side to side until, with both eyes open, it fuses at a site in between the monocular perspectives. It is obvious that the "real" binocular object is virtual. The three-dimensional object does not correspond to what either eye sees (also a virtual image) but is based on the disparity between the eyes. Parenthetically, the creation of a virtual spatial image from an ocular disparity is comparable to the creation of a virtual temporal image, i.e., of duration,

from the disparity of surface and depth in the becoming and decay of objects.

Not only is an object a virtual image but the image that is the object "reflects" the real world as a kind of mirror. Perception is an adaptive process that is shaped by the physical world. The object is a residue of unrealized potential. In mind as in evolution, organic form is parsed by the environment to give an entity (object) that survives. In evolution, an organism competes to survive and reproduce. In perception, an object is sculpted to actuality. What survives in perception is likely to be reproduced in a subsequent percept. This is the coherence, the causal-like continuity, of evolution and of perception. The perceived object is not a substitute for a real object "out there" but is its complement. The object fills a sensory niche defined by limitations. The world of perception is a *negative* image of externality.

What then is an object? I am an object, so is the world. The world object grows out of the space of my body. Every object is a part-object in a context. What is part of an object and what is an object? A leaf is part of a tree; a tree is part of a field. I am part of the field of the tree. A leaf on the ground is a separate object; a piece of a leaf is an object. Where do objects begin? When do parts become wholes? A discontiguity of part and whole is a condition of individuality.

The individuality of an object depends on the modality in which the object is perceived. An auditory object is not lost in the interval between sounds but a visual object will disappear if there is an "empty" interval between presentations. The silence between sounds is the background of an auditory object, as the space between sights is the background of a visual object. The difference between sounds is the change that auditory objects undergo. This difference is equivalent to the replacement of visual worlds.

What, then, is the meaning of a discontiguity? Every individual object implies the ground of a larger whole. The object is not individual to begin with but is realized as an outcome of process. Objects are part-objects that isolate through the analysis of wholes and their meanings. An element (object) acquires its individuality from some portion of the meaning allocated to the whole (universe).

Space and time are relations of adjacency and precedence. The object is defined by these relations. An infant must learn that adjacent or overlapping objects are not part of the same object, that an object can change in the sequence of its appearances but remain the "same" object, or that an object persists in spite of transient obscurations. These are stages in object individuation. Meaning develops as the spatial and temporal boundaries of an object resolve. In a sense, an object that moves from here to there is a different object, for the world of "there" is a

different world. The world changes when an object moves. Every change in position is a changed world. The identity or invariance of objects across change, for example with a moving or recurrent object, is linked to the meaning of the object. The object has to be *recognized* to remain the same object.

A chair is the set of its conceptual and spatial relations. Object boundaries are salient, i.e., meaningful, features. An object has to mean something to be perceived. In normal cognition, a chair is not perceived without being recognized. The chair and its meaning are part of the same percept. An object that fails to be recognized, e.g., an unfamiliar or impossible object, or a nonsense word, is still recognized as unfamiliar. For an object to be recognized as unfamiliar or meaningless, a phase of object meaning has to be traversed for there to be a mismatch with another class that is recognizable. The chair, the observer (self), and the space between the chair and the observer, including the sequence of buried phases leading to these (inner and outer) entities, are moments in the derivation of one object.

The microgeny of an object is unidirectional, with successive forms of the object realized in an obligatory sequence. Successive phases are not slices of the object though each phase, with its infrastructure, has the potential to deposit as a different world. An observed object in nature is the full set of phases over which it actualizes. The object does not exist for an observer without all of these phases, so the set of phases (phase spaces[10]) of the object is its minimal duration. An imaginary slice through becoming is an artificial segment. The entire process is necessary for a veridical object. The process *is* the object, i.e., its momentary time-creating set of developing spaces.

An object is an entity in space-time. An unexpected change in spatiotemporal context is an occasion for a new object. The temporal context is the deviance across becomings. The deviance is anchored to a "decay" point from which a duration is computed. The spatial context individuates an object from others in the vicinity. Adjacency distinguishes a tree as a unity; duration incorporates the unity in the manifold of the field. The meanings that arise in these relations guide the actualization process. Meanings in the early object experience reappear as submerged phases in maturity. The final object is what remains after its relations are specified. A tree is a figure nested in context. Essentially, an object is a set of contrasts.

Every object requires a duration sufficient for its existence. The initial phases of an object begin as potential during the present of an actual object. These phases reconfigure what will momentarily become the past of the object toward which they are heading. An object grows out of the past as a preparation for the next wave of object formation.

The immediate past of an object, i.e., the "track" that guides its development, does not include prior *actual* objects but the sequence leading to them. For example, in a series of final objects going from A to B, the past of B is not A but the process leading to A. The immediate past of an object, say a chair, is not the present (actual) chair less its change in the interval.

Mead has written,[11] "The organization of any individual thing carries with it the relation of this thing to processes that occurred before this organization set in. In this sense the past of that thing is "given" in the passing present of the thing, and our histories of things are elaborations of what is implicit in this situation. This "given" in passage is there and is the starting point for a cognitive structure of a past."

Direction and Change

In the elaboration of mind as in the course of life, change is in one direction. There are forward and regressive tendencies in matter[12] but change is always asymmetric. Involution is not the reciprocal of growth; growth and decay are different ways of characterizing the outcome of change. The outcome depends on the form that is changing, not the change the form is undergoing. The direction of change differs from the direction in which an object changes. For example, dying is an aging in growth, not a decline as a reversal of change. Involution is a judgment as to direction by an observer who interprets change as an external effect on a solid object.

Could change and time be symmetric considered independent of what is changing? Symmetric time depends on the distinction of time from objects in time, not from change, since any reversal of change would entail a reversal of time. The opposite of asymmetric change is time reversal. Since a lack of change is an absence of time, the opposite of change in any direction is timelessness. A thing without change is a nonentity, or not in time. An entity must change to exist. Objects perish when they cease to change, so change prevents an object from disappearing. Stasis is minimal change in the appearance of an object; the actual form of the object is similar from one becoming to the next. A rapid change is a dynamic shift in form. The apparent rate of change differs according to the deviance in shift in an actual object, i.e., the degree to which a configuration is self-replicating.

Rate of change is relative to a reference. How can one determine whether a rate differs between two objects? To extract a rate of change in one object, two successive points in the same object have to be compared. This requires an interval. Two events over the interval compared

to other events over the same interval might reveal whether the number of events per interval is rapid or slow. However, a rate requires a duration to measure events per unit time. Duration is mind-dependent, and so, therefore, is rate.

The form of an object is the configuration deposited by change. A shift in form, or rate of change, is a shift in configural properties. When snow melts to water, the rate of change is measured by the configuration (snow) that is changing (to water). The configuration is altered by an increase in temperature. This is an intrinsic change in the object; for example, an increase in the kinetic energy of its constituent elements. Whether snow melts slowly or rapidly, every moment of the object is a form that departs to some extent from its model in the immediate past. The object—snow, slush, water—is what it is at that moment. At one moment the object has the configuration of snow, at another, of water. The rate of change is determined by a comparison of two times (T1, T2) and two objects (snow 1, snow 2), not an observation of one object that is changing. If snow persists, it appears to be a solid object. The configuration of the snow is relatively constant every moment it is reproduced. The constancy is the coherence in the track over which the object develops. The greater the coherence, the more likely the object will persist (recur). An object persists when the outcome of change is a new object of minimal (but some) difference from that of a moment ago.

CAUSATION

Causation is related to time only if change is related to time and only if all change is causal. If change is not uniformly causal, if there is causal and noncausal change, then either causation does not inhere in change, or change is not related to time. Independence of causation and change implies a Laplacian universe where everything is determined[13] and nothing changes, while independence of time and change entails a time before the universe began and a time that would continue were the universe to stop.

It is difficult to imagine time without change. This is not because the existence of time makes change possible, or because there must be change to "fill" empty time, but because the perception of change and the concept of time are deeply interwoven. This is not true of causation, which is a theory of change that seems to be independent of time, even if causation involves temporal concepts that assume a (discrete) theory. These concepts are embedded in the idea of causation. For example, a *cause* is in the present, an *effect* is in the future. These concepts, including those of precedence and reproducibility, are implicit in causal

theory. While causal relations can be discussed independent of time, the concept of causation issues from that of time asymmetry. The idea of causation, therefore, is due to the concept of time, not the reverse.

Precedence

The concept of precedence is fundamental to causation. In fact, causation is precedence in a special type of connectedness. But precedence cannot be asserted in the absence of a theory of time. A before-after relation is a weak type of precedence—actually a mode of succession—that assumes time but does not give it. Before and after do not determine a past and a future. Time in the world might be circular or recurrent, e.g., $A \rightarrow B \rightarrow C \rightarrow A \rightarrow B \rightarrow C$. Event B could both precede and follow event C with no distinction of multiple recurrences of any event. To have a past that is not a possible future, i.e., a past that is forever past, one needs a present to fix that past at a given temporal "distance." Past and present are codependent, since the present emerges out of the past, and the past is implicitly dated and revived in every present. There can only be a past if there is a present the past is prior to. If the present is mind-dependent, so is the past and so too is precedence, which is a relation with (at least) one limb fixed in the past.

Change is a transition of events, precedence is the sequence in which this transition is located. As change is not independent of the objects that are changing (i.e., something has to change for change to occur), precedence cannot occur without an event pair. Precedence requires change between events (and times) that are dissimilar. Precedence obtains when $A \rightarrow B$, not $A \rightarrow A$. A dissimilarity between events is the basis for a determination of the sequence. Two events at the same time are simultaneous, while two identical events are the same event. The dissimilarity is the product of a mind-dependent comparison in a present that incorporates the events that are compared, even if the events, e.g., Columbus leaving Spain and arriving in the New World, are all in the historic past. Since past and present are part of the same mental construct, the past does not exist independent of mind unless there is a "God's eye" present—thus, a past in God's mind—everywhere at once in a block universe.

The physical "tick-*tick*" of a clock is succession, the mental "tick-*tock*" is precedence. Pure succession is not possible in mind, for precedence is imposed on successive events. Events are identified and prioritized. The identification is the conceptual selection of the event and the assignment (extraction) of meaning. The prioritization is a personal valuation and the position of the event in a sequence. The establishment of precedence is an essential activity of mind and is derived

from the event-layering of memory that is responsible for the phenomenal present. To know, at B, that A precedes B, not just that B is altered for having been preceded by A, in which case A need no longer exist, is to be conscious of a duration that extends beyond (before) B; i.e., to have a present in which A survives *as an event* in B.

Precedence, therefore, is mind-dependent but not succession. Precedence is order, succession is nonsimultaneity. Events that are not simultaneous are successive. Precedence is the sequence unstated in succession. It is a strong form of succession where events are discriminated in a temporal relation. Causality requires (minimally) the determination of cause and effect or the demonstration that causes have effects or that effects have causes. With the exception of "backward causation," i.e., that present events are caused by future ones, or the improbable idea of a future that attracts the present, i.e., that the present does not determine the future but is "sucked" into it, causation requires a direction, at least for a given causal transition. The direction is assigned in the relation of precedence between cause and effect. The cause is prior to the effect and distinguishable from it. Since precedence is mind-dependent and causation depends on precedence, causation is also mind-dependent.

Probabilistic relations are potentially symmetrical or reversible. Is causation asymmetric? In a causal world, every cause is an effect of some prior cause. In such a world, the distinction of cause and effect is arbitrary and depends on where the causal chain is probed. Since causation depends on relations of precedence that are mind-dependent, causal asymmetry is also mind-dependent.[14] The asymmetry is related to the feeling of agency. But why should agent causation be asymmetric unless the mind is constrained in evolution by an asymmetric nature?

Discrete Time

The concept of causation may be incompatible with the hypothesis of authentic change. Causation requires solid objects and discrete time. If objects actualize out of process, if there are no physical solids, a change does not act on an object but melts into another change. It is easy to imagine an object changing to another object, but what is a change becoming another change? At what point does change become something other than what it is? Objects and changes (or events) require demarcation to specify causes and effects. A change cannot terminate and another change begin, for there is no way of characterizing what happens between changes. This (false) problem of change between events has occupied the minds of many fine thinkers, and is not resolved

by the nesting of causal links within chains, e.g., A → A1 → A2 → B since the difficulty is the transition between minimal pairs.

Causation requires change that is extrinsic to elements. But if objects are constantly changing, what is the change that is the object and what is the change that is the causal effect of the object? Either the epoch of change is contained in the preceding one, in which case there is no demarcation between the object and the change the object undergoes, or the change is introduced between epochs, or between one object and the next. In any event, causality requires demarcation. In a continuum, change is a four-dimensional becoming. An interface, thus a cause, cannot be isolated in the stream of change, only a segment that precedes or follows another segment. This leads to the conclusion that cause and effect are *neighboring* segments of change, a segment being the formativeness of an object (see below) where the preliminary structure of the preceding object (cause) dominates the formativeness of the ensuing one (effect).

A cause should be more, however, than an adjacency of formative sequences. An object could then cause itself to reappear. Each instance or realization of the object would be the effect of a becoming over the infrastructure of a prior (causative) instance. The idea of causation seems to involve the introduction of energy to a system with a change in state or position or the coming together of disparate objects. Still, the precedence of one of a pair of becomings with some regularity is a first approximation to a causal relation. Causation, then, is a type of anticipation with a high degree of probability, leaving aside for the moment the nature of the objects involved in the causal action.

Reproducibility

The concept of causation tends to disregard an object that is self-replicating but it thrives on reproducibility across encounters of a similar type. One cannot establish that A causes B without multiple instances of A followed by B. The need to verify the causal effect, that a match lights when struck, requires that some (causal) change can be repeated. In novel change there are no exact recurrences. One can strike many similar matches in a similar way and determine they will all light, but one cannot strike the same match twice, or identical matches in an identical way, and determine they will always light. Suppose one could repeat the identical sequence and establish over innumerable trials that the outcome is always the same. Could one then say with absolute certainty the next action will not be different or is there some possibility, however remote, of an unforeseen happening?[15]

A billiard ball that does not move is not caused to remain where it is in the same sense that the cue causes it to change position. An unlit match is not the cause of the "same" unlit match a moment later in the same sense as striking the match causes it to light. The match or the motionless billiard ball is moving with the earth and is a density of active particles, but something other than the structure of a ball or match is introduced to cause a change in these objects. Russell[16] argued that persistence is a form of (intrinsic) causation no less than interaction. For most philosophers, however, the idea of causation usually involves an extrinsic effect. This way of thinking is strengthened by the perception of a collision of one object with another, or the impact on an object of an "external" force.[17] The appearance of an interaction between different objects, combined with the relative suddenness of a change, enhances the impression of a causal relation and, conversely, weakens the impression of causation in objects that are changing more slowly. We do not ordinarily consider a child the cause of the adult, or each moment in the life of any object the *cause* of the next.

The inability to reproduce exactly the same cause and effect and, were it possible, the inability to claim that the same conditions would always lead to the same outcome, along with the influence of quantum theory, have led many thinkers to speculate that causal laws are probabalistic. For some theorists, causation is a relation of probabilities. For others, the probabilities are surface indicators of underlying deterministic relations. A probability close to zero or one is a best approximation to causal certainty.[18] As Emerson wrote,[19] "The dice of God are always loaded."

If causation is probabilistic, causes do not necessitate all of their effects.[20] If causes do not necessitate all of their effects, can it be said that they necessitate any of them? If every effect is a probability, is any effect a necessity? Presumably, causes increase the probabilities of their effects but this increase may not be the result of a causal process. Put differently, the probabilities that reflect a causal relation are descriptive of an association between cause and effect, or the liklihood an effect will follow a cause, but are not accounts of the causal relation itself. In this sense, probabilities are not laws but measures of regularities in nature.[21] In fact, a probability, even zero or one, is not a law even if built into all change in the world. The probability would indicate that many (or all) A-like events are regularly (or always) followed by B-like events, so an event in the A category is likely (or certain) to be followed by an event in the B category but it could not give the change, e.g., whether causal, emergent, and so on, that characterizes the A to B transition.

Every object is a complex of momentary instantiations. Suppose we study the series of instants or slices of an object, say a match as it is

lit and burns, or a billiard ball that is displaced. We would observe a set of microstates in gradual and coherent transition from one configuration to the next, whether from the perspective of a change in state or a change in position. The momentum or energy transferred to an object transforms it to another object, but an object is an energy state to begin with.

To identify some change as causal—the change from unlit to lit match—and other change as noncausal—the persistence of an unlit match—is to deal with change in an arbitrary manner. It is also a displacement of change from the interior of the object to its external surface. In a word, it focuses on the apparent change between objects rather than the intrinsic change inside them.

The problem of causal change, then, is that it posits interaction in the silent gaps between final objects. Since these gaps are changeless, causation is a folk account of the changing appearances of things, not the authentic change through which object replacement occurs. Moreover, if change is not universally causal, causation is not a universal theory of change. To isolate A- and B-like events in a causal relation is to claim that some events are causally related while others are not. Swallows are not causally related to summertime. A tree hit by lightning falls as I shiver in the rain but not because of my shivering. A theory of change that is a partial account of all change is probably the wrong theory. The problem for such a theory is not the local occasions of swallows and summers, or lightning and trees, but the passage of one slice of nature to the next and whether this passage is everywhere causal.

Causation and Becoming

A becoming is a microgeny or an actualization on the way to completed form. Microgeny refers to the process of becoming in relation to phylo-ontogeny. Becoming is an actualization in relation to time and change. Actualization is the realization of actual (present) objects. These are all concepts in which the process of change is within the developing object, i.e., the local density in mental space and time.

As mentioned, if causation exists, it does not characterize the sequence of appearances of mature objects but the becoming through which the appearances develop. Object B develops out of the microgenetic path of object A, not the transition of A to B. The process of realization leading to an object, not the actual object, is the basis for the ensuing object. Imagine a fountain with an observer looking down from above who sees only the configuration of the spray at the surface. Each new configuration seems to change into the next. The observer, who cannot see the upward-moving stream, will conclude that the shape

of the spray at A leads to the shape of the spray at B when, in fact, A is replaced by B and it is the actualization of A that determines, in part, the shape that B will assume.

So in the growth of a mental object, A leaves a track that is the basis on which B develops. B is guided by this track and the constraints of sensation. Causation in such a system is the relation of tracks A to B, and the alteration in track B induced by the external world, not the final or surface objects these tracks generate. In other words, the actualization process, not the object that actualizes, is the "track" or vector that guides the growth of subsequent objects. This track accounts for the relative constancy of passage from one object to another—the persistence and slow change of the world—and is the basis for the appearance of a causal linkage between objects.

Epistemological Dualism

Epistemological dualism[22] refers to the inexact correspondence between percepts and real objects, psychophysical dualism to the duality of the mental and the physical. The former is a mind–object distinction, the latter a mind–brain distinction. Since mental states are more immediately related to brains than to other objects, epistemological dualism assumes a causal series from the real object through brain process to the mental state.

Causation is deeply committed to a theory of external objects. It requires a mode of realism that excludes the mental or entails a duality of material and mental objects. Objects in perception are, proximately, configurations in the brain. A causal linkage is assumed between the "real" object and its brain state correlate—the "physiological object"—which is taken to be the equivalent of the real one. Whether the mental object is identical to the physiological one, or caused by it, is secondary to the causal inevitability postulated in the world-to-brain linkage. Since mental objects are assumed to be direct effects of material objects, mediated by the brain state, the assumption of a causal, but not existential, subjectivism is necessary if a complete subjectivism is to be avoided. The retreat to causal subjectivity seems, given our present knowledge of the brain, the last refuge of the realist. Causation requires solid objects to impact on the mind. But the concept of solid objects independent of mind is implicit in a theory of causation. Causation, therefore, is circular in that it requires as a given (interaction between objects) what is imputed as a proof.

Novelty

The idea that objects do not endure but are continuously renewed—the renewal being the process of change—is the basis for the inference that

novelty is the essential characteristic of change. The first step is the disavowal of objects as solid entities that occupy the present, for a historical process in which objects are generated from past to present. The object is not the outcome of a prior chain of events. The past does not "do its work" and disappear, but is revived in every actualization. On this view, the past of an object is not the object of a moment ago but a phase in the formation of the present. The past of an object is engaged in the becoming of the next object, even as the present object disappears.

There is novelty in the birth and growth of objects and their unfolding from depth (past) to surface (now). Since an object does not cause something to happen, a discussion of novelty cannot be framed in the context of entities that change, or whether there is novel change in an entity, but whether the change that leads to the entity is novel. This change is the thread of the world while the entity in perception is its residue.

In principle, change could be novel, emergent, causal, or recurrent. Change could lead to the same object or a different object, and the different object could be the result of a fixed or open succession. To say that some changes are novel presumes that other changes are not. I would think that genuine novelty is possible only if it is true of *all* change. If this is the case, novel change does not occur only in complex systems but at the most basic level. If novelty were limited to complex systems, causation might characterize the component elements. Emergent states would be resultants of causal relations that were obscured by the complexity. Novelty as a function of complexity is related to the (im)probability of an exact recurrence. The more complex the system, the greater the novelty and the more likely a departure from sameness or predictability.

Novelty is not randomness since there are strong constraints on the succession of objects; i.e., every mental state is biased by the preceding configuration with a high probability of replication.[23]

Moreover, the possibility that an object might recur on the basis of chance would imply that novelty is not a universal feature of change. Novelty should be found at every level of organization. There is novelty in the transition across two successive states of an electron. An electron is part of a system. The novelty is in the "adventures" of the particle in relation to the system and in relation to larger domains of organization. Novelty is the changing temporal and spatial relations of a particle, and the impossibility of replicating exactly (in context) the inner and outer relations that even the smallest particle enjoys. Context is the key.

Unlike causation, novelty excludes repetition, but causation might not exclude novelty if what is novel is determined by the prior state. Universal causation entails that events can be predicted from prior

events, but the prediction of novelty does not exclude novelty. The combination of two nuclei of hydrogen and one of oxygen to a molecule of water can be construed as a causal interaction with a novel effect. The effect may or may not be predicted by a knowledge of the constituents, but if the effect differs from the sum of the activity of the constituents and if the constituents change in the transition to the effect, there may or may not be novelty in the outcome of the change, but this would be independent of its predictability. Causation can still be invoked if the outcome of a novel change is predictable, even if the novel or emergent step involves a relation that is not fully explicable.

Causation, Novelty, and Free Will

Free will presumes that consciousness and choice intercede in the causal flow of events; that mind can penetrate a causal chain and effect a novel outcome. Free will requires physical causation in two ways: for conscious intervention in a causal chain of events; and for causal interaction between consciousness and the effect of that intervention. In the first instance, causation is that mode of change to which consciousness is opposed; in the second, causation is invoked in order for a change to occur. In that mind causes something to happen, free will enjoins causation but substitutes the self or consciousness as a prior cause (see Chapter 6).

If there were novelty or indeterminacy in the material world, free will could be an influence on or attunement of the probability of a given outcome. Free will could *determine* what was uncertain in the world through an intrusion of mental (agent) causation into physical indeterminacy. If physical change were novel, there would be little to choose between novel change in the physical world and novel change in conscious behavior other than that the complexity of the latter would give the impression that an agent was steering the novelty in one direction or another. Indeed, free will might entail the intuition or anticipation of novel outcomes achieved through determinate change.

If conscious choice were determined by prior events, i.e., if the mental determinants of a choice were themselves determined, or if given a choice the decision to act (the act that is chosen) was a necessary conclusion of a causal chain in mind, consciousness would be interposed as a node in a causal sequence. In the decision to blow out a match and the exhalation that follows, consciousness would play the same role as a sudden gust of wind that extinguishes a spontaneous fire. The difference is the decision to intervene and whether this decision obeys causal "laws." Free will depends on a strict interpretation of causation and a loose interpretation of choice. Free will requires the

possibility of choice irrespective of whether the bases of the choices are causally determined. There is no choice in nature. This is why, finally, free will resolves to the analysis of choice and the autonomy of the self that chooses.

In brief, we act as if an intended action causes the effects we desire. Whether the decision is based on a projected outcome or the probability of the outcome, given a set of present and past known facts, i.e. an anticipation, the problem is not what goes into the decision in terms of the facts that are considered but how the decision is made. In other words, free will does not depend on the (rational) basis on which the facts are decided, or how the individual sifts through (deliberates) various alternatives, but the final precipitation of choice by or through the self.

NOTES

1. A. Lovejoy, *The Revolt Against Dualism* (La Salle, IL: Open Court, 1930).
2. The fascinating topic of inverted vision is discussed by M. Solms, K. Kaplan-Solms, M. Saling, and P. Miller. "Inverted Vision After Frontal Lobe Disease," *Cortex* 24(1988): 499–509.
3. See E. Pöppel, *Mindworks: Time and Conscious Experience*, English trans. T. Antin (Harcourt Brace Jovanovich, 1988).
4. See *Self and Process*, 127–146, for a microgenetic theory of duration and the nature of the specious (phenomenal) present.
5. A. N. Whitehead. *An Enquiry Concerning the Principles of Natural Knowledge* (Cambridge, U.K.: Cambridge University Press, 1919).
6. There is a good discussion of this point in J. Lango, *Whitehead's Ontology* (Albany, NY: SUNY Press, 1972).
7. *Self and Process*, 127–46.
8. W. Newton-Smith, *The Structure of Time* (London: Routledge and Kegan Paul, 1980), 17 et seq.
9. The distinction between objects as solid entities that undergo change, and change as an atomic sequence that lays down entities, is captured by some quantum theorists in the distinction of particle and wave: "We are *directly* aware of the particle aspect of the universe through the senses (while) the more subtle wave function is *inferred* by thought about our sensory experience in the domain that is manifest to the senses" D. Bohm and B. Hiley, *The Undivided Universe* (London: Routledge, 1993).
10. See K. Pribram, "Prolegomenon for a Holonomic Brain Theory," in *Synergetics of Cognition*, ed. H. Haken and M. Stadler. *Synergetics*, vol 45 (Berlin: Springer-Verlag, 1990).
11. G. H. Mead *The Philosophy of the Present* (La Lalle, IL: Open Court, 1932).
12. A. N. Whitehead, *The Function of Reason* (Princeton, NJ: Princeton University Press, 1929).
13. Unlike some philosophers, such as P. Van Inwagen, *An Essay on Free Will* (Oxford: Oxford University Press, 1983), I do not find the distinction between determinism and universal causation persuasive.
14. For example, H. Price, "Agency and Causal Symmetry," *Mind* 101(1992): 501–520.
15. The principle of alternate possibilities holds that an action is free if a person could have acted otherwise or refrained from acting; for example, S. Cahn, *Fate, Logic and*

Time (New Haven, CT: Yale University Press, 1967). However, if the same conditions can never be exactly repeated, how can the occurrence of an "equally weighted" alternate possibility ever be established?

16. B. Russell, *Human Knowledge* (New York: Simon & Schuster, 1948).
17. Acceleration is simultaneous with force, though, as Earman has noted, "these immediate effects 'spread out' in the future direction of time," J. Earman, "Causation: A Matter of Life and Death," *Journal of Philosophy* 73(1976): 5–25.
18. See D. Papineau, "Causal Asymmetry," *British Journal of Philosophy and Science* 36(1985): 273–89.
19. R. Emerson, (1841), "Compensation," in *Essays*, 57.
20. W. Salmon, "Probabilistic Causality," in *Causation*, ed E. Sosa and M. Tooley (Oxford: Oxford University Press, 1993).
21. See B. van Fraassen, *Laws and Symmetry* (Oxford: Oxford University Press, 1989).
22. Lovejoy, *Revolt Against Dualism.*
23. The argument that freedom is the interplay "between something almost random or haphazard and something like a restrictive or selective control," K. Popper, *Objective Knowledge* (Oxford: Oxford University Press, 1972), restates this bias without addressing freedom.

CHAPTER 3

Asymmetry of Past and Future

ARGUMENT: *The past constrains the present, and is revived in it as a "track" in its development. The present departs from reproduction through novel change in the revival and constraints on becoming imposed by the external world. The future is the set of possible presents that might be reasserted in the perishing of the current present. Belief in the future is impelled by agency, regularities in becoming, and the asymmetry of microgenetic process.*

PROBLEMS CONCERNING THE NATURE OF THE PAST

What is a past event and how are we to conceive of time past? What does it mean for a past event to have once existed or to exist "for all time"? We assume the existence of events in an *objective* past independent of the present in which the (past) existence of the event is represented. For example, we take as a given the objective occurrence of the billions of years of the universe prior to human awareness. Another mode of pastness is the representation of a past event *in the present*, e.g., our (current) understanding of the early history of the universe. In addition, there is the personal or *experiential past* of an observer.

The objective past owes its characterization to a body of (scientific) knowledge. This knowledge is not a true representation of the objective past but an interpretation of what the past was like. The interpretation

is active or recurrent in the present even if the events inferred to have occurred are outside the causal history of this planet. Since knowledge or an account of knowing entails a personal engagement, the objective past is always filtered through the experience of an observer. A consensus as to the nature of the past reinforces the observer's belief in its objectivity but does not add objectivity to the past event. The past of personal experience, therefore, includes not only the events (episodes) of a life, their recollection and documentation, or the part they play in the derivation of the present, but the representation within a present of all the conceptual knowledge that is available to a given perspective. Whether or not there is an objective past, the past, objective or not, is ultimately known in the phenomenal present of an individual mind.

Put differently, what we know of the past is "factual" knowledge accessed in the present moment. The objective past is a story based on this knowledge. Indeed, the nature of the past changes as our knowledge of probable events changes; the story is constantly being revised. This is not to say there is no objective past, or that the past did not exist, or does not in some sense continue to exist, but rather that the basis for our knowledge of the past is through present experience. This is not a trivial observation.

IMMUTABILITY

The problem of objectivity is linked to the question of permanence. In causal theory, a past event is the cause of another (ensuing) event. The distinction of cause and effect is a necessary condition for the postulation of immutability. If change occurs *between* objects, the objects themselves can be conceived as events that, once having occurred, persist without change. Since change in causal theory is between objects or events, objects are not necessarily changed in the passage from one state to the next. The postulation of extrinsic change is consistent with the idea that objects are immutable but does not necessitate that an unchanged object persists for all time. The concept of an immutable past is not obligated by causal theory. Immutability, however, is in accord with determinism or fatalism, in the idea of a chain of past causes and effects (objects, events, instants) that is forever unalterable.

Consider a past event, such as (a moment in) the life of a dinosaur. We know that dinosaurs once existed, but in what sense did (do) they exist? Suppose we see an image of the light record of a dinosaur streaming to a distant galaxy like an image of a far off exploding star. Such images remind us, like photographs, that the events actually occurred. But such a recording, like a movie of one's childhood, is perceived in

a perspective. This perspective is the present (now) of an observer, not the actual past of the event. The ability to record the past does not settle the question of whether or not a past object is immutable, even less whether it is an "immortal" fact or an "eternal" truth. We have to ask what we mean by fact, by immortal or eternal. What does it mean for a thing to exist at any moment, past or present?

The atoms that constituted a dinosaur persist in the bones of the dinosaur, and elsewhere, as objects in the present. The bones of the dinosaur allow us to reconstruct a story about a once living organism. The atoms of the dinosaur are immortal. They are redistributed but never die, but is the *fact* of the dinosaur immortal? This involves, in part, a theory as to the irreducibility of the base constituents of a thing and the varied configurations such constituents can take on.

The Facts of Existence

We cannot approach the fixity or immortality of the past through its documentation in perception or memory. These are transient events in the present of an observer, really present objects or concepts and only secondarily past events. A present object raises another set of questions regarding existence. I think we mean by existence, past or present, that an object or event is an irreversible or unalterable *fact*.[1]

Moore[2] distinguished the various meanings of "to exist" such as is *real*, is *true*, or is a *fact*. These distinctions are important. A unicorn is a fact in so far as it exists as an idea or a pictoral representation. A dinosaur is a different kind of fact than a unicorn, but what is the difference? Neither dinosaurs nor unicorns exist. The fact that we know that dinosaurs once existed adds belief or knowledge (but not existence) to the existence they presumably once had. Dinosaurs and unicorns are ideas in the present. What is the difference between the ideas of two objects, one of which was once "real"? We could say that dinosaurs and unicorns are concepts but only the former were objects. Does a concept exist in the same way as an object? To become an object in perception, a concept usually needs a material entity that serves as a provocation or model.

Both objects and concepts have properties. Does a thing have to exist to have properties? A unicorn has properties. It exists more than an event in the future. At least we know what a unicorn looks like. The idea of a unicorn and the idea of a dinosaur might be on an equal footing, but we would say a unicorn has less existence than a dinosaur or that it exists in a different way, e.g., as an idea rather than a fact. Dinosaurs and unicorns are ideas but dinosaurs are facts. What is a fact?

We cannot say that to exist is to be a fact if to be a fact means to exist. A fact is a (momentary) configuration of enduring atomic elements. What, then, is the meaning of existence for the past configuration of a dinosaur? Is there not a more concrete sense in which the expression "dinosaurs exist(ed)" is meaningful?

Dinosaur bones appear in perception but a dinosaur is a theory on where the bones came from. The bones are objects in the now of an observer. The dinosaur is an idea in the now of a thinker. An idea has a quality of existence. Moore argued that the idea of a griffen exists. There is such an event as that idea. To exist in the mind as an idea and to exist in the world as an object could be two different modes of existence. Or, one could say, these are qualitatively different contents (concepts, objects) in the mind, an idea (dinosaur) and an object (bones). What then is the connection between an idea (or concept) and an object (or fact) in relation to existence? When a concept becomes an object, does it also become a fact? If concepts are also facts, a unicorn is a fact. If only an object is a fact, what becomes of facts if an object is the realization of a concept? How does a non-factual concept become an objective fact?

The Revival of the Past in the Present

The problem of the past is not resolved by a distinction of idea and object: whether objects are derivations of ideas and exist like ideas; whether objects exist independent of minds; or even whether mental events can be said to have a "substantive" existence. Nor does it depend on an accurate picture of dinosaurs or if there "really were" such animals. The problem is in what sense the past can be said to exist or remain a fact that is forever unchanged.

If I type an x on my word processor and erase it and replace it with another x, what remains of the existence of the first x? An x in a word processor, like a footprint washed away on a sandy beach, may be a trivial fact in relation to a dinosaur, and unlike a dinosaur may have no consequences and leave no traces, but like a dinosaur an x is a fact in the world. More accurately, the world was a fact that contained the x or the dinosaur. Suppose I only think about typing an x and then forget I even had this thought. The thought was once a physiological event in my brain. Does the past world contain its past ideas? We are searching for a sense in which the past x, the past dinosaur or the past world—objects or ideas—exist as past facts.

The reason for these speculations is to gain a better understanding of the nature of the subjective past. From the standpoint of microgenetic

theory, a moment in the life (or concept) of a dinosaur *replaces*, or is a transformation of, the preceding moment. If every moment perishes in its replacement, what becomes of the preceding moment when it is replaced? How can a moment in the life of a dinosaur continue to exist as an irrefutable state of affairs if that moment—more precisely, its actualization—was a path for a reinstantiation of the next moment? A theory of change in which objects perish and are replaced by other objects has a decisive impact on the way the past of an object is interpreted. On a replacement model, a past object is erased by the one that follows, since the latter grows out of the one that is being replaced. An object (A) persists as a constraint on an ensuing object (B). The configuration of B is that of a reconfigured A. When B perishes and is replaced by object (C), there is no actual sense in which A can be said to persist in C. What survives of A in C is its reconfiguration through B into C, but what is actually in C is B reconfigured.

Since an object is a fact that has being or existence in a way that differs from ideas, and since every object emerges from the past of a previous object, i.e., from the becoming of that object, history is not propagated by facts or objects but by a covert fact-generating process in relation to which facts and objects are surface manifestations. The past reappears in the body of the present and dies in the final shape the present takes on. Every past moment is transformed as it propels an object into the present. The past that is revived, e.g., the concept of a dinosaur, is delivered into the present as an actual fact, e.g., that dinosaurs exist(ed). However, this still leaves suspect the fact of the past dinosaur. The problem is not a break in the causal history of the subject at birth, or the onset of a personal subjectivity, but the *immutability* of any facts antecedent to present experience.

Memory and Pastness[3]

The past experience of a subject, whether recollected or long forgotten, contributes to every present. The present is guided toward actuality by the history of all prior presents, i.e. the personal past which, cumulatively, deposits the preceding present. The intrinsic past is not a memory of the past but the shaping effect of the preceding moment. The personal history exists only as a constraint on the occurrent present. Put differently, the present moment—its becoming, not the actual world that is deposited—is all that exists of my past in relation to the present of the next moment. Memory is the process through which past objects become actual, i.e., the becoming of objects into (as) the present. It is not a mechanism or set of operations by which the past is looked up

and remembered. The different expressions of memory are phases in the actualization of a present object. In this way, memory deposits a living organism.

The past is revived (re-*presented*) as an image or an idea. Images and ideas are present objects that are incomplete, i.e., have not exteriorized. The present is created by the revival of the past to an actual object. An actual world is necessary for a present moment. When the past achieves an objectlike status, as in dream or "veridical" hallucination, an image can become a present object. But this is a different present than that of an actual (real) object. When an actual object fails to be realized, a world of mental imagery becomes the entire world of the observer. The observer too is incompletely specified. However, when an actual object (world) does develop, incomplete objects (images) embedded in the actual ones constitute the mental (introspective) content. Mental content must be in opposition to a real world or it will become a "real" world itself. This content, e.g., a memory image, an idea, is intermediate between past and present, with one foot in the past (in personal memory or world knowledge) and one in the present (the now of the observer).

The study of intermediate phases confirms that past and present are segments of becoming, not receptacles, and that past and present are established in the process of object specification. The closer the becoming to an image, an idea, a proposition, then finally an object or an utterance, the more presentness it takes on. Conversely, a premature termination or an enhancement at a preliminary phase gives a content in the past. More precisely, pastness is evoked by such contents through an attenuation of their actualization. Past and present are created in the becoming. A content that is submerged in the present is apprehended as a memory. A memory that actualizes in the world is apprehended as a fact.

Time, Space and Memory

The world is experienced in the space of mental objects and interpreted in the time of mental process. Time is ingredient in this experience, but covertly. For an object to persist in spite of change or to undergo change without changing, i.e., to continue to be the "same" object, change has to be stabilized over some duration. The persistence of an unchanging object is not a persistence over time of a solid object but is the compilation of phases within the object as it makes its successive appearances, i.e., in the coherence of its formative moments.

Space and time are layered. There are multiple times and spaces. This is evident in studies of altered states and cases of pathology.[4] A

phase in mentation may actualize prematurely in a different space and time. The space of dream or dreamlike hallucination is foreshortened, volumetric, and viewer centered. The dream is (almost) timeless, or at least not serialized as in waking perception. In waking hallucination there is time alteration. Events may be static (as in palinopsia), discontinuous, slowed, or accelerated.[5] These preliminary segments are normally traversed to persist as virtual layers in the "clock time" and Euclidean space of a conventional object. Their reappearance in pathology or dream supports the inference that multiple times, and spaces are implicit in everyday objects. Every space and every time is a different world. For example, dream and wakefulness are contrasting modes of cognition with different times and spaces, apparent times and apparent spaces that, like apparent objects, range between possibility and fact.

The asymmetry between past and future is brought out in cases of pathology. Patients with a memory disorder have an altered sense of time (duration).[6] In severe cases, the feeling of a continuity of past to present is lost, the subject having an experience that is limited to the actual present, i.e., present objects. Such cases may still retain a lively sense of the future and its possibilities. Other patients, e.g., with frontal lobe lesions, are drawn to stimuli in the immediate surround. They may have a good memory yet tend to be disinterested in, but do not lose, the past and the (idea of the) future, and seem to "live for the day." In such cases, time awareness and duration judgments are probably unaltered.

Amnestic cases suggest that the ability to recall events punctuates a duration and prolongs it. A duration in which events are forgotten is an epoch of relatively changeless, therefore nonexistent (contracted) time. This is not true for normal individuals where "filled" time passes quickly and "empty" or "unfilled" durations, e.g., periods of boredom, are perceived as longer. The difference pertains to a duration in recollection as opposed to immediate experience, i.e., duration judgments versus the feeling of time passing. There is also a difference between an inability to recall and an "eventless" period with normal recollection.

Past and Future

The nature of the past, or rather how the past is interpreted, establishes our view of time and history. If history is an accumulation of events *over time*, and the present is the leading edge of the process of historical change, the past is a foundation on which the present is overlaid. If history is a chain of events, like a movie reel, the present serves to transmit the sequence one step further. The facts of history are then eternal and the present could not have been otherwise. On the other

hand, if history is an excavation of the depths of the present, i.e., a search for the (lost) series of configurations that survive abstractly in the present moment, the past is approximated by a retreat through this depth, which corresponds to pastness. The deeper a content, the more glacial the change in its reconfiguration. The surface of mind consists of actual objects that are constantly "changing." The depth consists of a past that changes slowly, i.e., is a *memory*. That is why a withdrawal from surface to depth is a way of recapturing the archaic sources of the present.

The way we conceptualize the past is important in that many theorists, taking a causal chainlike view of history, see a symmetry between past and future. The "once was" exists in the same way as the "will be," with the present moving along a line between them. What is the basis for this idea? Does the future, like the past, have a shaping effect on the deep structure of the present? Some philosophers have speculated that foretelling the future is like remembering the past. But memory has a structure and a direction. The foretelling of the future has the quality of a memory, like a vague image in the "mind's eye" for distant (past or future) events, but prediction does not have the structure of recall. For example, what would be the equivalent for the future of long and short term memory, or the microprocess of "retrieval" and forgetting? Unlike perception, memory is consciousness of something other than facts. The contents of memory approximate but are never true to the "facts" of experience which are the illusory vehicles of change.

A memory is an abbreviated phase in the becoming of the present that encloses a meaning and a valuation rooted not in facts or events but in the immediate prehistory of the events that are remembered. If foretelling the future shares some of the properties of memory, it is not because prediction is memory reversed in time, but because all mental content, including ideas about the past and the future, is generated out of a common core, through phases having the collective character of what we call memory.

Thus, for the future to develop toward the present in a time reversal, the future would have to play the role of the past as a potential within the becoming of the present. For the future to play this role, i.e., to constrain the becoming of the present as a revival through layers of pastness, the future would have to take on all the properties of the past, and would then no longer be distinguishable from it. The past would also assume the properties of a (nonexistent) future. If this were the case, either there would be an inability to distinguish within the present whether the present was generated out of the past or the future, in which case the present would be nondirectional, thus incoherent, or the reversal would be an exact duplicate of the standard direction. If

the latter, why then is there a feeling of asymmetry from past to future rather than the reverse?

The concept of change is crucial to a theory of past and future. If change in the present is the outcome of a fixed past, i.e., if change is deterministic, how would novelty enter a sequence that is fixed to a given point? If the immediate past is fixed, and the present is inevitable, how does change move that state to the present, i.e., where does the newness of the present come from?

From the standpoint of the present, which is, after all, the only standpoint we have, the present gains novelty in a transition to the future. Objects are constantly flirting with change, and this change is perceived as a movement to the next moment. Change is perceived as the coming into being of the edge of a future penetrated by a now that is advancing. But if this edge is the surface of an oncoming now, with the transition from one present to the next eliminated because it is timeless (changeless), what future could be imagined other than the surface of the now that is coming into (creating) awareness?

The belief in the future is conditioned by plans and expectations, which are ideas in the present, and by probabilities, which are inferences based on regularities. Plans are by definition about the future since the past cannot be changed. A plan is a structure that contains a behavior as a potential or a concept that actualizes in parts. The belief in the future is occasioned by concepts as preparations for action. A concept leading to an action could be a logical plan or a bizarre fantasy. The content of the concept does not determine whether it is a fact. The content does not incline the concept toward or away from fact. It is the objectification or enactment that makes the concept factual. Since any percept or action is a present act, the plans and expectations on which a belief in the future is founded have the quality of (nonfactual) ideas. In this, we come full circle back to unicorns.

The idea of the future is based in part on regularities. These regularities, including the influence of the past on the present, arise in the coherence of becoming. A minimal deviation from one microgeny to the next, i.e., the relative similarity of perceptual content across a sequence of mental states, is the basis of personal identity and object persistence. A greater deviation in a relatively stable context is the basis of apparent change and "interaction." Duration establishes constancy, and constancy is the ground of prediction. However, the belief in a future, or the conviction in this belief, unlike the past which is based on memory, is not based on prediction—the future is unpredictable—but on the momentum and asymmetric direction of becoming.

In sum, the past is active in the becoming of the present to the extent that it fails to become a fact. The past is a potential that actu-

alizes in the present. The present is the surface of a becoming and consists of mental existents (e.g., propositions) and experiential facts (actions, objects) that survive after the past (potential) has been realized. The difference between a dinosaur and a unicorn resolves to the context or spatiotemporal relatedness in which these ideas occur. This context is their reality. The difference between the idea of the present and the idea of the future relates (1) to the content of the becoming; (2) to the feeling of agent control, intentions, concepts as plans, etc. (3) to the extraction of regularities, i.e., prediction; and (4) to the feeling of a forward movement to the present and the asymmetry of past-to-present, which extends to an asymmetry in the direction from present to future.

The Idea of the Future

The possibility of time travel, e.g., visiting the future before it happens, or the more problematic travel to the past, suggest to some that the future, as well as the past, is fixed and "waiting" to be discovered by a now that moves along a segment of a world line like a searchlight on a river of time.[7] For many, this is a defining metaphor, in Emily Dickinson's words:

> "Down Time's quaint stream
> Without an oar,
> We are enforced to sail. . ."

The immutability of (eternal) objects in the past, and the presumption that eternal objects "exist" in the future (otherwise, what is the meaning of eternal?), the hypothesis of time reversibility, and the concept of a moving now, prop up the idea of a future that already exists and comes into awareness as the now approaches its posterior limits.

For some philosophers, a fixed future is not the same as a future that is determined. Such a future might be unrealized and conditioned on the present. It would seem that on this view the present must occur in advance of the future for the fixation of future events, i.e., the future could not be fixed until all prior presents were realized. This is an open, not fixed, future that becomes fixed by the actualization of the immediate present.

However, it is unclear to me why a fixed future should differ from a future that is determined. A fixed future is required if time is reversible so the present can be recaptured by running time backwards. If change between presents is indeterminate, the indeterminacy would apply in

either direction and the present would not necessarily be encountered again by a time reversal beginning at some future point. This would also apply to the past, since indeterminate change confounds a reversal at any point in the time series.

If the direction of time corresponds with the direction of causation, a fixed future entails the possibility of a reversal of causation. In reversed or backward causation, an effect precedes its cause; for example, foreknowledge of certain danger in the future causing an avoidence of peril in the present. Though conceivable in principle, arguments for backward causation tend to rely on superstitious thinking[8] or the insertion of psychic phenomena into a description of physical causation, such as precognition and telepathy or phenomena such as waking before an alarm clock goes off.[9] There is also an appeal to "causal dependence" where a weak concept of causation is stretched to include instances of relatedness or contingency.[10]

Precognition would seem to be based, not on the effects of the future in the present but on a *belief* as to the fixity of future events. The future can be predicted or it can appear to be foreseen because it is assumed to be fated to occur as it does.[11] If the future could truly be foretold, not as a probability but as a fact, this would signify the fixity or fatedness of future events. Precognition entails fixity, fixity is compatible with reversibility, and reversibility is compatible with backward causation.

A theory of the future as fixed ignores the unequal share of the past in the genesis of the present. We live in the present and our concepts of past and future radiate outward from that perspective. Since the present is the window through which past and future are conceived, an understanding of the present is essential if arguments as to the fixity of past or future, or the direction of change and causation, are to be persuasive. If objects are created by a past and then "perish," how is a future object instantiated without an interior past? If, in a reversal of time, a more distant future serves as the past of a more proximate future, one would have a progression from the past (the distant future) to the present (the recent future) in which the (current) past becomes the open future. In that case, how would one chose between asymmetry and reversal if reversal becomes asymmetric in the opposite direction?

Finally, if the recent past of the remote future is still part of the present future, and if the "location" of the now is arbitrary, the fixity of the future would depend on the fixity of a "prior" future, not on the transformation of a past. A theory of the future as fixed requires that the fixedness of past and future be identical, for any point in the future is a possible past or future for any other point. That is, for time sym-

metry there is no real past or future, just events that can be aligned in either direction independent of the event. Such a world is not a world of becoming in which the "not yet" does not exist and the "once was" becomes the now that "is." Symmetric or isotropic time needs an arbitrary now. Anisotropy needs a future that is qualitatively different from the past to conform with the asymmetry of change.

Isotropy is an outgrowth of a theory of causal change. It owes to the fact that many physical reactions are time-symmetric and is in accord with the concept of a fixedness of past and future events.[12]

It is also consistent with the possibility of time travel suggested by relativity theory, which entails that time (space-time) is perspectival, but not necessarily mind-dependent. The combination of isotropic time with a perspectival now has the consequence of a present that is viewer-dependent, thus a moving or arbitrary now. Indeed, a moving now is necessary for the idea of symmetry and fixedness.

Microgenetic theory is also perspectival but within a perspective a now of some thickness is an everywhere-at-once that is simultaneous for all (shared) perspectives. Within this motionless now, time does not flow, it pulses, each pulsation establishing the past, then the present. Becoming is asymmetric from past to present, with each cycle generating a novel state. The future is an idea that is invented from regularities in the becoming. In some respects, the future is like a prolongation of the timeless duration between adjacent presents to an infinite series of subsequent nows.

Is the Future Relative?

If the past is activated and established at the onset of each becoming, and if the terminus of a becoming is the present, could the future depend on the point in becoming from which the terminus (surface) is observed? From a locus in the past, the present is the future. The future is unattainable within becoming for it cannot achieve an endpoint beyond the "knife-edge" of the present. But to a self that is anchored in the past, i.e., to earlier phases in becoming, a surface object that has not yet materialized would appear to lie in the future. To such a self, a subsequent (transpired) past might seem to be a present now. The present would consist of forgotten objects. The series of relapsing events—the decay of a stretch of the present—might be apprehended as a sheet of oncoming events spread out as a future for a self that is fixed in a still remoter past. In other words, the surface of the present, or a subsurface (past) phase could serve as a future for a phase that represents a still more distant past. This follows if pastness and present-

ness depend on phases in becoming and if the self (observer) can "occupy" a different phase in relation to the habitual surface. For a deep self deprived of objects, i.e., a self beneath the "space" of introspection, the future would unfold out of a still more recent past.

This suggests that past, present, and future might depend on the locus of a perspective between phases in an object (or subject). If the self, which is the source of the perspective, can be "located" at more than one phase in becoming, there would be no necessity for the perspective to lie in the "real" present. If this is conceivable, then it is equally conceivable that the future is relative to the locus of the self, and that the world experienced in the present is not necessarily the actual present of the world, nor even the actual present of that self.

To some extent this state is replicated in dream. A dream is a revival of a past memory that is experienced in the present (of the dream). The attentuation of becoming creates a false present out of contents that, in the waking state, would be submerged, and thus past. A deep self deprived of a "normal" surface experiences the images that come up as surface (i.e., present) objects. In this way a memory image becomes, for the deep self, a perceptual event in the present. This is possible because the self is derived from configurations that anticipate the content of the dream.

Such a scenario is not a distortion of "real" physical process but a clue to the way the process is organized. The complexity is so deep that what is plausible, or conceivable, can easily become what is possible, and from there it is a small step to actual theory. I think we have to hold fast to something and this something is the now of real events. This is because a waking self that is capable of observing such events cannot exist without the real present. In dream, there is a present experience without a past or future. In this respect, dream is timeless. The self of a dream is not an intentional self. What, after all, do past, present, and future mean to the self of a dream?

Moreover, an epoch of becoming—the minimal duration of an object or a self—is *not incremented* into successive phases. The precedence of self before world, or past before present, that is responsible for the distinction of mind and world, is *felt* within the becoming. The unity of the becoming is the realization of a self in relation to a world object in the present. The unity establishes the perspective. The perspective is not "located" within the becoming, for the becoming is not divisible into points or loci. The self is not elaborated at a locus or segment but depends on a relation between phases. For this reason, speculations as to relative futures based on a perspectival theory of the point of observation *within* a becoming are probably without foundation.

The Open Future

Change is the potential of the past given up in the fact of the present. The subjective experience of change, however, or the feeling of transition and time passing, is the movement of a self ***through time***. This is a feeling of a shift from a potentiality in the ***present*** where "all things are possible," to a resolution in the ***future***, where this potential becomes fact. In other words, the microgenetic transition (becoming) from potential to actual is felt as a transition from present to future, not from past to present. The feeling of a movement from present potential to future possibility supports the sense of agency, in contrast to a feeling of a direction from past to present, in which the present would be felt as a terminus of a fixed sequence.

To common sense, a state of indecision in the present supposes that a choice will materialize, if it does, in the future. Indecision either resolves or persists, but the resolution or persistence occurs in the future, i.e., a future present. However, when choices are indistinct, the indistinctness has a quality of pastness, at once both potential and lack of resolution. The selection which resolves the indecision is aborted before the final act or object is realized—this is the indecision—and the depth of the choice corresponds with the past of the mental state.[13] The specification achieved in the completion of the process, a resolution that takes place in a future present, is the replacement of the occurrent state (of choice) by another state (of selection).

Accordingly, the future exists in idea, not in fact, that is to say does not exist as an actual state of affairs, not as a "fringe of anticipation" (the "fringe" of the present is its decay in relation to the next rising surface), nor as an elsewhere, another time or place toward which the present moment is heading.

Is a nonexistent future identical to an open future? This is like asking about space-time on the other side of the universe, or about the properties of unicorns. A better question is how the present is established by the immediate past, i.e., how does one microgeny drive the next, and what are the possibilities for novel change in the present; i.e., is change across presents governed by causation, probability, or intrinsic novelty.

An open future can be conceived as forked or branching like a decision-tree.[14] On this view, the forward edge of the present is a bridge to the adjacent border of the future, with each transition marked by a possible world. But if there is no absolute future, i.e., there is no time beyond the occurrent state, with the edge of the present giving way to another present edge, the future would correspond to a ***belief*** in the recurrence of a series of present states beyond the limit of the occurrent

one. On this view, a point in the (abstract) future would correspond to the final (to that point) becoming of the set of intervening presents.

THE FUTURE FROM THE STANDPOINT OF LOGIC

To some extent our views of past and future, and the idea of a fixed past or future, result from the application of logic and tense analysis to time theory. For example, fatalism is the thesis that the future is *logically* inescapable. On this view, the laws of logic suffice to refute the possibility of free will. Fatalism is tied to a correspondence theory of truth. So too is logic. Microgenetic theory entails: (1) a (nonexistent) future; (2) a recurrent present; (3) novelty in the emergence of the present out of a past that is revived; (4) a relation of derivation—not correspondence—between ideas and objects; and (5) a relation of approximation between events in the material world and thoughts or logical proofs in the mind.

Moreover, a logical construal of what are facts in relation to the fixity of past and future is different from the fact of their fixity or openness. Whitehead wrote, "the harmony of logic lies upon the universe as an iron necessity." This harmony is in the ultimate nature of things. The search for this nature is at once the goal of logical thought and an exploration of logical structure. The logic of the universe is implicit in its organization. Is the nature of the universe implicit in the structure of logic? Propositional logic is a language of relations. Wittgenstein wrote, "Am I not getting closer and closer to saying that in the end logic cannot be described? You must look at the practice of language, then you will see it"[15] Logic is a means to understand and communicate the relational structure of a universe that may not exhibit this structure in its own organization. And, language is full of traps.

Logic is a tool to decide among competing possibilities. Could novelty be one of them? Novelty implies indeterminacy. Suppose indeterminacy does not arise because an outcome is a probability, i.e., through chance, but because an event is genuinely novel. Can a universe of continuous novelty be given a logical description? Logic can be applied to indeterminacy,[16] or show that an outcome is indeterminable, but this was at stake before the problem was submitted to logical analysis. So-called change logic presupposes that "time is 'chopped up' in discrete bits."[17] Can logic capture the continuity of becoming or the ambiguity or contradiction inherent in a process model? Perhaps, but then it is a different world from the everyday experience that logic describes. Logic demonstrates the truth or falsity of statements with respect to systems

that have a structure in common with it; that is, logic rewards a system to the extent the system maps to a logical structure.

MIND AND NATURE

The nature that is scrutinized in perception is not the whole of what nature is. There is a nature of subatomic particles within an object, a nature of olfaction and echolocation and infrared, a perceptual nature that eludes organisms without the appropriate sensory detectors. We distinguish a real world that contains the objects of science, and a phenomenal world that is a window of the senses. This distinction is often framed in terms of a causal nature and an emergence in subjectivity. Mind is a complex objectification that seems in opposition to the simpler objects of nature. From inside a mind the world is a kind of spectacle. Yet the material world is active in mental process, not only impinging on mind from outside but generating mind from within. Wordsworth put it beautifully:[18]

> . . . a sense sublime
> Of something far more deeply interfused . . .
> A motion and a spirit, that impels
> All thinking things, all objects of all thought,
> And rolls through all things

If mind is governed by the same "laws" as material objects, a causal nature entails a causal account of becoming. Conversely, if becoming applies to mental *and* physical objects and if mind exhibits emergent properties, change in the material world would also be emergent.

This raises the question of whether phases traversed in the realization of an *actual* world correspond to "real" aspects of the material world embedded in an apparent object. In other words, does the becoming of perceptual objects mirror becoming in the external world? If the actualization of the mental is an instance of actualization in the material world, every mental world is a possible physical world. If phases in subjective space-time have their correlates not just in material brain process but in the objective world, dimensions of physical space-time would correspond to the continuum between buried and surface phases in mentality.

Whitehead believed that stages in the formation of objects were identical for mental and physical entities, every entity having a double aspect. A common organization distributes to the mental and physical out of a primordial unity. The illusory double aspect is created by the

emergence of objects through mind into physical nature. Subject and object objectify together as expressions of a common creative process. I would say that perceptual objects are complex entities in which the subjective is embedded. Subjectivity is a condition of feeling that has the actual world as a goal.

WHAT IS REALITY?

The present does not move, it is replenished. Every present, this one and those in the past, contains my entire past as its greater part. This past is fully in the present when an image approaches an object in clarity; for example, in some types of hallucination, or in memory imagery such as reverie. To recall the past is to revive with more or less exactness a former present within which that past occurred.

The world is real when all perceptions achieve an object that is fully in the present. While there is more than one present for an experience that seems real, there has to be a present of some sort to have a subject for an object to be real to. This occurs in dream as well as waking perception. An object saturated with meaning or feeling is a mental object that seems real when it is the only object one has. The feeling of reality demands a common (cotemporal) object across the different perceptual channels. An attenuated object in one perceptual system, e.g., an auditory or visual hallucination, can threaten the sense of reality in the other modalities. The generation of an object in vision or language can mitigate the effects of an auditory hallucination. The person will report the hallucination as unreal, or the judgment of what is real is overcome by (drawn into) the hallucinatory level. If an object fails to actualize in one system, there is a disparity in object worlds. One world will seem real, the other hallucinatory. If other perceptions join in the hallucinatory endpoint, the object world is replaced by an image world that is no less real to the subject. Reality is not what is out there, since the "out there" has to be generated in each perceptual system. As Bradley wrote, "the merely external is our ignorance set up as reality."

A hallucination is not necessary for a disturbance in the sense of reality. An emphasis on the conceptual or meaning content of objects provides a quality of pastness. The object is more like a thought in the immediate past. This confers a subjectivity on the object that undermines its actuality. The feeling of reality is determined by the degree to which all perceptions actualize to the same endpoint, regardless of what endpoint that is. The subject is a product of this process and does not choose what is real. If I have to choose or if I am uncertain, my world

is already in danger. All of the objects around me are real (actual) in a different way. Is a shadow less real than a tree? What kind of a thing is the shade? The shade needs me as a perspective. Does the tree need me as well? When something is real, what reality are we talking about? Reality is not gripped by a thing that is solid at one end.

NOTES

1. Existence refers to a finite world of temporal facts. The relation of the finite to the infinite is a topic for metaphysical study. For Whitehead, the concept of an absorption of temporal fact into the eternal harmony of God's nature provided a resolution of permanence with transience in such a way as not to lose novelty; L. McHenry, *Whitehead and Bradley* (Albany, NY: SUNY Press, 1992).
2. G. E. Moore, *Some Main Problems of Philosophy* (London: George Allen & Unwin, 1953).
3. For background, see Brown, *Self and Process*, 175–178, 187; and Brown, *Life of the Mind*, 335.
4. Discussed in *Life of the Mind*. On time, see J. T. Fraser, *Time: The Familiar Stranger* (Washington DC: Tempus, 1987).
5. For example, see M. Bender, "Dysfunction in the Visual Perception of Space and Motion," in *Physiological Aspects of Clinical Neurology*, ed. F. Rose (Oxford: Blackwell, 1977).
6. For example, see W. Richards, "Time Reproductions in H.M.," *Acta Psychologica* 37(1973): 279–82.
7. See P. Horwich, *Asymmetries in Time* (Cambridge, MA: MIT Press, 1989).
8. For example, the tribal chief who continues to pray for a successful hunt after it is concluded. See discussion by A Dummett and A. Flew, *Proceedings of the Aristotelian Society*, Supple. 28(1954):27–62; Symposium, "Can an Effect Precede Its Cause"?
9. The transition from simultaneity and timelessness in subsurface mentation (e.g., dream) to seriality and time in waking consciousness may underlie some of these examples. See *Self and Process*, 132, for a discussion of precognition in dreams in relation to levels in time awareness.
10. The possibility that something like backward causation underlies mental phenomena deserves closer scrutiny. Microgenesis assumes a progression from whole to part. A reversal of this direction, i.e., a shift from part to whole, is comparable to backward causation. Moreover, even from the standpoint of whole-to-part action, backward causation may be involved in agent control.
11. See discussion in G. Feinberg, S. Lavine, and D. Albert, "Knowledge of the Past and Future," *Journal of Philosophy* 89 (12)(1992): 607–42.
12. See J. Earman, *A Primer on Determinism* (Dordrecht: Reidel, 1986), on isotropy and a fixed future. Anisotropy also occurs with certain physical phenomena, such as radiation, gravity, and exponential decay, or phenomena such as the wave of concentric circles made by a stone falling in a pond. See also H. Zeh, *The Physical Basis of the Direction of Time* (New York: Springer-Verlag, 1989).
13. Choice is not always at a deep or subsurface phase. One can entertain two or more clear, well-reasoned ideas in consciousness without the background of vagueness or anxiety that is characteristic of preliminary cognition. However, such an individual is confronted with a different type of choice. There may be a relative indifference to outcomes, or the active competition is at a depth beneath the conscious alternatives.

Ideas differ as to their affective valence or "charge," i.e., their valuation, and the depth of their driving concepts. This difference is qualitative. The conscious idea is not a conscious replica of its unconscious precursor.

14. See discussion in W. Newton-Smith, *The Structure of Time* (London: Routledge and Kegan Paul, 1980) and Horwich, *Asymmetries in Time*.
15. Wittgenstein, *On Certainty*, 501.
16. S. Cahn, *Fate, Logic and Time* (New Haven, CT: Yale University Press, 1967).
17. G. von Wright, *Philosophical Logic* (Ithaca, NY: Cornell University Press, 1983). Classical logic is imputed to be the doctrine of the laws of thought, but those laws are dynamic and relational, unlike logic, which has one ontological theme: being as objective permanence. See G. Gunther, "Time, Timeless Logic and Self-Referential Systems," *Annals New York Academy of Sciences* 138(1966/67): 396–406.
18. Tintern Abbey, 97–101.

CHAPTER 4

Privacy

ARGUMENT: *Freedom differs from free will in that the former is defined in relation to rational decision and extrinsic compulsion, while the latter is a private experience centered on agency and choice independent of the rationality of one's options. Agency is a relation across contents in the same or successive mental state(s). The relation is between self and image, including the body image. Privacy is essential for the exclusivity of foreknowledge, the ability of the agent to predict his or her acts and to recollect a prior intention.*

> Thou soul that art the eternity of thought.
> WILLIAM WORDSWORTH

This chapter seeks to address a confusion in the popular mind about the nature of the free will problem, namely, the idea that freedom and free will are equivalent concepts. The distinction is important because these concepts are centered at opposite sides of the mind–world divide, and the epistemics of the distinction will govern the argument in the successive chapters.

To begin with, one can say that a condition of freedom or its restriction does not necessarily depend on the rationality of the subject in that condition. A rational person is necessary to make a judgment as to the presence or absence of freedom, but this person need not be the subject to which the freedom or lack of it pertains. A canary in a cage is not free. Of course, this is not the canary's opinion, but a determination by a rational being familiar with the concept of freedom. Still, a person in a prison cell could be said to lack freedom in the same sense as a canary in a cage. This sort of freedom is largely a matter of limits on the natural behavior of an organism, whether an animal or

a human. In this example, the lack of freedom is irrelevant to whether the canary or the prisoner has free will. The prisoner still has free will even if he is unfree. He can decide whether to stand or sit, to call out or be silent, what to say, and so on. The lack of freedom is with respect to the pursuit of certain activities. There is freedom for one interest or appetite but not another.

In constrast, free will would seem to be an all or nothing capacity related to an awareness of the context and limits imposed on a behavior and the alternatives that are possible. Freedom would then be the realization or *exercise* of free will. This would mean that a state of free will might occur without the conditions for freedom, though if free will is lacking in a given circumstance freedom could be a meaningless concept for that occasion.

On this interpretation, free will relates to a decisional process while freedom relates to the ability to act on some or most of one's decisions. In the above example, free will might involve the intention to escape from a cell, and freedom the effort to realize that intention or its occasions of satisfaction. In a word, free will is the ability of the self to choose and decide, while freedom is the implementation of an act that is freely decided.

If a canary is happy in its cage and does not leave if the door is open, would one continue to say the canary lacks freedom? Perhaps it might be argued that the range of behaviors open to the canary is limited by learning or experience, but this is true for all of us. People who are abducted can learn to love their captors. In the case of a human prisoner, the contribution of rational thought is important, not for the application of reason to decision making, but because the presence of reason is a benchmark for whether a person has awareness of his or her options and the limits that are imposed on those options by the environment. If one is indifferent to a hindrance it ceases to be a hindrance. The definition of a hindrance requires a block or impediment which in turn implies an inclination that is impeded, precisely what is lacking in indifference.

Suppose a person desires to remain in a cell because he understands he is a danger to others. Does he have less freedom by remaining in the cell if the desire to remain in the cell is based on an evaluation of his condition? Since there is a conformity of judgment or decision with action, free will and freedom are congruent. If the person chooses to be confined, the condition of confinement satisfies his rational choice. The fact that he is not "free to leave" is irrelevant to his freedom, defined as the opportunity to pursue a rational goal.

Freedom is *specific* to an occasion while free will is a *generic* capacity. If there is no free will, there is no freedom for any occasion

though a lack of freedom for a given occasion, e.g., confinement, if freedom depends on reason, obligates a capacity for free will for any occasion. The absence of the conditions for free will presumes the absence of freedom, at least in humans if not in canaries. If a person cannot choose, what freedom is there in the action chosen?

There are, however, occasions of deficient free will where it is unclear that freedom is obstructed. A person asleep in a prison cell, a lunatic or an imbecile, may not be aware of a lack of freedom, nor are they rational in the sense of having an ability to choose. If we assume such an individual is unaware of being confined, could we say that the individual does not experience a deprivation of his freedom to leave the cell since the awareness of confinement and the option of leaving are not present in consciousness? Free will is lacking in a person who is asleep, but is the concept of freedom relevant to such a person?

Suppose a person is free to travel but lacks the money, or has the money but can't obtain a ticket, or has a ticket but is ill on the day of departure. In each case, there is an obstacle to the realization of a desire to travel. We would not claim in these cases a lack of freedom, since travel in the future remains a possibility. Yet we would not wish to say that a deprivation of freedom requires a condition of some permanence.

Irrational action is another matter for it accentuates by contrast the excessive importance given to reason in the determination of freedom. Irrational action may be unfree if the irrationality is extreme, as in psychosis, but this does not imply the contrary, that free will increases with the rationality of an act, leaving aside the question of whether or not there is free will. An action that is irrational is inconsistent with free will since irrationality obviates choice, but choices need not be rational to satisfy the conditions of free will, they need only to be uncompelled. Reason presumes choice, while free will requires choices but not necessarily rational ones. However, irrational thinking does not ordinarily include competing irrational options. The psychotic is not usually confronted with a decision between two equally irrational acts. The very presence of choice supposes some level of rationality.

Indeed, one could say that free will occurs when there is choice regardless of the choice that is made. Every choice is more or less rational, or has a stronger or weaker emotive basis, so to say that choice must be rational for the will to be free is to deny what is inherently human in choice. A computer could be programmed to always make the most rational choice, but the program would not be construed as capable of free will since human choice is always motivated by character, values, and so on.

Many problems are created by an account of freedom in which rational thought, not just the absence of irrationality, is an essential ele-

ment. In addition to those discussed, there is the opposition of reason to desire or the role of beliefs that are deemed to be false. Is a decision unfree if it is based on love, faith or loyalty? If lovers and soldiers, as Shaw quipped, are irrational by definition, does one cease to have free will the moment one falls in love or joins the army? A person who makes a decision on the basis of a passion or a desire, or someone in the grip of an unopposed belief, even the possibly false belief in the existence of a world of real objects independent of but identical to our perceptions of them, might be judged as not fully rational and to that extent lacking in free will.

It is a useful exercise to review some of these issues because they permeate our thinking about freedom, and it is not an exaggeration to say, contaminate the literature on free will. Reason is important to freedom because it implies the capacity for choice, not because an action must be rational to be free. Indeed, most of our actions are guided by habits, preferences, and desires. The absence of reason is an absence of choice or an absence of the awareness of choice, which implies an absence of free will and, by inference, an absence of freedom.

These two aspects of freedom, the communal or interpersonal and the cognitive or private are often entwined in philosophical argument. The basis of free will in choice infects the idea of freedom in the requirement for rational decision, while the role of external objects in the concept of freedom infects the idea of free will in the degree to which a behavior is compelled. The antisubjectivism of much current philosophy also tends to bias the interpretation of free will to the degree of liberty available to a person in a given situation. There is a displacement of the concept of free will from the mind where it belongs to the interaction of the person with the object world. Free will comes to be interpreted as the balance between the rationality of an action and the limits on its enactment.

The effect of this trend is to inhibit thinking about free will from the perspective of the agent, that is, as a phenomenon centered on agency and choice, essentially a private experience involving the potential for a multiplicity of acts, the survey of options, the decision on a goal, and the feeling of self-initiation. Still, we are so transfixed by the role of the external that any theory of free will must undertake to resolve the interiorness of decision making with its effects on objects, including the parts of the body. World events recur as targets for the instantiation of choices. The world infiltrates the free will discourse in the form of occasions for the satisfaction of the goals of an intention.

This chapter considers free will as an interior state, the components of which owe their description to a reconstruction from the symptoms of brain pathology. A symptom is an exteriorization of a private mental

content. The privacy of the content is less a property of the state than an ingredient. For example, it is more accurate to speak of an expectant privacy than a private or personal expectation. Privacy is more than an intrapsychic locus or a subjective standpoint. Privacy makes a fundamental contribution to each of the main components of a volition, including self-awareness, the feeling of agency, anticipation, recollection, foreknowledge, and the realization that an action discharges in the *body image*.

What Are the Objects of an Action?

The belief in the existence of free will rests on the assumption that consciousness or choice can intercede in the flow of events. The event flow is a description of the perceptible change through time that occurs in the progression of self and nature. Essential to the idea of free will is the feeling that consciousness can penetrate a chain of physical events to induce an outcome that is an escape from the chain, i.e., that consciousness "makes a difference" in a causal sequence. There is an awareness of two distinct series of events, one private (mental), one public (physical). The mental series seems to effect the physical series, and the physical series seems to include the agent's own body. The will, say, to lift the hand, is apprehended as a conscious state in which the self effects the body as a physical entity which in turn causes a change in an independent external object.

The self or will does not have to effect a physical object for the feeling of agency. Inner speech and mental imagery have a volitional character. The self feels it can search out a memory image or picture an elephant on the head of a mouse. Images are revived and altered "at will." Thoughts can have a volitional character even in the absence of action. An amputee can will a phantom limb to move and feel the phantom movement. The effect of the will on mental content is similar to the effect of the will on the body. The will as an interior agent is (feels) identical to the will that acts on objects. There is a subtle difference in the experience of willing on an image or on the body since the target of the willing seems to be mental in the former and physical in the latter. But the self that speaks or lifts the hand is the same self that "looks up" a memory or "manipulates" an imagination image. Since the self can act on other mental contents with a feeling of volition comparable to that of an action on a body part, even if a perceived body part is not construed as a mental object, which it is, interaction from self to world is not a condition of willing.

The will acts on images and the body, not on external objects such as a chair; that would be telekinesis. The body is not an ordinary object.

It is a mental image[1] that is part internal, part external. The body image partly exteriorizes as the perceived body. This was captured by Mark Strand in the lines: "Our images blaze a path/That our poor bodies must follow." Other images—illusions, hallucinations—may also partly exteriorize. An illusion is usually defined as a perceptual distortion of an external object. The spatial or temporal distortion indicates that the object is incompletely specified. The illusion is not a change in a real object but is an image that has not been fully objectified. The distortion is a sign that the object is an image in the mind of the observer. According to the prominence of the imaginal content, i.e., the degree of distortion, of form, or of meaning, the individual may be more or less uncertain whether the object originates in the mind or the world. Hallucinations may also have this quality. The observer is unclear whether the hallucination is an objectlike image or an imagelike object.

The body image (p. 129) is similar to a visual illusion or hallucination in that it is transitional from an intra- to an extrapersonal locus. The body image exteriorizes but not to the point where it is independent of the self.[2] The feeling that the body is one's own image is probably due to the somaesthetic or kinaesthetic sense. With a loss of joint and muscle sense in a limb, leaving only visual sensation, i.e., seeing the limb but not feeling it, the limb may exteriorize like a visual object. Patients say their body part is in the world or belongs to someone else. In cases with denial of hemiplegia, patients can see their paralyzed limb but may believe it no longer exists or is a part of another person.

The opposite condition occurs in amputation. There is a loss of visual perception of the limb, i.e., a perception of its absence, yet sensory feedback from the stump or the action discharge is maintained. The subject sees the stump and ***knows*** the limb is missing in spite of a somaesthetic or kinaesthetic phantom. That the feeling of the limb arises in recurrent perceptions generated by the central action discharge is confirmed by cases where the phantom is lost after a stroke in the opposite cerebral cortex. The difference between a real limb and a phantom is that the visual absence of the limb undermines the kinaesthetic feeling of its reality. Vision disconfirms the false belief in a real limb that arises through intact perceptual systems. Normally, vision carries the limb outward and transforms a kinaesthetic phantom (image) to an actual object.

In sum, the agent or self acts on mental images and the body image and through the body on external objects. The self is apprehended as distinct from the imaginal contents it generates but the experience of agency is essentially a relation between *intra*-personal contents. Agency is not action on objects but an effect of the self on an image. When I

move my finger, my self acts in relation to a perceptual image (of a body part). When I imagine an elephant dancing on a mouse, my self also acts in relation to a perceptual image. Since an image is a derived part of the self, the self acts in relation to its own generative content. Agency is a relation across successive moments (self, image) in the derivation of the mental state. An image that is continuous with the self is the mental correlate of the action. Agent and action are segments on a continuum of change.[3]

FOREKNOWLEDGE

Agency, therefore, is a private feeling of causal effectuation by the mind on its own content even if the self feels an engagement with external objects through the body (including vocalization). The "objects" of agency are private objects, i.e., images. The privacy is needed to establish the experience of agency.[4] An action can be announced in advance, and a person other than the agent can be informed of an ongoing intention, but the experience of agency depends on the anticipation of an action by the agent, i.e., the foreknowledge of the effect, and the recollection of a prior decision to act, i.e., the memory of the cause.

The privacy of foreknowledge—knowing at the present moment that a personal action will occur in the future—contributes to the feeling of agency and self-initiation. The privacy is as important as the foreknowledge. If my action is known beforehand to others, or to God, my action is not truly free. If another mind knows how I will act, and this knowledge is prior to, or even simultaneous with, my own knowledge, my action would have the appearance of being destined or predictable and my sense of freedom would seem illusory.[5] The possibility of my having done otherwise, which is essential to choice, would also be thrown in doubt. The privacy of foreknowledge is essential to my belief that I instigate the action. With shared foreknowledge how would I determine who was the agent of my actions or intentions? If another mind knows of my action in advance, would I not wonder if the other mind provoked the action or the intention to act? If the other mind "caused" my intention to act, the process leading from that intention might give rise to an action for which the intention was a necessary preliminary. The entire complex from an alien intention or my own intention to an alien action or my own action might elaborate a feeling of agency that would be indistinguishable from a purely intrinsic sequence. In the former, the feeling would be a deception since the original intention was induced by another agent. The privileged access to personal foreknowledge,

therefore, not only heightens the sense of agency but is crucial to the belief that one is the sole agent of one's own acts.

Shared foreknowledge is a bit like fortune-telling. Freud thought the secret of the fortune teller was to intuit a person's innermost wish. A prophecy based on that wish might tend in due course to be satisfied. The self-fulfilment depends on the concordance between the prophecy and the wish, since discordant prophecies, like odd psychoanalytic insights, would tend to be rejected. To know a person's wishes or desires is to be close to a foreknowledge of their intentions. The possibility that someone knows my desires or intentions is vaguely unsettling to my belief in freedom, but not as disturbing as the ability of a psychic to predict my future behavior or to forecast future events in my life.

For example, a psychic might reveal either my wish to go to France next year or the *fact* that I will go to France next year. In the first case, my intention to go to France is known and its outcome (going to France) is inferred. In the second case, the outcome alone is foretold. The intuition of a desire by a psychic is less discomfiting than the prediction of events in my future, or my future actions, since a desire, say to go to France, does not regularly lead to an action but an action that is accurately predicted appears to be unavoidable. Yet so many people are drawn to the pronouncements of psychics that the fatalism and despair to which infallible predictability could lead does not seem incommensurate with a belief that foreknowledge is available to one supposedly gifted with such powers.

One seeks a fortune-teller voluntarily. The belief that others can surreptitiously "read one's mind" is common in psychotics. They believe (and fear) that other people are reading or inducing their thoughts. They believe that personal foreknowledge is shared with other people or spirits, generally malicious, or that external thoughts are initiating and guiding their actions to which they are passive onlookers. In other words, they have lost the feeling of being agents to their own thoughts or actions. This is a sign that a thought, an intention, or an action plan is beginning to objectify as an actual entity. Phases that serve as potential for objects now actualize as termini. This is the meaning of an objectification of thoughts. The thought or concept behind the object becomes an object and is apprehended as extrapersonal. The psychotic has lost the ability to prevent the same intrusion into his intentions or future actions that is welcomed by the person who seeks a fortune teller. Here we see the contrast between a *shallow belief* in the possibility that one's foreknowledge is shared with a psychic, and the conviction in the psychotic that such sharing actually occurs. Feeling transforms a belief to a *conviction* that can shake the very foundations of one's being.

One consequence of the objectification of thoughts is an alteration of the feeling of an opposition between the self and the external world. The boundary between self and world requires an actualization over the full set of microgenetic phases. An actual world has to materialize. The process has to eventuate in an external object. A concept that has not individuated to an object is an objectlike concept or an incomplete object, in any event, a concept that is no longer felt as personal. For a concept to actualize and contaminate an object experience is equivalent to the withdrawal of the object to a concept. Concepts must remain interior and opposed to objects for private knowledge to remain private. This is a fragile process. Feeling penetrates belief to give a conviction of the existence of *internal* concepts and *external* objects. This conviction, like that of private foreknowledge and agent-causation, is eroded in pathological states.

Novelty, Prediction, and Privacy

Privacy is essential to the interpretation of an action as *novel*. The novelty is a function of the timing (delay), the possibility of choice and deception, and the departure from repetition, habit, or obligation. Novelty requires a comparison with prior acts. A person who cannot recall a prior action is unable to say if a present action is novel, or what it is novel in relation to. There may be a reduction of novelty in behavior in some cases of brain pathology, as in the perseverations of frontal lobe cases, or the repetitious speech or behavior of Alzheimer's disease. This disorder may reflect the inability, with severe forgetting, to represent the recent past as a cumulative baseline to compare with a behavior in the present. The Alzheimer patient does not know that his behavior is repetitious. Whether a behavior that is accompanied by a lack of awareness for its repetitiousness is construed by the subject as novel is a question for further study. In any event, this form of repetition is not the deep recurrence that is obligated in theories of time reversibility, just as the "clinical" novelty that is lost in such cases is not the *deep* or genuine novelty that characterizes change.

Privacy is essential to choice but no guarantee of novelty. There are obsessive thoughts, habits, routines and compulsions. Habit is the ruin of novelty. The will that is "applied" to an image or an action is under the constraints of the self and its constitution, i.e., character, and character is a constraint on the unexpected. The will is also under the omnipresent control of others, in suggestion, in social obligations, or in peer pressure, no less than in hypnosis or "brain-washing." Privacy and consciousness of the private state are necessary for the feeling of vol-

untary action, but their presence does not entail that an action will be volitional. Compulsive or habitual thoughts and actions with awareness and foreknowledge still do not have the same claim to volition as novel actions. The more obsessional a behavior, the less choice in action and the less free the act appears.

The absence of choice in action should correspond with an increase in predictability, but this is not clearly the case. The behavior of an out-of-control psychotic who does not contemplate choices is unpredictable, at least by others, and an action that is unpredictable to others, in psychotics or as a result of drugs, fever, etc., is the very opposite of deliberate, rational voluntary action. Such behavior is perceived as random or chaotic. Conversely, a knowledge of human nature can often enable one to predict the course of action that will be followed by an individual even if many options are available. Novelty requires choices as alternatives to the certainty or the prediction of an action, but novelty is a judgment based on a succession of acts. The choice for volition is between two options in consciousness; the comparison for novelty is from a past to a present action.

There is a complex relation between choice, novelty, and prediction. With an increase in the novelty of an action, its predictability declines. The predictability is an impersonal judgment by others. This differs from a prediction by an agent in the context of a personal foreknowledge. Since free will would appear to be incompatible with a high level of observer predictability, an inverse relation should exist between the ability of others to predict an action and the actor's feeling of agency.

The ability of an agent to forecast an action that to others is unpredictable, based on their knowledge of the agent's character, for example, a respectable person who announces to their astonishment that he will commit a criminal act, is no assurance the agent is acting volitionally. Indeed, in such a case the prediction is probably grounded in a passionate impulse or an irrational belief. The foreknowledge by the agent, the dissociation from character, the lack of predictability by others, the novelty, all properties of a freely willed action, do not guarantee that the action will be volitional. This is why *choice* no less than foreknowledge is essential for volition, choice in the mind not in the world, i.e. different concepts not different objects, including the "objectless" choice of not acting. Prediction and privacy interact in complex ways in the judgment of agency, a judgment we are constantly making about the actions of others and implicitly extracting from our own volitions.

Privacy is the self's commentary on subjective novelty in the transition across moments in life. Conversely, the publicity of states that

anticipate an action is apprehended as a measure of its predictability. The more public an action the less it is *my* action. The predictability of an action by others implies that the action is highly constrained, therefore less "free." But the prediction of a personal action also constrains one's options. If I announce that I will move my finger in 30 seconds and then after 30 seconds do so, the action is motivated by the prediction. In some sense, even in the privacy of the agent, the more predictable an action the less "free" it is than when guided by tacit deliberation or personal decision. This is partly a result of the loss of spontaneity, but mainly because an unannounced action is apprehended as more easily rescinded. As every politician knows, an agent who announces a future action constrains his choices. This is paradoxical, since the ability of the agent to forecast an action enhances his sense of agency. The paradox arises in the conflict between the personal and impersonal standpoints and the fact that to verbalize an action is to convert the private foreknowledge or potential of an idea to a statement of fact. In becoming a statement, an idea becomes an object with causal powers. This is the basis for the apparent increase in predictability. As the King in Shakespeare's Hamlet says:

> My words fly out, my thoughts remain below,
> Words without thoughts do not to heaven go.

Still, an agent must be able to predict his own actions. It is not enough to have foreknowledge of a forthcoming act. The foreknowledge must entail a feeling of empowerment and a direction, a feeling of inevitability once the self makes a decision. This feeling is a kind of agent fatalism. The direction is the inevitability of the oncoming moment. Prediction is a direction toward a future state, whether in the immediate or the distant future. The knowledge that I will go to France next year has a certain conviction. The conviction is not for the outcome but for the state of current belief. There is agency in uncertainty. A trip to France is always contingent. Will I be alive next year, can I afford the trip, etc. Whether I will go to France is irrelevant. What matters is that the present belief contains a conviction even as to its uncertainty, i.e., the conviction that I have the current belief that I will go to France and the conviction that this is my belief.

Privacy and foreknowledge are features of thought or choice before acting. To think about an act before acting, i.e., to deliberate before a decision or intention, is to explore the undersurface of the intentional content. Thinking is an outpouring of the conceptual antecedents of what will be the final object (or goal). One could say that the iteration of a mental state to a phase coextensive with the conceptual basis of the intention extracts additional content from the original concept. The

delay before action and the privacy of the delay, enclose options that expand the phase of potential before the act or object actualizes.

RECOLLECTION AND AGENCY

The revival of contents from the concept of the action-to-be, i.e., the initial intentional state, at the moment of the volitional act, the occurrent state, is a revival through the present self of a prior state of intention. The revival of this "past" content through the present self accounts for the belongingness of the contents, the feeling that the concept or plan of the action, which is an idea from a prior state, is *my* plan. The prior *self* that intended the act is not revived with the intention, yet the occurrent self remembers the prior intention as belonging to the prior self. That is, the concept or intention, say "to lift the finger," is revived but not the self, the "*I* (will)," that gave rise to that intention.

The temporal tagging or coding of prior events, episodic recall, is not for the self of a prior state. We remember an experience or idea in the immediate or distant past but not the self of that past. We can revive multiple ideas at the same time but not multiple selves. The price of the oneness or identity of the self is an inability to recall the self of one's past. The identity of an idea *is* its contrast with other ideas. For the self, such a contrast would be oblivion. We forget the self of the past in order to have a self of the present. Put differently, the idea of a past self is weakly distinguished from the idea of the occurrent self. The self-concept differs from other concepts. Why this should be so, i.e., the problem of the identity of the self through time, is a topic in need of closer study. Indeed, a self apprehended as an idea that is equivalent to other ideas is inconsistent with a unitary self-concept. If the momentary self is like a fleeting idea, and it is no simple matter to say why it is not, there would be no self to have an idea in relation to. Perhaps this is the basis of a multiple personality disorder, in which the prior selves of a person are not recollected in the current self but "exist" in the form of isolated ideas (p. 124).

In a volition, there is a delay before action, however brief. A failure to recall at the moment of an action the prior state of deliberation, or the decision or intention to act, or a loss of the relation of the decisional to the acting state, i.e., the sequence and direction of the states, or a disturbance in the duration of the delay, or an inability to control the timing of the act over this duration, would abolish the volition and consign every action to a durationless present. The test of this supposition would be a case of *severe* amnesia where forgetting erodes into immediate memory. If an intention is forgotten the moment it is conceived,

e.g., the idea to lift my finger, and there is no rehearsal of the intentional content, even with a continuity of the self-concept, there should be a loss of the capacity for volition. Such an individual would lose the feeling of agency if this feeling depends on a present knowledge of a prior intentional content. Presumably, the causal link between an intention and an action, thus the feeling of being an agent to the action, depends on the before of the intention and the after of the action. This feeling is established in the stacking of prior presents in the occurrent state, with the intention and its "location" in the past reconstructed in the becoming of the present.

Finally, privacy enters recollection as it enters foreknowledge, except that at the moment of action the anticipation that is necessary to foreknowledge becomes a retrospective that is necessary to memory. The intention (to lift the finger) has this double aspect. It is apprehended as past idea in relation to the present, and it is apprehended as a past idea that is directed toward the future. The present self "looks back" at the intention, while the intention "looks ahead" toward the action. The transformation of anticipation to recollection, a prior self that predicts an action and a present self that remembers the prediction, is an outcome of the embedding and replacement of the series of mental states that constitute a volition.

NOTES

1. P. Schilder, *The Image and Appearance of the Human Body* (New York: International Universities Press, 1950); *Self and Process*, 112, for discussion of the body as an intermediate object; S. Gallagher, "Body Image and Body Schema: A Conceptual Clarification," *Journal of Mind and Behavior* 7(1986): 541–554, on the distinction of a deep or unconscious body *schema* and an explicit or conscious body *image*.
2. The transitional status of the body image in action recalls the interpretation of the function of the frontal lobes as integrating the internal with the external environment. Put differently, an action deposits at a phase intermediate between the internal and external segments of the mental state.
3. J. Kim, "The Nonreductionist's Trouble with Mental Causation," and J. Hornsby, "Agency and Causal Explanation," in *Mental Causation*, ed. J. Heil and A. Mele (Oxford: Clarendon, 1993).
4. On some of the various meanings of privacy in the philosophical literature, see T. Sprigge, "The Privacy of Experience," *Mind* 77(1969): 512–521.
5. See the discussion in J. R. Lucas, *The Future* (Oxford: Basil Blackwell, 1989).

CHAPTER 5

Will, Agency, and Constraint

ARGUMENT: *Primitive will is the core given over to action, affect and subjectivity. Subjectivity is elaborated as the will diverts from instinctual routines bound to environmental objects through an intrinsic development to the drives. The drives partition the will into desires, needs, wants, etc. These are tributaries of drive as it distributes into the hierarchy of values in the self. The goals of a drive become the concepts of a desire. External constraints are simultaneous with the contents they elicit. Intrinsic constraints apply within the mental state and from one state to the next. Constraints alter probabilities of outcome but are not direct causes.*

* * *

The unfolding of the human mind begins in a *core* of primitive will. From this core, the will leads outward in instinct or inward in drive. Phylo-ontogeny lays down the instincts and the drives. In the course of this process, the prominence of highly constrained innate patterns of behavior shifts to the prominence of the effects of learning and the role of self-initiation, and a relative freedom from objects in the surround. The specification of drive out of instinct is largely innate.[1] The specification of desire out of drive depends on acquired values and beliefs. The seemingly disparate phenomena of drives that are innate and values that are learned can be interpreted in relation to an intrinsic process of fractionation through the constraints of sensation (experience) at successive phases in the derivation of the mental state.

What is a constraint? Is a constraint a cause? The question is important to our understanding of causal effects in the generation of the

mental state, i.e., whether a mental content is *caused* by the constraints that delimit its features. If constraints are causes, what is their relation to agent causation? If an agent constrains a developing action, does this constitute a causal role? Before this problem can be discussed, however, we need to consider the concept of will itself and the manifestations of willing in thought and action.

PROBLEMS OF DEFINITION

The primitive or instinctual will is the engine or motive of the *deep* self.[2] The deep self is a primitive concept that encloses the instinctual will. Will and self, drive and concept, are different aspects of the same construct. The will is the dynamic or energic aspect, the self the substantive or conceptual aspect. Or one could say, the self is the will from the standpoint of representation, the will is the self from the standpoint of kinetics or affect. Will develops through drive into desires and feelings. The self develops through beliefs and values into ideas and objects.

What exactly is the will? The meaning of the term is confounded by the different forms it takes on in the becoming of the self and the application of the term to many different types of mental events or actions. There is uncertainty about the definition, expression, or employment of the term, which may even be specific to a language. For example, the distinction of willed and voluntary action is not fully captured in the shadings of the French word *volonté* or *Wille* in German. However, most Western languages have words related to the English usage.[3]

As a modal verb the will is appended to the self as in, "I will lift my finger." This entails the meaning of will as an expectation of a future action. There is also the implication that the self, or its conscious implementation as the "I," has a plan or intention that is embedded in the "I will." This is evident in the fact that the meaning of, "I will lift my finger" strengthens or makes more definite the meaning of, "I plan or intend to lift my finger." This relation is also embedded in the use of the future tense. The French, *je lèverai mon doigt* accentuates in a similar way the meaning of *j'ai l'intention de lever mon doigt*.

As a regular verb, the will is what the self does to get an action started, as in, "I will my finger to lift." Here, the will is an activity of the self. It is the self's action or a link between the self and an action. In this sense, an act of will is a choice the self can make, one of many possible choices, such as guessing, thinking, or hoping. As with the *I will* in "I will my finger to lift," thinking and hoping are intransitive verbs that do not involve actions on external objects. In the above statement, the "I will" is nonmotoric. The actual movement is described by

the verb *lift*. Willing is comparable to thinking or hoping as an intransitive activity of the self. Hoping, thinking, and willing tend to prefigure (motor) actions. The self wills, i.e., causes an act of willing to occur, in the same sense as the self thinks, hopes, and so on. If thinking and willing are comparable, the activity of willing might underlie the activity of thinking. Put differently, if willing is an act of the self like thinking or hoping, are thinking and hoping acts of willing?

The will can also be interpreted as an impulse or effort that inheres in *willed* action. On this interpretation, the will is an attribute of certain but not all actions. The presence of this attribute implies that the action is self-generated and usually voluntary, but not necessarily that it is "freely decided." A willed action may be highly constrained so long as it is not involuntary in the sense of automatic. If someone puts a gun to my head and directs me to go somewhere, the going would be a willed action that was involuntary. Volition requires the possibility of choice, whereas willing concerns the conscious control of the action.

Conversely, one can say, "I went on the airplane unwillingly," e.g., if I was apprehensive about a plane crash, but this does not make the action involuntary. In this instance, the choice among competing options was not sufficiently persuasive to prevent the person from boarding the plane. The boarding is a willed act, even if done "against the will." Indeed, in this example, where the conflict is between the desire to leave and the fear of disaster, the more superficial derivation of will into desire overcomes its deeper derivation into drive (fear). Here, the presence of choice makes the action volitional, while the conflict with drive accounts for the feeling it is done "against the will."

The will as a noun is an entity or state that empowers an action. A state of the will is distinguished from a state of the self, as in "the will to lift the finger" or "a lack of will." For example, an individual can say, "I haven't the will to do that." This meaning implies an interaction between the self and the will or between the will and other mental contents. One could say, "I found the will to lift my finger," if finding the will is not itself a form of willing. The belief that the will is a kind of entity might (mis)lead one to search for a region in the brain where the entity is "located." A "loss of will" could then be associated with damage to this region, while impulsive (willful) behavior might be attributed to a state of hyperactivity of the region.

Suppose one says, "The will willed the finger to lift." This may sound odd but it is acceptable as a statement. Here, the noun *will* is identified with the self as a cause of willing, and the verb *will* is an "act" of the self or the will prior to an action. The noun concept of will does not entail an identification of will and self for, as mentioned, it can lead to statements such as, "I acted against my will," implying that

the self can will (initiate) an action other than the action desired by the will; i.e., that self and will are distinct, even conflicting, entities.

The self that wills an event is a self inclined in a certain direction. The inclination is the willing, the will the measure of the self's inclination. If a direction of the self is an act of will, is the act of will an action of the self? To say that when the self wills an action, the act of willing is itself an action, is not quite the same as to say when I think a thought, the act of thinking is itself a thought. We distinguish a thought, and the act of thinking a thought. The thought is the content or product of the thinking, and thinking is the activity or the process of thought. Is it the case that willing an action is in the same class of events as thinking a thought? If the action is a product of the willing, or the willing is the process that leads to the action, and if an action is a *content* in the mind like a thought, perhaps thinking and willing are the same *kind* of activity.

To say when "the self wills an act" that the act that is willed is one action and the willing is another action, is to say that willing and acting are actions of a different type. In thinking a thought, the thought is a perceptual content that actualizes in the process of thinking. An action is also a content in perception. To say the expression, "the self wills an act," is comparable to the expression, "the self thinks a thought," implies that willing is not an action distinct from the action that is willed.

Suppose one interprets *I think* as thinking is going on and I am thought up in the thinking. The self is a phase in thinking (willing, feeling, etc.) out of which the content of the thought is extracted. Put differently, the greater (deeper) part of the activity of thinking deposits the self and the lesser part deposits the thought. For the *I will* to be like this would mean that in the becoming of *I will,* first the conscious self, the "I," is deposited, with ensuing phases of decision or goal-directedness labeled as thoughts, hopes, willings, etc., or as actions.

One objection to the idea that willing is a form of acting is that this multiplies the causes of an action. If to say "the self wills" is to say "the self acts," and if a willed act must be preceded by an act of will, a state of will must be prior to every act of willing. The self must then will to will an act, and so on. This problem of regress is eliminated by recognizing that the self and the will are different aspects of the same entity. The self that feels that it wills (thinks, hopes) is the self as an agent or the agentive feeling of a self as it actualizes.

Varieties of Willing

As discussed, the expression of will differs in its actualization from core to surface, with a continuum from instinct and drive to desire, feeling,

and the affective tonality of images and perceived objects. The nature of will is elusive because it refers to different segments in this continuum, some of which are artificial entities, e.g., properties of social interaction or dialogue, others of which are "natural kinds," in that they count as correlates of mental process.

In the primitive or core self, subject and drive state are indistinguishable. The subjectivity of an animal inheres in its drives. One could say, "A bear *has* a drive (to feed, fight, etc.)" but we do not ordinarily distinguish a bear, as a subject, and its drive states as something belonging to the bear. Actions that realize drives expose a subject as the momentary drive state it instantiates. The subject of the drives pursues satisfaction. A bear is adorable or ferocious depending on the drive at a given moment. The subject revealed by such actions, e.g., patterns of feeding, sexual behavior, is the evanescent drive that is realized.

We could say, "A bear has a will to survive," but not "A bear wills an action." This is not because a bear is incapable of willing, but because a bear as a subject is identified with an act of will as a drive. For a bear to will an action would require the bear to have a self-concept distinct from the act of willing. The bear expresses will in the form of instinct or drive, but (presumably) lacks will in the form of desire or wish. For the feeling of agency, the will must be experienced as a link to an object or goal. As drivelike states of willing differentiate to desires, the core self differentiates to a (conscious) self-concept. A desire is not a muted or partial drive but a qualitatively unique individuation.

Choice seems to require the inhibition or suspension of will as drive. The idea that choice or deliberation develops in the delay before action corresponds with the concept of action as a motoric set that is empowered by a drive energy which can be released or withheld. An energic theory of drive entails a distributive or allocative model of energy. In such a model, a more or less unitary self is a source or avenue for the discharge of drive into acts, objects, and ideas. However, choice partitions the self as it partitions the will. The partition of the deep self leads to the self-concept with its explicit beliefs and desires. A different aspect of the self is at stake in every decision.

The energic theory of drive gives rise to the problems of weakness of will (akrasia) and intensity. Akrasia is not a reduced intensity of willing but a form of indecision or incomplete commitment. This is also a "weakness" of the self. We acknowledge the occurrence of degrees of willing but not degrees of selfness. Yet different "selves" or attitudes of the self are exhibited at different times in the "same" person (as in the "same" bear).

Degree of willing is considered a function of intensity, but intensity is a feeling of the degree to which an act is voluntary. This is not[4] a

matter of the effort expended in action for the feeling of willing can be acute for subtle acts such as the pause during or after a conscious decision to lift the finger just prior to movement. Nor is intensity related to conviction. There may be inability to act out of fear or futility in spite of a strong commitment. Nor is intensity a function of decisiveness, which tends to usurp rather than resolve deliberation and occurs in the absence of conscious willing; e.g., in a posthypnotic suggestion. The engagement of the self in an action entails some indecision regardless of the nature of the choice. The self-concept comes into prominence in volitional acts that, because they are volitional, involve choices.

There is a qualitative aspect to intensity since a difference of degree is also a difference of kind. There is also a distinction as to whether the intensity is a personal quality or an impersonal attribute assigned to an agent by an observer. From the impersonal standpoint, one can say the intensity of an *action* lies in its objectivity or display. In contrast, the intensity of *willing* is a measure of the subjectivity or interioricity of an action from the perspective of the self. From a personal standpoint, intensity is a judgment of choice and control, i.e., the extent of self-participation in a given act.

With regard to choice, willed action does not obligate that a choice is made in or by consciousness, only that one is conscious of the choice that is made. Suppose I say, "I am going to a movie now" or I simply get up and leave the room and go to a movie. There is no question these are willed actions, where the option of not going to the movie or doing something else is inferred to be in the background. These options are potential actions that are not in consciousness. One cannot infer that they are tacitly present in consciousness but unelicited in favor of the one that is chosen.

Will and Drive

The primitive will is the first actuality of mental process, the instinct to survive that is the unconscious urge to subjectivity. The inception is with a configuration in upper brainstem or diencephalon that deposits a ***core*** of incipient acts and percepts. This phase of willing, holistic, energic, animal, individuates to progressively less intense and qualitatively distinct states of drive, desire, preference, whim, and finally, the affective tonality of objects. The manifestations of willing, i.e., whether an act is motivated by drive, desire, or preference, depend inter alia on the segment that predominates in a given mental state (Figure 5.1).

The deep self is the precursor of concepts and objects and creates the inner life. The primitive will is the origin of drive and affect and

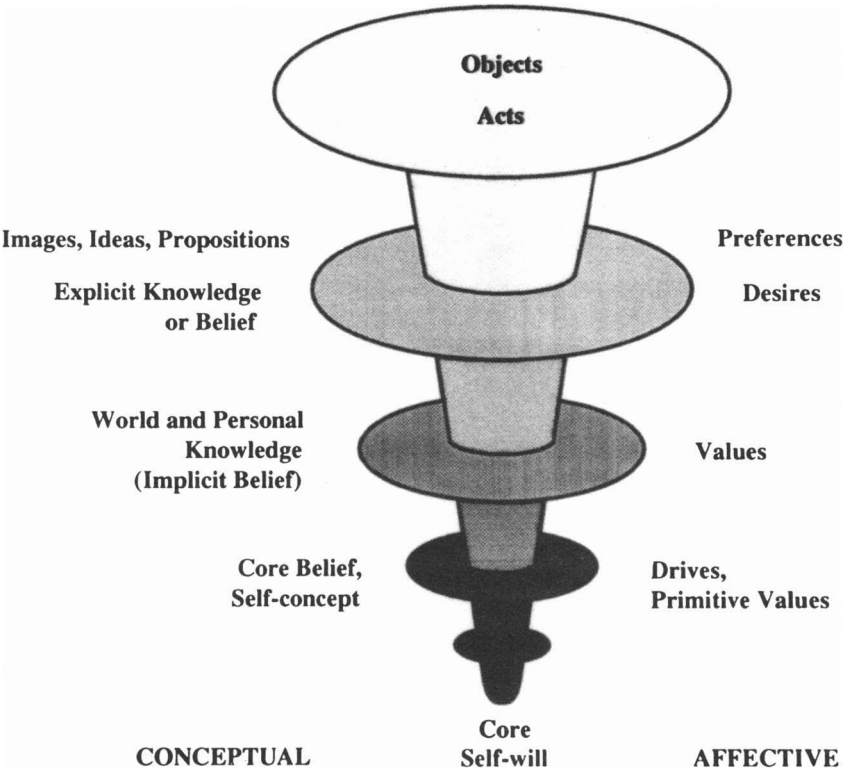

FIGURE 5.1. The mind–brain state fractionates from a core of the deep self and instinctual will through implicit belief and drive, to explicit belief and desire, to act and (mental and extrapersonal) object.

creates a subject.[5] The feeling that inheres in action is the basis of subjectivity. The primitive will announces the deep self in the commitment to action. Every mental state has its onset in primitive will. An action contrary to or "against one's will" points to a conflict at subsequent phases in becoming, not in the primitive will itself which is prior to choice and selection.

The primitive will develops outward in partial instincts and inward to the drives. There is an archaic line to instinctual repertoires; e.g., nest-building or courtship routines, that are repetitive and reproducable stimulus-dependent patterns directed to the surround, and another path to the drives. Instinctual routines have gestaltlike percepts. The all or none properties of the action and the tightly locked character of stimulus and response, reflect the potential of the core for an immediacy of realization that is diluted at later phases. Displacements from one instinct to another[6] reveal a common ancestry for the dif-

ferent patterns. Instinctlike machines such as insects lack subjectivity; there is no development into drives. An instinctual organism is an object, not a subject. In organisms with drive expression an active principle is at work that distinguishes the actions of a subject from what are physical movements in other animate or inanimate creatures. This active principle is primitive will.

The primitive will individuates to the drives which go on to replicate the two directions of commitment: an inner course to desire and subjectivity; an outer course to objects. The drives differ from the instincts in having less rigid conditions of satisfaction. There is a direction to a goal through an intrinsic, subject-building route. The subjectivity is the nucleus of an inner life. The transition to drive is mediated by archaic components of the limbic formation. The whole to part shift that characterizes this transition diminishes the force of instinctual will, prolongs the duration of the forming mental state, and permits a "delay" in the expression of the act and its object. The delay accompanies a greater specification or selectivity of behavior. Unlike instincts, the drives do not exhaust the core but elaborate it. The drives and their routes of individuation are the first step in the articulation of a self.

In the transition to drive there is a shift from whole to part or from potential to actual at every phase in becoming. Displacement phenomena reveal the potential implicit in *instinct*. Substitutions by way of satiation or obstruction, e.g., sexual drive passing into sleep, reveal the potential implicit in *drive*. This potential reappears in the conflict or choice between competing goals in *desire*.

The drives are the tributaries of instinct diverted from immediate discharge in stimulus-bound patterns of behavior. In mammals, there is an admixture of instinct and drive. Routines such as nest-building in a beaver combine instinct and drive in intricate goal-directed sequences. The delay in action is filled by a progression to a target. The inward growth enlarges the "mental space" of the organism. Actions no longer bear the stamp of rigid determination; they are more adaptive and can be adjusted to the circumstances.

A subject is a rudimentary self occasioned by the evanescent drive state. The growth of a subject to a *self* occurs as (precursors of) the perceived objects of drive satisfaction are individuated by learning to the imagined goals of desires. In the development of drive to desire, the subject is again enriched by the inward path. The shift to desire anticipates the many shadings of likes, dislikes, tastes, and preferences. Each of these attitudes incorporates a belief and a value that express a different facet of the self. Whether the self is interpreted as the cumulative expression or average of values expressed over time, or the distribution or priority of values in a given mental state, unlike a subject

that is its momentary drive state, a self is more than the preemptive value at a given moment.

The self of the desires is more complex than the subject of the drives. A subject is innately given and identified with the drive state of the moment. A self has to develop.[7] A bear or a human infant is more of a subject than a fish. This more is the intimation of a self in the nuance of behavior, i.e., the individuality of the animal. A bear has a "personality" that differs from that of other bears. In the chimpanzee, there is not only individuality but a primitive self-concept.

In the evolution of mind, the development from the core to the drives elaborates the inherited endowment of the individual. Morphogenetic process lays down instinct and drive. The environment provides the objects that are anticipated in the drives. Objects are satisfactions of innate process but do not modify the fundamental nature of the drive state. In desire, it is not the object but an instance in the object category that is decisive. There is a shift from an object category to a category member. In this respect, there is a greater dependence on a specific object.[8]

The relation between behaviors that are genetically determined and those that are acquired is not an overplating of innateness by learning but a resolution of endogenous form with the constraints of experience (p. 198). The view that learning is an *addition* to a genetic endowment is not supported by studies of early development.[9] Learning modulates ontogenetic growth. For example, the lack of connectivity for foveal vision in animals blinded at birth reflects an attenuation of morphogenetic process by the absence of visual stimulation. Sensation (learning) selectively enhances or retards innate process at one or another segment. The spatial and temporal distribution of external constraints on context–item shifts accounts for the functional specificity achieved by the connectivity. Since an extrinsic constraint becomes an intrinsic constraint in the ensuing state, the distinction of extrinsic and intrinsic, or the comparable relation between the acquired and the innate, is not useful in the problem of free will. After all, even identical twins have a sphere of individuality that is still not a window of free will.

CONSTRAINTS

Two types of constraints, one extrinsic, the others intrinsic, shape the mental state:

1. The first is the "sculpting" of the mental state by sensory input. Sensory input refers to afferent excitation delivered to the brain by the sensory nerves. A sensation is a physical stimulus that is the inferred

cause of a perception or *mental* object. A perception is a mental representation influenced at successive phases by sensory input. The sculpting is the basis of the specification of a potential into the final object world. The percept is what survives the elimination (inhibition) of other possible (potential) worlds. On this view, sensation is not ingredient in a perception but is an *extrinsic* modifier of perceptual form.

The generation and pruning of form in object perception is a microcosm of evolution. Every organic form struggles for existence. An object is not an exception. A percept is an adaptation to a niche in the physical world by the pruning of maladaptive form. Put differently, the percept is shaped to simulate what is "out there" in a graded realization of potential. A configuration is delimited at successive points by sensation to deposit a "real" object.

2. The second type of constraint can be divided into two subtypes: (2a) an intrinsic effect on the series of context–item shifts *within* the mental state, and (2b) an effect on the iteration *over* this series of the ensuing mental state. These constraints are independent of sensory input even though an animal that is sensory deprived cannot sustain a series of mental states. Indeed, if one could eliminate all sensory input during the transition from a state at T1 to the next at T2, the decay of T1 in T2 alone (constraint 2b) would prevent an exact replication.

The initial subtype (2a), involves the effect of a phase *within* the mental state on the succeeding phase, i.e., the effect as an earlier phase on a later phase in a single process of becoming. This effect is the basis of the whole to part transition, the "mechanism" through which configurations "travel" over phases in a mental state. The whole–part transform devolves over segments. Instead of multiple operations on different components there is a single operation at multiple phases.

The other subtype (2b) refers to the effects of the configural pattern of a given mental state, or the residue of that pattern, on the mental state that follows. If mental states overlap, as is likely, the activity of an initial state T1 will not have concluded before the next state T2 develops. If T2 unfolds over (out of) the activity of the immediately preceding state T1, the effect of the residual activity of T1 on the generation of the T2 state constitutes a type (2b) of intrinsic constraint. The occurrent (present) state is highly constrained by the preceding state. The baseline activity of the just prior state is the basis on which the occurrent state develops. Only a marginal deviation between two sequential states is possible, but some deviation is inevitable since a state can never be exactly replicated.

The deviation from one mental state to another that is temporally contiguous reflects this mix of internal and external constraints: the whole–part sequence within each state, the decay of prior states within

the present, and the sculpting of the present by occurrent sensations. These constraints guide the derivation of the mental state and influence the next mental state in the series that overlaps it.

In sum, the objective world constrains the mental state through a process of sensory parsing. This sculpting effect is imposed on the becoming so that objects emerge from contexts as inhibitory contrasts.[10] The parsing explains how an intrinsic transformation can create objects that are adapted to the physical surround. An occurrent state undergoes an obligatory passage over the track or residue of the immediately preceding state. An iteration over the prior state reactivates some features of those just traversed neural configurations. The binding or synaptic relations (strengths, refractoriness, etc.) that establish the configural properties of a population of nerve cells at a given moment are also the properties of the ensuing state, since the ensuing state is the prior state together with the qualitative effects of sculpting, deviation, and decay.

These constraints are exercised through graded patterns of momentary connectivity that underlie a given configuration. Since the constraint is *configural* it is the cell population not a particular synapse that is the relevant unit of change. This configural aspect has a qualitative basis that is not captured by accounts that involve thresholds, neurotransmitters or explanations at the cellular level. What it requires is a theory on the dynamics of neuronal populations as wavefront or fieldlike effects that are *intrinsically* generated.

Given the constraints on becoming, what are the causal properties of a constraint at its interface with developing content? If a content (configuration) develops through the inhibition of irrelevant or maladaptive form, in what sense is the surviving content *caused*? Since Hume, causes and constraints have usually been distinguished, and it remains unsettled whether constraints are causes. A constraint is a "passive" inducer, not a prior cause that prompts a future effect. A content that is an inhibitory contrast is what is left after rival possibilities are extinguished. The causal relation is one of inner necessity, not lack of options. Moreover, the constraint is simultaneous with the content it delimits. The exercise of the constraint and the appearance of the content are conjoint occurrences. Since the constraint forms the boundary of the content, the constraint is as much a part of the content as any other part. In a very real sense, the constraint and the content are different perspectives on the same event.

Constraints and Causes

A constraint is not a sieve that detains certain elements. In a whole–part shift (2b), the problem of whether the constraint is a cause is not the same problem as, for example, the causal nature of a content *C* that is

exposed by eliminating other contents, A, B, D, and E. In this example, the exposed content (C) is not changed by the removal of the other contents. In a context–item shift, what is eliminated is not yet actual. It is a possibility imminent in background potential. In the unfolding of the mental state, A, B, C, D, and E do not exist when a constraint is applied. The occurrent item (C) is what the constraint accounts for, while the items that fail to develop are inferred as the context in which the developing item is embedded.

With regard to *external* constraints (1), or sculpting, there is no exact correspondence between the item, e.g., a perceptual object, and the physical stimuli that shape its appearance. The stimuli delimit the individuation of the object but do not give the object directly. The sculpting that drives the process of becoming differs from sculpting per se, e.g., in marble, in that a form is not exactly "exposed" but realized. The contrast is essential for the content. In such a system, there is no copy from world to mind. Within mind, there is no copy from perception to memory. The constraint increases the probability of a development in a given direction. If one holds, as argued, that *external* constraints contribute nothing new to the actual content of cognition but accentuate existing patterns of mental process, a concept of free will in relation to extrinsic constraints (see below) does not confront the more profound issues lurking at the heart of will and action.

With regard to *intrinsic* constraints (2b), in which the state of activation at T-1 biases the properties of T2, the change from T1 to T-2 will necessarily be slight since the process tolerates a minimal departure from the immediately preceding state. Otherwise, there would be no coherence and continuity of mind across discrete moments. But in what sense does T1 *cause* T2? The activity of a neuron, for example the threshold for discharge, firing rate, or refractory period, are parameters that determine whether and what type of subsequent configuration results. In what sense are these parameters causal factors in the resultant configuration? Is the configuration (mental state) caused by the baseline of the dynamic that precedes it? A given configural effect, i.e., the occurrent brain–mind state, is constrained by the baseline of the just prior configuration. If we are to include baseline conditions as causal features of any occurrent event, we would have to include the resting state of the universe to describe the causal relations of every event that occurs. This is a form of indirect or passive "causation" that is more like causal persistence (p. 31).

To take an example, if someone is shot in the chest and dies, the bullet is a cause of death, or a cause of more proximate causes of death, such as hemorrhage. If the person survives being shot, the bullet is not a cause of the survival. If the person survives because of, say, an unusual

lung capacity, the capacity does not *cause* the survival the way the bullet causes the death or injury. The capacity enables the person to survive or is an account of why he or she survives, but it is not its cause. The person was already surviving when he or she was shot and continues to survive after the shooting. The lung capacity or the physiological correlate to which survival is attributed is a certain level of function, say 50 percent of normal, which is sufficient for survival even if it is not its cause.

To make the comparison clearer, if the lung capacity is 100 percent prior to the gunshot injury, this would represent the baseline at the moment of injury. In this respect, the baseline is like the brain state T1 just preceding the ensuing state T2. Let us assume that a reduction of this capacity to 50 percent is compatible with survival. Let us further assume a baseline capacity 75 percent of normal, not 100 percent and that a capacity after injury of 25 percent of normal is insufficient for survival. Would we then say a reduced capacity is the cause of death? In other words, is a baseline or resting state a cause? Every object in the universe persists in the context of a resting state, but this is not ordinarily construed as part of the cause of the persistence of the object or a change in its state. Indeed, the change is *from* the resting state. In these cases, whether a gunshot or a sequence of mental states, cessation of activity is a failure of causal persistence.

Is there causation over segments in a single becoming (2a)? In addition to the problem with causation in a whole–part shift, there is a difficulty as to the temporal properties of a becoming sequence. There is a precedence of phases, but is there causation from one phase to the next? The temporal seriation required by causation, past to present to future, does not exist until the becoming actualizes. There is a before and after, or earlier and later, but these temporal distinctions do not seem sufficient for causes and effects. One needs a present in relation to a past and future, and a present requires the actuality of a complete becoming that gives a whole object.[11]

In these examples of external and internal constraint, it is difficult to specify the locus of causation for events in a causal pair. External constraints influence biases to determine whether intrinsic constraints are muted or accentuated. Intrinsic constraints have a comparable action. Since these effects have in common an impact on context–item shifts, it is the nature of these shifts, i.e., whole–part relations, that is decisive with regard to whether the relations are causal.

A physical constraint such as a sieve or filter differs from a psychological constraint such as a gun to one's head. In a psychological constraint the *meaning* of the object and situation are crucial. Brain-damaged patients with *agnosia* may have an alteration in object

meaning. If the meaning of an object or a situation is altered, the situation should no longer exert the same constraint on action. Monkeys with bilateral temporal lobe ablations and a Klüver-Bucy syndrome[12] have a similar disturbance. They may calmly manipulate and mouth objects like snakes which normally terrify them. Is freedom enhanced if a snake or a gun loses its meaning, and no longer constrains behavior? A gun is an object in the world. Its meaning is up to the observer. One cannot assume that the awareness of a perceptual object is the cause of the ensuing behavior without a theory of how the world "gets in the head" or how meaning influences action.

The ubiquity of extrinsic constraints, from gravity to traffic lights to peer pressure, makes it reasonable to collapse a constraint to an effect on the configural properties of the mind–brain state. Extrinsic constraints exert their effects on endogenous (intrinsic) process. If the mental process is fundamentally *endogenous*, even if an extrinsic constraint *caused* an intrinsic constraint to be exerted, if the latter is not equally causal, a causal theory of behavior is unsupported. This might well be the case in a series of context–item transitions where the context is not definitely causal in the elicitation of the item; i.e., if whole to part shifts are emergent.

Disjunction of Will and Reason

The problem of sensory constraints on willing, i.e., the expression of the will in relation to the conditions of life, is a species of the more general problem of the effects of experience or learning on brain process. Perceiving and learning have in common the fact that both require a theory of the relation between mind and physical sensation. The difference between perception and learning is the depth of the object that is sampled. Memory relates to earlier, perception later phases in the same object. The mind–world relation pervades every aspect of psychological thought, from the obvious quesion of how a (raw) sensation becomes a formed and meaningful percept, to the less obvious relation between will and reason. In all cases, the problem devolves to that of sensory input and percept formation. Conventional thinking has it that percepts are generated from sensory elements by a compilation of informational bits. For microgenetic theory, a percept is an endogenous image shaped by sensation to a model of "reality."

In the case of will and reason, the convention is that primitive will is an endogenous capacity suppressed by rational thought. Reason is an outcome of a learning process that is internalized to modify the drives. Microgenetic theory treats will like perception, as an endogenous process that undergoes fractionation into action and the desires. The neglect

of the endogenous in mental process leads to a variety of errors that derive from mistaken assumptions based on everyday experience; e.g., the postulation that objects are *projected* into the world, or the opposition of reason and will (drive). In the case of will and reason, experience reinforces the belief in their incompatibility, that reason is a product of culture, will an innate disposition, or that the will is inhibited or modified as a response to external censure and instruction.

The opposition of will and reason is interpreted by many as a conflict between an innate or endogenous process and an exogenously derived constraint. The primitive will, like the "id" of psychoanalytic writings, is a force or energy inhibited by higher voluntary centers. The concept of a disjunction of will and reason is embedded in ordinary language, e.g., in the description of willed (nonautomatic) acts that are contrary to reason and therefore involuntary, as when one says, "The Devil made me do it." One can substitute for the Devil the effects of an abusive childhood, hypnosis, persuasion, etc. Acts that are construed as voluntary but unwilled, such as the sense of duty, or obediance, can be explained in a similar way.

Conversely, an action motivated by an obsession or intense desire that is "against one's will" is an action that "defies reason," such as a crime of passion. Such an action is construed as irrational or involuntary. In states of conflict between the will and the self, or between desire and reason, reason is understood as an acquisition through learning (extrinsic constraints) of a conceptual ability that becomes an ***intrinsic*** constraint to modify an innate faculty of will. In such cases it might be argued that reason or will *alone* can effect an action. If will and reason are distinct they might be subject to independent coercion. Reason can be overcome by false beliefs, submission to authority, seduction, brainwashing, hypnosis, etc.; will can be overcome by mental illness, drug effects, compulsions, and so on. When reason is overcome by will or if there is coercion of the will, the rational self "knows better" but is powerless to act in a rational manner.

Such anecdotes comprise our commonsense vocabulary of will and action. However, compulsions, coercions, disjunctions, and the like are folk descriptions of interactive entities in behavior, not explanations of covert mechanisms. An internal or external constraint is not, for the reasons discussed, a clearly "causal" explanation. If a constraint is not a cause, theories of freedom based on constraints are not causal theories.

Freedom and Coercion

Many philosophers argue that freedom is simply the absence of coercion, defined as a restriction of some sort imposed on an agent. For example,

Schlick[13] has written that "freedom means the opposite of compulsion; a man is *free* if he does not act under *compulsion*, and he is compelled or unfree when he is hindered from without in the realization of his natural desires." This account of freedom appeals to common sense. It also seems neccessary in some sense for a theory of moral responsibility.

However, there are many difficulties with this idea. It hardly bears repeating that there is no procedure to distinguish a hindrance or compulsion from any event in the experiential history of the individual up to a present action, including the internal (e.g., hormonal) and external (environmental) context in which the act occurs. There are no *fundamental* "causal" properties to distinguish the "no" of a parent, the pressures of habit, the fear of reprisal, or a gun to one's head. How does a "no" differ from a "yes" or a "maybe"? How does a gun differ from a threat? A command? A plea?

There is also a circularity in the argument. A constraint intervenes according to its strength, but its strength is inferred from the control it exerts. Constraints limit or expand the scope of action, some more than others, but it is futile to specify that point where a constraint eliminates freedom or, more fundamentally, the role of constraints in the generation of action. If the complex of constraints on an agent accounts for the choice of an action, even if the agent believes he or she is acting freely, the action is caused by that complex and is not "freely" decided. Since the constraints on action are so pervasive compared to the narrow range of (possible) freedoms, freedom on this view is at best a sporadic addition to causal nature.

Ayer[14] follows Schlick when he writes, if I "am compelled by another person to do something 'against my will' . . . I should not be said to be acting freely." Again, what is meant by compelled? Does the external compulsion *directly* control the will or does it *appeal* to values in the self-concept? A gun to the head is a challenge to one's courage, skill, and desire for self-preservation. The gun limits the range of possible options but the choice is still governed by what is important to the agent, for example whether to survive or act bravely. Indeed, in that when a gun is pressed to one's head it seems *irrational* to act other than instructed, the agent who acquiesces is acting according to reason. If freedom is equated with rational action, the agent who acquiesces is acting freely.

Ayer argues that a constraint is a cause on the grounds that an action would not have occurred without it. For example, I would not have handed over my money were the gun not pointed at my head. Nor would I have dined at a certain restaurant if I didn't like their shrimp salad. Is the shrimp salad a cause, like the gun? The "cause" of dining at the restaurant is a desire for the shrimp salad. The "cause" of

acquiescence to an assailant is the desire to escape unharmed. In both instances, it is a desire or value that motivates the action. In my view, an argument to causation from such counterfactuals is vacuous.

Whatever action a rational agent chooses in such circumstances, whether to submit, to fight, to run, etc., is motivated by his self-concept, the values that permeate his beliefs and desires, and only indirectly by the situation. The gun sharpens the conflict in values that is the basis of any free action. Constraints bring to the fore the values, i.e., the makings of character, that are tacit in every conscious choice. Indeed, the anxiety and indecision, the struggle over what to do, the awareness of the difficulty, whether to submit, to fight, or to run, are all options without which there is no free will, or without which there is no awareness of free will or lack of free will, which is probably the same as an absence of free will. Conflict is a litmus test for free will whether or not a freely willed act leads to freedom "in the world."

If freedom is defined in terms of constraints, and an agent is never free of constraints, freedom is always a matter of degree, a quantity, or a function of the probability of deviation from the expected. With a gun to my head, I am still free to contemplate my fate, think of options, or wiggle my toes. Freedom then depends on the situation, not on the agent. It is a property of the distribution of constraints in a given setting in which a particular act occurs. But this is an odd philosophy. An account of free action should not be hostage to the availability of occasions in the world for its implementation. One should say that extrinsic constraints account, trivially, for the difference between free and unfree choices from the standpoint of *freedom*, while intrinsic constraints are *the* issue as to whether any choice is free from the standpoint of *free will*.

Will and Causation

The perception of causation develops out of "sensorimotor" function at a similar age in all children. Piaget[15] found that visual causation requires initial contact between objects, and follows on the perception of causation in tactile and kinesthetic systems. He thought that the perception of causation originated in action. This line of thought was anticipated by Guyau,[16] who believed that actions such as reaching were the seed of the idea of the future. The concept of the future and the direction from a (present) cause to a (future) effect is based on the action experience. The feeling of agency gives rise to the perception of causal relations.

Hume[17] maintained that spatiotemporal relations between objects were insufficient for the idea of causation, which depends on the "ne-

cessity" or power in these relations. The power in the "connection" between objects is extended to objects from the mind. Power is a mental addition that gives life to mere association. But to allocate part of causation to the world and part to the mind is to confess the inadequacy of the association concept. Power or necessity is the momentum of process in mind and nature, while spatiotemporal relations are the private and public worlds that process lays down. Cause and effect in the mind have a necessary "association," since they arise in the actualization of a mental state where the cause is preparatory and the effect terminal.

In mental process, i.e., agent causation, a cause is private; an effect is either private or public. In physical process, i.e., object causation, an effect is public and its cause is inferred. In nature we observe effects. Every cause is an effect of a prior cause. The cause is in the past and inferred from the effect. If one begins with an event such as lightning, and labels it the cause, the effect, a fire, will always be a probability. Or one can begin with the effect, the fire, and work backward to the cause. There is greater certainty for past than future states. But the certainty of a cause is only achieved when the effect is already known. In mind the certainty of the cause is obtained at the expense of the inevitability of the effect. An agent can have a "change of mind." The possibility of reconsideration distinguishes the options available to the agent from the inexorability of physical causation.

In the progression to an extrapersonal object, there is a loss of privacy. The effort of will in the imagination, the specification of will in speech and action, the direction of will to objects, represent a progression from potential to fact, from self to world, from privacy and agency to publicity and commitment. Self and will are left behind as their "effects" actualize. Causation in becoming is a progression from the concealed to the actual. In science, causation is from one object surface to another;[18] causes and effects are not fundamentally different. In becoming, cause and effect are states of the same object.

Goals, Reasons, Plans

Agency differs from physical causation, inter alia, in the feeling of the present being drawn into the future, as when a goal seems to be the cause of an action. The ends not only can justify the means, but seem to cause the very actions that are necessary for their achievement. This is purposefulness, not backward causation. Physical change does not look beyond the nexus of the moment. There are no *goals* in the material world.

A goal such as to strike a match this moment or go to France next year seems to cause a series of actions that lead to its attainment. The

goal is the object (idea) of an intention (e.g., desire). The intention guides the act toward the goal. The goal seems to require and, in this way, determine a series of actions that lead to its satisfaction. The goal is both the starting point and the destination of the act. Though the goal is an objective toward which the action is striving, it is still an occurrent concept or proposition in which an incipient action is embedded.[19] Whether the goal is for an immediate action, such as striking a match, or for an action in the distant future, such as going to France next year, the goal is still *an idea in the present*. The idea is for a future action. It is not a future idea that attracts an action toward it but a present idea that gives way to a subsequent action. The direction of "goal-oriented" action is not from goal to action, or even from action to goal, but from a goal (object) as an idea in the mind to an action as an idea in the mind (or world).

With a commitment to act there is a necessary relation to the goal. I decide to lift my hand to throw a ball and feel a "connection" between the willing, the lifting, the throwing, and the target. The potential in willing gives way to a decision, and the decision gives way to an action. Once the goal is chosen and initiation is committed, the action and its movements follow more or less automatically.

The components of the action or their anticipatory forms are submerged in the action "plan." A plan is an action that is preset for a desired outcome (object). The presetting is prior to a conscious description of the plan as a script for a forthcoming act. The plan seems to be the driving force of the action. When I throw a ball, I am not conscious of calling up a plan or a schema nor do I access the individual movements into awareness. Indeed, were I aware of each movement, the action would be hesitant and less skillful.

In a sense, the presetting of an action is an outcome toward which the action is directed. The action unfolds or discharges to satisfy the plan, which is both a cause (of the unfolding or discharge) and an outcome or realization. In this sense the plan is the goal of the action. The problem with goals or plans as determinants of actions, however, is that imagined outcomes in the future, e.g., a trip to France, can influence actions in the present, buying a ticket. Moreover, the "association" between effect and cause in this situation has to bridge many gaps in contiguity and over a prolonged period of time.

Ordinarily, the will selects goals not plans. The "what" of an act is more emphatic than the "how." Certainly this is true for simple actions such as blowing out a match. In more complex acts, it is possible to construct a scenario that comprises a sequence of acts leading from an idea, such as a trip to France, to the final realization. Perhaps there are people who conduct their life on the basis of such scenarios, or an

assessment of alternatives in the course of complex action, or even a decision tree at each step in a sequence. However, the ability to analyze a program in advance of an action, or elaborate a script, does not obligate that the action is a product of the analysis. Scripts are perceptual or verbal images that are generated out of the same concepts that give rise to the action. A script is not a precursor but a surrogate action.

Moreover, each step in a script is a composite of part acts that reduce to the same problem as, how does the will blow out a match? For example, one can scrutinize and consciously direct an action such as pursing the lips before blowing out a match but each movement in the act of pursing is then a kernel of automaticity where before the entire movement was automatic. There is a preset core in any action no matter how individuated. This implies that the conscious plan is not the plan by which the action is guided.

This way of thinking has implications for our concept of the "causal surface" between willing and acting. If every action is a composite of constituent acts, and if each act is a hierarchy of nested automatisms, the will could not effect every (any?) constituent. In fact, some constituents of a volition cannot be duplicated by voluntary effort. For example, I cannot purse part of one lip. If I can control one element in an action, that element is no longer an element but the whole volition. Further, the element contains subordinate components that are themselves refractory to voluntary control. For the will to induce an action, it should be able to access the plan of the action even if planning is not obligatory in implementation. The will, it seems, must incorporate the goal as an object of an intention, even if the goal does not drive the action, and it must activate the conceptual base (plan) of the action, even if the constituents of the action fall into place automatically as it materializes.

These considerations might suggest that agent "causation" involves the junction of the forward edge of the will with the rear edge of the action (plan). According to the common account of the will, the self or person has the will to act which then activates, implements or elaborates a goal and/or a plan of action.[20] Can a juncture of the will with the action or action plan be identified? If the plan engenders the action, is the plan the locus of the will? If a plan is preset and nonconscious, how does a (conscious) will select or activate it to get the action going? Would this necessitate two forms of willing, one a conscious decision and/or initiation and another that is a nonconscious trigger?

The goal or purpose of an action is very close to being the same as its intention. I intend (have the goal, purpose) to blow out a match. The goal or intention seems to be present before a conscious effort of the will. The will accompanies the concept (plan) of the action in relation to which the action is itself an outcome. Each of these preliminary

steps, the deep self-will, the intention, the plan, are events in the present that have a direction to the future. Because of this future direction, the events seem prior to the action and in a causal relation to it.

Many related states are like causes of an action, including beliefs, reasons, motives, as well, of course, as affects such as wants, wishes, and desires. Each of these states might be the cause of an action. I act in such a way *because* I believe, desire, reason etc. But if an action is set in motion by such a variety of causes, or if the will is provoked to act by such a multitude of states, how is it possible to isolate a causal surface?

These states are expressions of the self at or before the moment of an action. The self acts and part of its action, i.e., another mode of action, is to give reasons, goals, and beliefs that explain or justify why or how the action occurred or may occur. The intention is a surrogate action that precedes and is replaced by a motor action though the precedence of the intention does not necessarily entail a causal relation to the action.

In studies of pathological behavior, explanations by the patient of abnormal actions are often given with conviction, even if the beliefs that underlie the conviction are inaccurate (guesses, fabrications). Such explanations are discounted by a normal observer and are attributed to the pathological condition. Similarly in normal people, one should be suspicious of introspective reports as explanations of behavior, even if they conform to a shared rationality. The various psychological states that seem to cause an action are in some sense themselves "actions." A conscious belief or desire actualizes the concept underlying an action even if it is not instigatory in the generation of a motor act.

Environmental constraints delimit an action but do not enter its structure. An external stimulus to action, e.g., jumping aside to avoid an oncoming car, is not the cause of the action. The cause could as well be the awareness of danger or the desire for self-preservation. If I am depressed, suicidal, inattentive or preoccupied, or hopeful the car will stop, I may act differently. These states, not the car, become the reasons for whatever action is chosen. These reasons are descriptions of attitudes, so ultimately the self elects to act.

The self actualizes to conscious content. A reason given to motivate an action is a self giving an accounting of itself, and, through the accounting, fractionating to a proposition and becoming "more" conscious. Reasons, goals, beliefs are different shadings the self-concept takes on as it individuates. Desires, needs, wants are different tonalities the will assumes as it individuates. If the reason for an action is a proposition derived from the self-concept, and if the action is derived from the self-concept, the reason cannot be the cause of the action since both

are products of the same core. Reasons have explanatory power but lack causal potency.

Indeed, a choice can be voluntary without a reason. I can decide to lift my left hand instead of my right hand, yet not be able to give a reason for the choice. This is a type of impulsive rather than deliberative action, yet it is a freely willed one. I can lift my finger now or in thirty seconds. Could I give a reason for the timing? There is an inability to give a reason for a simple action that is freely willed. This is because reasons are surrogate or supplementary actions that do not occur where the concept guiding the action is impoverished.

Does Free Will Require Determinism?

Free will is often held to be incompatible with determinism. There is incompatibility if free will is fully identical to physical brain events and such events are causal. If consciousness or agency is nonphysical or emergent, different "laws" might obtain for the physical and the mental. Free will would then be something new in the world. Yet free will seems to require determinism or universal causation[21] in at least two ways:

1. For conscious intervention in a causal sequence. The sequence might be a chain of physical events (e.g., stopping a billiard ball), or a chain of behavioral states (e.g., restraining a child), or an effect of the mind on the will of another person (e.g., hypnosis or giving advice); but from the agent's standpoint the intervention can influence an outcome that is more or less certain. If the chain of world events is determined, why intervene unless agent control is an exception to determinism.

This, of course begs, the question of whether the intervention is also determined. Human action ought to be bound by the same laws as physical nature. From the subject's point of view, however, physical causation is necessary as a contrast and rationale for agency. Indeed, the fact that both the idea of causation and the idea of free will originate in the mind may be a clue to the nature of the mental process from which these conflicting yet interdependent ideas emanate.

2. For causal interaction between consciousness and physical events. Free will enjoins causation in assuming that mind causes something "nonmental" to occur (e.g., a limb movement), but substitutes consciousness or will for a physical event as a prior cause. The interface between consciousness and action is felt as causal. My limb will move where and when I direct (cause) it to move. The self-will is the causal agent and the action is the effect.

Free will entails agent causation as a justification and an explanation: a justification because causation in the world is required for the assertion of free will in the mind, i.e., the will is free but the effect is

caused; and an explanation because free will is an alternative to determination as an interpretation of the nature of conscious mental activity. The fact that all physical causes are effects of prior causes does not obtain in the interpretation of the will as a cause. The will is not felt as the effect of a prior state of will but as a truly free capacity.[22]

Does Free Will Require Indeterminacy?

Indeterminacy implies that change is probabalistic but it could also pertain to change that is fundamentally novel. If all change is probabalistic, free will amounts to an adjustment of the probability of a given outcome. When I blow out a match I alter the probability the match will spontaneously extinguish. This may be the case but as an account of free will it is counter-intuitive. If extrinsic factors are excluded, such as a sudden gust of wind, the probability of extinction of the match is altered from close to 0 to close to one. How can the weighting of probability by an agent from almost zero to almost one be explained?

A change in probability is a description, not an explanation that resonates with private experience. I believe I directly cause the extinction of the lit match. Every action has some probability of occurrence without the intervention of an agent, like the probability, however small, that a chimpanzee could randomly type a Shakespeare play. But agents do not interpret their actions as probabilistic.

For the same reason, there is an incompatibility between free will and novelty since the will does not induce an unpredictable event. Novelty in the world is like free will in the mind. Free will is the ability to induce novel change. If change in the world is novel rather than causal, free will would bring (causal) order to an unpredictable world. Free will might involve the definiteness of an outcome that is otherwise uncertain. Choice makes a situation explicit.

Novelty is possible in the absence of mind but free will requires consciousness and choice. If novelty in change is independent of willing, free will could be an extension of novelty to human experience. If the will is truly free, determinism and causation cannot be "laws of nature"; there must be essential novelty or indeterminacy. If the will is not free, change could still be novel and free will could be illusory. Free will would be an awareness of the possibilities that novelty affords, not the initiation of a genuinely novel act.

Thus it seems either the world is determinate and free will is creative, or the world is creative and free will is determinate. If consciousness and the world are expressions of novelty, what is there to choose between novelty in the world and novelty in consciousness? Novel change in the world, e.g., a miracle, is explained by the action of an

external agent (God), while novel change induced by consciousness is explained by the action of an *internal* agent (the self).

If all change is novel, there is nothing to distinguish the change associated with free will from any other type of change. Free will would then be the self's interpretation of inevitable novelty as a product of its own agency. However, the freedom from determinism would be at the mercy of continuous novelty.

The issue devolves on choice and its interpretation as an opportunity for novel change. If conscious choice is determined by prior events, and the decision to act is a necessary conclusion of a causal chain in mind, consciousness is like an object interposed as a node in a sequence. In the decision to blow out a match and the exhalation that follows, consciousness plays the same role as a gust of wind. The difference is the decision to intervene and whether this decision obeys causal "laws." As discussed above (p. 35), free will depends on a strict interpretation of causation and a loose interpretation of choice. There is no choice in nature. This is why, finally, the idea of freedom rests on the analysis of choice and the nature of the self that chooses. In this respect, one must agree with Hannay that "the place to look for the origins of the kind of control that human brains exercise is not tropism . . . but something constituting a hiatus in the input–output series (that) has the status of a genuine explanatory node."[23] The search for this "explanatory node" is the task of the following chapter.

NOTES

1. The distinction of innate and acquired is by no means a simple matter. There is a microenvironment that influences the action of the genes. Contextual effects are present at every stage in development. It is a question of degree of externality at a given stage or what constitutes the external world for a given element. In microgenetic theory, the innate refers to patterns, such as morphogenesis (Chapter 11), that develop early and are predominantly endogenous, while the acquired refers to the shaping or modulation of innate process by experience.
2. There is a full discussion of the deep or core self in Brown, *Self and Process*, 61ff. The concept is of a core self that distributes into acts and objects. The core undergoes slow change, the surface, rapid transformation. On a prior distinction of a matrix (essential or core) self that endures and a focal or peripheral self that changes, see D. Parker, *Experience and Substance* (Ann Arbor: University of Michigan Press, 1941).
3. V. Bourke, *Will in Western Thought* (New York: Sheed and Ward, 1964).
4. Though see W. James, *Principles of Psychology* (New York: Holt, 1890).
5. F. Schopenhauer, *The World as Will and Idea*, vol. 3, ed. R. Haldane and J. Kemp (London, 1907–1909). See B. Brewer, "Self-Location and Agency," *Mind* 101 (1991): 19–34.
6. K. Lorenz, *Studies in Animal and Human Behavior*, vol. 1 (London: Methuen, 1970).

7. Henri Bergson, *Creative Evolution* (London: Macmillan, 1911), maintained that instinct concerns actions on *things* while the contents of intelligence are *relations*. The microgenetic equivalent of this distinction is that, given an organism capable of intelligent acts, "things" become internalized as attenuated objects (preobjects) *with* their relations. Actions and objects withdraw to or do not individuate from their relational background.
8. The argument has been made that will is for an action while desire is for objects; e.g., T. Reid, *Essay on the Powers of the Human Mind*, vol. 3 (Edinburgh: Bell and Bradfute, 1819), 75. This distinction corresponds with the *derivation* of will (drive) to desire and the parallel derivation in perception from unconscious beliefs to concepts and objects in awareness.
9. See S. Carey and R. Gelman, eds., *The Epigenesis of Mind* (Hillsdale, NJ: Erlbaum, 1991).
10. The hypothesis that an object is an *inhibitory* contrast derives from clinical observations rather than physiological research (see Brown, *Life of the Mind*, 257; *Self and Process*, 10,55), though there are points of contact with studies in Gestalt psychology.
11. Studies in the philosophy of time have shown a comparable distinction, e.g., J. McTaggart, *The Nature of Existence* (Cambridge, U.K.: Cambridge University Press, 1927).
12. For references, see *Life of the Mind*, 183.
13. M. Schlick, "Causality in Everyday Life and Recent Science," in *Readings in Philosophical Analysis*, ed. H. Feigl and W. Sellars (NY: J. Earman).
14. A. Ayer, in *Free Will*, ed. G. Watson (Oxford: Oxford University Press, 1954), 19.
15. J. Piaget, *The Mechanisms of Perception* (London: Routledge and Kegan Paul, 1969).
16. J.-M. Guyau, *Guyau and the Idea of Time*, trans. J. Michon (Amsterdam: North Holland, 1988); originally published as *La genese de l'idée de temps* (Paris: Felix Alcan, 1980). J. R. Lucas, *The Freedom of Will* (Oxford: Clarendon, 1970), 62, has written, "the notion of cause is based not, as Hume thought, on our *observing* constant conjunctions of events, but on our being able *to make things happen*."
17. D. Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge (Oxford: Clarendon Press, 1888).
18. C. Ducasse, *Truth, Knowledge and Causation* (London, 1968). There are alternative views, e.g., David Bohm's account of an implicate subtext out of which explicate order emerges (D. Bohm and B. Hiley, *The Undivided Universe* [London: Routledge, 1993]).
19. Some philosophers consider goals future states that do not yet exist and cannot explain present actions. See the critique of this view in N. Care and C. Landesman, eds. *Readings in the Theory of Action* (Bloomington: Indiana University Press, 1968).
20. According to J. Velleman, "What Happens When Someone Acts?" in *Perspectives on Moral Responsibility*, ed. J. Fischer and M. Ravizza (Ithaca, NY: Cornell University Press, 1993), the accepted story is that a desire and a belief jointly cause an intention which then moves the agent's body.
21. See R. Hobart, "Free Will as Involving Determination and Inconceivable Without It," *Mind* 43(1934): 1–27; B. Berofsky, ed., *Free Will and Determinism* (New York: Harper and Row, 1966).
22. Spinoza wrote that human freedom "consists in nought but that men are conscious of their own will, while ignorant of the causes which have induced it."
23. A. Hannay, *Human Consciousness* (London: Routledge, 1990), 132.

CHAPTER 6

Autonomy and Agent Causation

ARGUMENT: *Agent causation differs from event causation in the feeling of spontaneity and the delay and demarcation between cause and effect. Agent autonomy may depend on a (virtual) duration that spans developing actions. The feeling of spontaneity is related to the precedence of the self, the depth (past) to surface (present) transition, and potentiality prior to actuality. Delay and demarcation reflect state decay and/or revival between intentions and actions. An intentional state is replaced over an intervening series with a final depletion of conceptual content in motility. Agent causation corresponds with persistence in event causation, in that a subject undergoing minimal conceptual shift is construed as causal across the boundaries of an interval.*

* * *

The ability of the self to cause an action (or retrieve a memory or introspect an idea or image) is the problem of agent causation or control, the presumably causal relation between an intentional (agentive) self and its actions. Agent causation does not require that an agent is autonomous, but autonomy is necessary for free will. Autonomy refers to the fact that a self, or agent, is not fully explicable in terms of prior or occurrent mental events nor fully reducible to events in the brain.[1]

AUTONOMY

Events prior to a state of willing, whether they begin with the birth of the individual or the inception of the universe, can be interpreted in

line with object or event causation. Experience and inheritance determine one's personality and character. For autonomy, the self must achieve spatiotemporal independence from its "causal" history, from the immediate environment, and from the accompanying brain state.[2]

A fixed causal history is assumed to be incompatible with autonomy. Causation is incompatible with free will but necessary in some sense for agency. For the autonomy of the agent, the causal ancestry of the self must rupture at one, the prior or agentive, side of an action and then, for agent control, be reasserted at the other, instigatory side. Autonomy cuts both ways. If the self must dissociate from the past to be autonomous, it must reconnect with the immediate future to induce an action. The postulation of an emergent self in the act of choosing, i.e., a self emergent from its causal history or occurrent brain process, does not resolve the difficulty. If the self is emergent, one still needs to explain the connectedness from past to present and the transition from self to act. The emergence necessary for autonomy must be resolved with the cumulative history of the self and its continuity with action.

Regardless of whether change is emergent, causal, or probabilistic, the difficulty seems insurmountable. Causation requires something like emergence for autonomy, emergence requires something like causation for agent control. A dissociation from a causal past, or the continuum of dissociation that marks an emergent or probabilistic series, entails a dissociation from a causal, emergent, or probabilistic future that is incompatible with a self that can "cause" or implement an action. These different theories of change all seem to undermine the capacity of a self to freely choose and act.

Kane[3] discussed the possibility of a special cause as a species of non-occurrent causation in which the self in an act of choice is not the effect of antecedent or occurrent states. Nonoccurrent causation is a form of causal indeterminism, where the past is not causal in relation to the self in a state of choice, and where the self of that state controls its actions and chooses among alternatives. In the momentary freedom from its causal past, this self is reminiscent of the noumenal self of Kant, a cause outside of time. This is, Kane admits, a mysterious idea. A self that dissociates from physical brain process could not influence the putative (brain) causes of mental states. Dissociation in time is an indeterminism of the moment that permits autonomy of the self (from its ancestry and the occasion) but excludes agent control.

Such confusions have made free will a sinkhole of complexity for generations of scholars. An intrapersonal or subjectivist account of free will that entails a truly autonomous self is incompatible with universal causation. However, even a refutation of determinism or a theory of change as emergent, would not guarantee free will because of the prob-

lem of dissociation. What is required is a different concept of change in the mind and perhaps in the world as well. I believe that a concept of *emergence in becoming, in which a virtual duration is extracted from the appearance and extinction of actual objects*, offers the possibility of a resolution of autonomy with agent control. On this view, emergence is not a deviance from causal certainty along a "horizontal" axis of change, i.e., from past to present to future, but is linked to the momentary actualization of the present state.

Some Clinical Problems for Autonomy

We distinguish cause and effect in the interaction of physical objects and the body. We describe the relation between a mind and a body in causal terms. The mind moves the body, e.g., an arm. The body, through the brain, affects the mind. An illness of the body can induce a change in behavior, as when chemical agents produce a disturbance of thought or action. When an illness affects a person's mind, we say the illness *causes* a disturbance in thinking. The illness of psychosis causes a thought disorder. An epileptic illness causes a bodily movement. In such cases, the illness is said to cause the behavior directly or, it might be said, to influence the will or self which then causes the behavior, e.g., a biological depression that affects the self-concept and secondarily alters behavior.

When the will or self is the target of a physiological disorder, as perhaps in depression or psychosis, the disorder is interpreted as acting on the self as a mental object. Psychotics and epileptics may disavow a feeling of agency for their thoughts and actions. Similar phenomena may occur in hypnotic states or intoxications. In the symptom of thought control the psychotic feels his mind is an object for other people's thoughts. The epileptic may observe his arm to move involuntarily. When a physical or mental illness affects a thought, an act, or the self-will, the self feels its mind or behavior is a victim of the illness. We tend to share this view and do not hold the person accountable for his acts.

When behavior is an effect of an external cause that operates through the will, it is not always clear what are the proximate causes of the behavior or its substrate. When an epileptic lifts an arm during a seizure, we say that an excitation of the brain has an effect on a motor pathway. If the epileptic commits a crime during a seizure or carries out a complex act in stages, where does the excitation act? Would we say the excitation causes a complex plan to occur? Does it cause the self to enact such a plan? Is the cause of the altered behavior an action dissociated from the will? A will that is degraded? There are cases where the will seems to induce a behavior but the self is not responsible. We

say, "He acted against his will," or "He didn't know what he was doing." It seems natural to distinguish the self and the will in this way.

These various examples illustrate some of the difficulties from a clinical standpoint with the concept of autonomy. The feeling of an autonomous self underlies the assumption that autonomy can be perverted by a variety of factors, from a deviant maturational history to a malevolent influence in adulthood to an acquired pathology. Moreover, we also assume that the perturbation of the self is a matter of degree. This is the case whether it is due to an alteration of the brain or a psychological disorder.

What are we to make of such observations? Is genuine autonomy corrupted by a psychological or neurological affection of the self or does the corruption of the self only magnify the existing constraints on self control? That is, does an acquired disorder cause a lack of autonomy or does it bring to our attention the lack of autonomy that is implicit in everyday life, and therefore unnoticed? Is what is lost in pathological cases the illusion of autonomy? The presence of a continuum from what seems to be full autonomy to a lack of self-control suggests that intrinsic constraints are present even with a feeling of full autonomy, but that these constraints are relatively inapparent in the normal state. Thus, while autonomy is required for free will, it is not clear how the concept can be supported, nor, if one assumes autonomy, is it evident in what sense a self can be said to be autonomous.

As noted, the pathological material suggests a distinction between the self and the will, e.g., the will as a motive power of the self. In a similar vein, Frankfurt[4] has argued that a disconnection or lack of identification between an agent and a volition is possible. This seems to obligate a model of the self in which the self and its will (or volition) are distinguished. In my view, the self and the will are complementary aspects of the same entity. The primitive will and the deep self are the action (affective) and perceptual (conceptual) properties of a common entity. The appearance of a dissociation, e.g., a self unable to will, or the overcoming of the will with a self that is presumably unchanged, can be attributed to the many uses of the term *will,* from instinctual drive to conscious initiation. For example, a state of indecision can be interpreted as a failure or weakness not of the self, which may apprehend the best course to follow, but of the will which cannot carry it out, even though deeper phases in the generation of the will, such as drive, are unaffected.

Past and Present Causation

Most of us would agree that the self-will can degrade or become weak (I would prefer to say *dissipate* into a multiplicity of derived elements)

or that it can be altered or affected by extrinsic factors. But we recoil at the idea that the history of the individual, i.e., the life experience, gives the state of will at the moment of a decision. The self-will can be driven by intrinsic or external events, from passion to brainwashing, but in a present action, the will is not felt as an outcome of its own history. The will seems vulnerable to all sorts of influences but not the effect of its own ancestry. This difference in feeling is part of the experience of *spontaneity*; i.e., a will that is felt to be free and directed toward the future, not the outcome of a causal past. Spontaneity is the feeling that the self departs from a causal sequence that includes prior states of willed action to initiate a novel act as a free agent. This feeling is an important part of the conviction of freedom. The will is not apprehended as "caused" by the past, but is felt as a "cause" of what is to follow. The bias in agent causation from present to future seems to overwhelm the sequence of event causation from past to present.

Moreover, as noted, the apparently *causal* effect of prior events on the present is not distinguished as to whether the event included a freely willed act. The effect of a past dream or an episode of explosive rage on the present may be trivial or profound but the effect does not differ, intuitively, from that of a prior volition. I may reflect that my past choice to get in a car could have been otherwise, but given that I was driving the car, the accident that occurred, say, when the brakes failed, was unavoidable. However, the choice to get in the car and the occurrence of the accident are equivalent events in my causal history. The fact that a prior act was voluntary does not imply a break in a causal chain leading to the present. Indeed, I may feel that given the circumstances my choice to get in the car was as inevitable as the accident.

Whether a past act was involuntary, forced, or deliberate, apart from the events to which that act leads, is not decisive in the constitution of the present. That is, events and decisions leave their traces on the personality but not in relation to their voluntariness. Willed and unwilled acts are part of a (causal) history that constitutes the life experience, but their prominence depends on their impact on the subsequent career of the self-will, not on whether they were freely willed or automatic.

More precisely, at a point in the future, a willed action in the present becomes nested in a prior causal series. Every present act is shortly part of the past. The cluster of the willed act (the will and, however conceived, its nexus to action) is an event-whole surrounded by causes and effects. For example, the *decision* to get in a car that subsequently is involved in an accident, though prior to the accident, and prior to the many decisions made en route, e.g., turning on the ignition, releasing the handbrake, stepping on the gas, etc., is subsequently fused with other intervening events, some voluntary, some automatic, some intru-

sive, as an event whole. The combination of decisions and acts or outcomes is embedded in a sequence of such events. From a future perspective there is a lack of distinction between voluntary and involuntary acts in relation to their causal role. This perspective allows the causal history of an occurrent self to include all prior moments of (possible) autonomy. In this way, the problem of agent causation *within* an event whole is finessed, and with it the causal history of the person. In a word, moments of autonomy do not leave gaps in one's causal history.

From the standpoint of the agent, *event* causation includes the effects of the body on other objects, e.g., the effect of a limb movement on a physical object. The chain of events that follows the movement of a limb is explicable in causal terms; e.g., opening the door, putting the body in the car, starting the ignition, stepping on the gas, and so on. The history of a given action after its initiation, the effects or consequences of the action, including the actions of an object or individual other than the observer, e.g., the failure of the brakes, the mental life of the driver in an oncoming car, can be interpreted in accord with event causation. Event causation can be applied to events prior to a state of willing and to events *subsequent* to a state of acting, to an act that seems to bypass the will or to one that arises from another cause, but not, except in cases of a depleted or degraded (akrasic) will or self, to the step from will to action. This step is the domain of agent causation.

Components of Agent Causation

There are several reasons why the presumed causal relation between the self-will and its thoughts or actions (agent causation) can be distinguished from the ordinary object or event causation of causal science. The most important concern (*1.*) the duration of the self; (*2.*) the emergence of the self within a duration; (*3.*) the feeling of spontaneity in decision-making; (*4.*) the feeling that the self as a cause and the action as an effect are distinct occurrences; and (*5.*) the problem of causal persistence (p. 84).

Duration

On a purely subjective basis, I do not feel that I drive or propel an act forward, as my finger turns a key, or as one object pushes against another, but rather that I surround an act such as moving my finger that issues out of me. This act arises in my self. I would like to say it is embedded or ingredient in my self. There is no awareness of a causal surface or interface between the self and its actions. This experience

fills a certain duration. The experience is of a self that envelops its acts, possibly in the same way that an instant is embedded in a duration. A duration, like a self, passes "through time" as an envelope that continuously advances.

A duration is not a suspension of the past and present over time. It is not a segment of physical passage. It does not involve the persistence of the past as a stretch or line of memory to an occurrent present.[5] Rather, a duration is a virtual bridge across change that frames the content of the absolute present. The question is whether a subjective duration could uncouple mental and physical events for a self to emerge from or supervene on brain process. The mind-dependence of duration, i.e., its absence in physical passage, suggests that in free will, indeed, in every state of full self-awareness, autonomy might arise in duration as a "special cause" that intervenes in the decision to act.

There are two ways of conceiving agent control in a microgenetic framework. The first, conservative possibility is that values or the neural configurations that correspond with them infiltrate the deep self, articulate it, and constrain ensuing acts to conform with the character of the individual. Values are mostly instilled and compete for supremacy at a largely unconscious level. They can be viewed as weighting a decision by determining the configural properties of a self at a given moment. The effect of values on choices is a compatabilist approach to the problem of freedom, but not a legitimate source of free will, in the sense of a conscious choice undetermined by subsurface process. The effect of values is compatabilist because the influence of the unconscious on subsequent phases is comparable to the influence of one's causal ancestry. But the "level of consciousness" is not decisive in this matter. The conscious self might constrain the action in a similar way, just later in its microgeny, and this effect would also be a species of compatibilism.

The second, more speculative possibility, involves a direct control on choice by a *conscious and autonomous* self. This effect would constitute a special cause and would differ from that of a bottleneck on emerging contents or a filter of "bottom-up" constraints. The self that is responsible for the feeling of agency and choice spans the duration of the present. This duration incorporates a succession of physical brain states but does not correlate with the physiological events that intervene in that succession. Through the overlap of physical states the duration of the self becomes independent of physical passage and thus achieves autonomy.

However, autonomy is one thing, control of an action another. For agent control, more than a duration is required. Animals or young children presumably have a present, i.e., a duration or now, but lack the

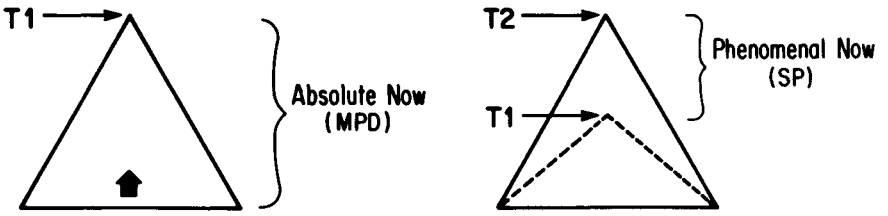

FIGURE 6.1. The absolute now, or minimal perceptual duration (MPD), constitutes the actualization of a single mind–brain state at T1. The phenomenal or specious present (SP) is a virtual duration that is extracted from the disparity between the actual (present) surface of the state at T2 and the depth (past) to which the T1 state decays.

capacity for genuine agency or volition. An agentive self occupies the present though the latter does not guarantee agency. While a rudimentary self-concept is probably necessary for a conscious now, agency requires a self in consciousness in a particular context.

I would speculate that the control of an action is linked to the relation between a duration and the process from which the duration is extracted. A duration is a *virtual* frame for the transformation over which it is derived. This is the basis of its emergence. Agent control obtains if the self can effect contents within the duration boundaries. This effect would be a whole–part relation in which the whole (self) that supervenes the duration influences the parts (acts) that fractionate out of it.

Consciousness is relational. One aspect of this relation is that of before and after, the relation of a self to the contents it generates. There is a transition *from* a conscious self *to* an image or an act. This transition resembles the relation of a duration to an instant. The whole of the self encloses and drives or modulates the part acts or instants that individuate within it. The individuation of the self to acts in the present resembles the individuation of instants in a segment of duration. There is in fact a deep analogy between the partition of a whole (self) to acts, images, or objects and the partition of a duration to instants. An instant or an act is a temporal "fact" that differs from the more conceptual quality of a duration or a self.

Another property of duration that contributes to a self "outside" of space and time is that a duration incorporates phases in becoming that are simultaneous until the duration is established. The phases are simultaneous because the becoming must be complete for the events to be seriated. There is no past for a given present until the present actualizes. The past cannot be causal for the present if the present must first appear for there to be a past in relation to it. The present is not like an effect of a prior cause that must materialize before it can be

known. The past has to be continuously renewed in every present. This is the personal, historical past, not just the forward edge of the immediately prior instant.

Events in mind or perception are not strung together in temporal sequence but are generated out of simultaneity and deposited in temporal order. Before a sequence actualizes, its phases are paradoxically atemporal[6] since time is created in their actualization. The resolution of simultaneous phases in becoming is diachronic with respect to relations of precedence, and synchronic with respect to their simultaneity prior to actuality. The precedence of phases in the mental state, e.g., the laying down of the self before an object, accommodates the whole–part transition, i.e., the self as whole and its acts as parts, while the noncausal nature of the becoming is grounded in the simultaneity that "exists" before the state actualizes. One can say that "mental causation" reaches its lower limit at the minimal duration where seriality disappears.

If the minimal duration of the mental state is not divisible, all occurrent events, including memories, are simultaneous with respect to that duration. The replacement of simultaneities is the engine of succession and *apparent* change in the world. Causation is inferred from contiguous replacements. Each mental state is a change within an irreducible simultaneity.[7] This means there is no cause or effect for occurrences in the mental state.[8] The price of incompatibilism is not a refutation of free will but of universal causation. The reality of free will entails a revision of causal theory. For one thing, since duration is established *after* a mental state actualizes, agency may require something like backward causation (p. 48). Since temporality and the "string" of world events actualizes out of simultaneity, for the agent to influence the string requires a complete derivation before the present enclosing that string is extracted.[9]

Emergence

The dynamic cascade of whole–part shifts in becoming is a microcosm of emergence and a precursor of agent control. This dynamic is always from the whole to the parts, i.e., fractionation, not the reverse (Chapter 12). The whole to part analysis over phases in the mental state constrains emergent form. The nested whole–part progression, e.g., the specification of gestalts to features, provides a basis for the evolution of agent control. Specifically, the whole–part transition is a paradigm for the relation between a self and an act. The shift that underlies the derivation of a self to an act is more than pure succession. A succession of phases occurs in nonagentive cognition with control by way of stepwise

constraints. In agency, the self is "spread" over the duration of a present within which the parts (acts) individuate.

The individuation of sequential acts within the simultaneity of the present creates objects in temporal order to conform with the expectations of event causation. That is, acts are serialized as they "come up" and the *resultant* serial order is perceived as the sequence of causal change in the world. Because the self is a simultaneity prior to the actuality of temporal "facts" it is atemporal, and thus satisfies the definition of a "special cause" outside of time.

Mind is transformed at every stage in the evolution and development of the self. The appearance of a self obligates an attenuation of object formation. This has the consequence of an elaboration of conceptual phases prior to object selection. A growth of conceptual phases generated by the self-concept fractionates the object at its depths and proliferates its underlying concepts. The parcellation samples the potential in the concept and is the basis of choice. This leads to the individuation of more than one option and the resolution of a decision from a state of choice, i.e., the zeroing in on a target concept. The growth of the self is accomplished in a series of whole–part shifts of extraordinary complexity.

Spontaneity

The spontaneity of choice and the openness of actions available to a self acting without compulsion are in striking contrast to the fixed sequence of event causation. The presumption of a free, autonomous, or spontaneous self that is not entirely the product of its causal history carries with it the problem of the unpredictability of behavior given a self that is undetermined. Hobart,[10] following Hume, critiqued the idea of a self that acts independent of character. Many scholars have argued, decisively in my opinion, that a self that is not founded on character would be capricious and unable to motivate or rationalize a decision. How could a self be wrenched from its history and personality at the moment of a decision. Character accounts for the decision a self makes. One's identity or uniqueness *is* one's experience and what, apart from this experience, is the self?

Independence from character is not required by autonomy, and certainly, it is not a part of the experience of spontaneity. Autonomy, one can say, is a postulate or a theory while spontaneity is a feeling. Since autonomy depends on an emergence *across* phases in an occurrent brain state, not across successive brain or mental states, and since the self is revived out of the past in every mental state, autonomy does not preclude the preservation of character at the moment of decision. The

question is not whether an occasion of autonomy has the implication of randomness (it does not), nor whether options have their sources elsewhere than character (they do not), but whether a decision supervenes brain process and is thus incompletely determined by it or if mental contents are epiphenomenal or identical to brain events. If the former, the feeling of spontaneity is an intuition of genuine free will. If the latter, one can ask, what accounts for a feeling of spontaneity that would support the delusion of free will, and what is the reason or justification for a false belief that is so widespread?

Independent of whether the belief is true or not, the *feeling* of spontaneity and thus the personal belief in free will has several possible sources. These include an intimation of the process of creative becoming, the iteration in the process from past to present, and the priority of the self in potential as opposed to the factness or actuality of present content. Together, these give the feeling of a potential prior to the realization of world objects or acts. The feeling of potential cannot belong to the past that leads to the present (the objective past is fixed) nor in the immutability of the present state as it actualizes (potential is relinquished on becoming actual) but in the process through which the past is revived in becoming the present.

Mark Strand has written, "Ah, the potential past, how it swells . . . " These are the layers of pastness in the momentary mental state, not the past of science and history. Yet in a causal theory of change, a fixed past is constantly laying down the present in a commitment to unalterability. Somewhere between the fixity of the past and the finality of the present lies the potential of the self. This is a virtual "somewhere" in the becoming of the present out of a "temporal space" distinct from the presentless flow of nature. The potential in the self contributes the feeling of possibility, of spontaneity, and (future) unpredictability at the interface of actual events with (past) causation.

Self-control and potential are important to the feeling of spontaneity but so also is the experience that decisions are made and complex actions initiated without an awareness of forethought. There can be a certain freedom in a lack of deliberation. Frankfurt[11] has noted that a person free to choose anything, even his own preferences, would be unable to make a decision. Moreover, conscious reason and logic do not always (ever?) play a role in decision making. The surfacing of decisions and behaviors from subconscious process obscures the links, i.e., the microprocess, that would explicate the behavior. The lack of access to preliminary phases in the becoming process fosters the deception that actions are spontaneous states of will. Indeed, arguably the most free and spontaneous of all acts, creativity, is characterized by pas-

sivity and recipience to contents which surface *through* a self that feels involuntary.

Cause and Effect in Agent Causation

The problem of distinguishing cause and effect is similar in agent and event causation (pp. 27–38). The terminus of a cause and the commencement of an effect, or the point at which an effect becomes the cause of a later effect, are common problems of change between minimal elements in physical and mental systems. In both types of causation, the demarcation of causes and effects leads to great difficulties. Events demarcated in a causal pair have an intervening duration that permits the embedding of an infinite number of additional pairs. This problem is usually ignored in serial accounts of event causation, or it is eliminated by the assumption of simultaneous causation where the boundary of the cause is identical to the boundary of the effect. Still we cannot account for the *passage* from one event or simultaneity to another unless, as is unlikely, time and change are independent.

The crux of agent causation is the need for a physiological or psychological interface between the self (agent) and some component of the action. Some *thing* has to be effected and this *thing* must receive the cause, or the cause must impact at some point in its structure. Discussions of causation tend to sidestep this problem, just as they avoid the problem of change at the moment of the effect.

For example, in agent causation it suffices to give a description of an intention and a later complex action, ignoring the event surround of the intention or the action, or the events that fill their interval. These events can be construed as nested causal pairs, as segments of a continuum, or as replacements of states of becoming depending on a concept of change. However, a causal theory should make an effort to demarcate cause and effect in an event sequence whether change in the transition is simultaneous with the effect, part of the effect, or prior to it.

In event causation, events can be inserted in the divisible interval between cause and effect. When lightning sets a tree on fire, the series of intervening physical reactions can, in principle, be specified. Generally, event causation requires that causes be relatively proximate to effects. This is not the case in agent causation where the demarcation of a cause is facilitated by a delay in the effect. A quiescent or deliberative interval between an intention and an action enhances the feeling of freedom (thought before action), and the feeling of causation (the effectuation of an action well after a decision has been made).

From another perspective, Davidson[12] has argued that the indistinctness of cause and effect typifies only the "garden variety of causality,"

not the relation between agent and action. In agent causation, cause and effect are distinct. This is a result of the delay interposed between willing or deciding and acting and the feeling of a transition from one (mental) domain to another (physical) domain.

Davidson goes on to say "what distinguishes agent causation from ordinary (event) causation is that no expansion into a tale of two events is possible." By this he means that the act of willing is itself an action in addition to the action that is its effect. This is consistent with the microgenetic concept that every mental state contains an action, and that every act of will is initiated in the generation through action of subjectivity. On this view, even if willing or deciding is not itself a direct action, i.e., if willing precedes but does not obligate a movement, there is an implicit action that it accompanies.

Agency and Causal Persistence

Part of the difficulty in distinguishing agent and event causation is the confusion between a causal theory of interaction and a causal theory of persistence. In physical causation, in addition to the effect on other entities, every entity is the cause and effect of its own persistence.[13] The atoms of a billiard ball replicate the ball a moment later. Even if one grants that the atoms of the billiard ball are reconstituted the next moment as an (apparently) identical ball, persistence is not always viewed as causal, or at least not in the same way as interaction. When one ball strikes another, cause and effect are clear even when, as with billiard ball causation, a case can be made for simultaneous cause and effect. The cause is assigned to one ball and the effect to the other. For a causal theory of persistence, one has to accept that a given ball is not exactly the same at two successive moments. The ubiquity of interaction as the prototype for event causation is in sharp contrast with the situation in agent causation where cause and effect always represent different states of the "same" entity; i.e., the subject. Specifically, agent causation is a change in state within a single persisting object, with cause and effect referring to sequential states in that object (subject).

Put differently, the change in a single ball from one moment to the next is like the change in a subject from one moment to the next. In the subject, this change is the shift from an intentional state, i.e., a mental set or act such as deciding to turn on a switch, to an intentional action, e.g., lifting the finger. This shift is the assumption of a novel state by the "same" reconstituted agent. From the standpoint of event causation, this is causal persistence, while from a subjective standpoint, causal persistence is the basis of agent causation.

In contrast, the change transmitted from one ball to another is like the change, from a third person standpoint, from an action of a subject (lifting the arm) to the effect of that action on an independent object (pressing a switch). This is causal interaction or ordinary event causation. The difference between agent and event causation dissolves as soon as one accepts that change *within* an entity is or is not causal in the same sense as change across entities.

On this view, the action of moving my finger (agent causation) differs from my finger turning on the light (event causation) as the causal effect of one object on another differs from the persistence of one object. The persistence of a billiard ball represents a type of causation inherently similar to the impact of one ball on another. It is a question of the energy state of constituent atoms. The difference between interaction and persistence in event causation is resolved in the concept of the becoming of the object as the only locus of real change. Ultimately, in the ontology of the world, all change is "causal" persistence, i.e., near recurrence, as the world object actualizes and perishes, to be replaced by another world object in a creative act of becoming.

From this point of view, the process *between* cause and effect in object causation is comparable to that in agent causation. The (apparent) implementation by the self of a physical movement is a form of causal persistence, in the resurgence of an actual object (or subject). In causation, a slight difference across instances (causes, effects) is interpreted as causal persistence, whereas a marked difference across instances is interpreted as causal "interaction," though instances (events) in both cases are only surface markers of uniform deep change. In this iteration of states, the depth survives the surface. Concepts outlast their implementations in acts, objects, and utterances. The surface replacement tends to be interpreted as event (object) causation. The depth replacement is agent (self) persistence. Agent persistence with object change is the basis of (the feeling of) agent causation.

MICROPROCESS OF AGENT CAUSATION

What is the effect of a decision to act if the decision is a (the) cause of the action? The target of the decision appears to be the goal of the action, and the attainment of the goal appears to be the effect, even if there is a prolonged delay between the decision and the action. I decide to move my finger and then move it. The decision to act is the cause, and the action is the effect. This is true whether the action follows immediately on the decision or after a delay. But how can the decision to move a finger in thirty seconds be the cause of the action thirty seconds

later? This decision might be the cause of an immediately ensuing state, which would then, on the model of event causation, be the cause of the state that follows. This would imply a series of cause–effect pairs leading up to the finger movement, the cause of which would not be the state thirty seconds earlier but the state just prior to the action.

The decision to act might also be the "cause" of whatever action it accompanies. Willing in the form of deliberation is an action correlate. Every state of an agent, say the state of deciding and the state of finger movement, is an instance of both will and action. Action is implicit in will, will in action. Will can be expressed in reflection, desire, choice, decision, impulse, or motility depending on the nature of the mental state at successive moments in the same subject. Deliberation before an action replaces that action though it has an action of its own. Later, the final action takes the place of actions that accompanied the deliberation. One can probably not have a deliberation that is conceptually related to an action and an action that is the target of the deliberation at precisely the same moment. Perhaps Rilke had this in mind when he wrote; "I want my own will, and I want simply to be with my will, as it goes toward action."

If the will in choosing or deciding involves or accompanies some action, this action, not the goal of the decision, is coupled with the act of willing, even if the willing and the goal are closely (conceptually) associated. This would seem to be true irrespective of the delay in acting, at least within brief intervals. The action of willing, or the action that accompanies willing, is more proximately associated with willing than the action that follows willing. The latter can occur after a variable delay and the duration does not impact crucially on the voluntariness of the act.

After thirty seconds another act of will, say an impulse to act, corresponds with the movement of the finger. There is a difference between this impulse, which is more spontaneous and motoric, perhaps unconscious, and the preceding act of willing, which is ideational and deliberate. These are two different acts and two different mental states.

In other words, we have the impression that a state of deciding at T1 is the cause of an ensuing action at T2, whether the action follows immediately or after a delay. Deciding and choosing are expressions of the mental state at T1; the action (finger movement) is an expression of the mental state at T2. Each mental state is an encapsulated whole consisting of an action and a perception. That is, every mental state is a complex act-object. At T1, the becoming deposits an act and an object. The act is the motor accompaniment of the deliberation and the *affective* tonality of the intentional content. The object is the content of the intentional state, whether this content is labeled as belief, the proposi-

tional content of deliberation, a plan, a choice, etc. This content is the perceptual equivalent of the action at T1.

At T1, perceptual content dominates. The action is not delayed but discharges into movements congruent with the decisional state, for example into its postural, gestural, and articulatory manifestations. The next moment, the state of T1 is replaced by another state (T1a), and another (T1b), and so on through the series of states intervening between T1 and T2, until the penultimate state is replaced, finally, by the state at T2. This state, T2, develops over latent configurations that issue from the state at T1, revived in the contextual background of the T2 action. The T2 state consists of actions and perceptions, but differently proportioned than at T1. At T1, the perceptual contents of decision or belief predominate, while at T2 it is their equivalent in action.

Delay

The problem of action, like that of causation generally, is tied to a theory of time and change, specifically to the relation between past and present. In contrast to event causation, a voluntary action and its cause (will, belief etc.) need not be contiguous or simultaneous. In event causation, a cause is continuous with its effect. When there is a delay, e.g., between lightning and a clap of thunder, a causal chain is intercalated to account for the delay, e.g., the propogation of sound waves. The delay is filled with nested causal pairs. If there is a real delay, however small, one can assume that the proximate cause has not been identified.

In voluntary action, as discussed above, a delay between cause and action is fundamental to the feeling of agency. Immediacy or simultaneity of cause and action most often characterize automaticity, not volition. An increasing "distance" from the immediacy of action, i.e., the uncoupling of decision and action, is the hallmark of volition. The delay itself can be voluntarily prolonged. The self can extend a delay that seems empty. Yet even if the delay is "filled," it is impossible to specify the intervening chain of events. The control over the delay, the inability to recall or recount its contents, and the discontinuity between initial and final state, sharpen the independence of cause and effect as separate events and distinguish agency from ordinary causation.

There is also the matter of first and third person perspectives. From outside, from the perspective of another subject, agent causation is not a problem; the subject is a billiard ball with a machine intelligence. From inside, from the standpoint of the agent,[14] the delay in action and the distal ("end-of-the-chain") effect on external objects creates the illusion of a sequence from will, to action, to an effect on outer objects corresponding to a chain of physical reactions, while in reality the "links"

or events in this chain represent successive states of persistence of the same entity, i.e., the subject.

Delay and Memory

Consider the example of a decision to move a finger in thirty seconds. Suppose after I make this decision, I move my finger thirty seconds later. I know that thirty seconds before, I decided to move my finger in just that way. In the present of the finger movement, my decision to act thirty seconds ago is a memory image. I recall the decision and construe the action as volitional. This will be so even if the prior state did not actually occur, i.e., if I am deluded, or if it was the result of, say, a hypnotic suggestion, since the prior state accompanies the movement as an imaginary, (recollected) image whether or not it was once veridical. What is the subjective difference between a true and false recollection if the recollection is apprehended as the cause of a subsequent action? How does an intention to act, as a cause of the action, say, to lift my finger, differ from a hallucinated voice dictating the same action? That is, how do they differ in their causal roles? How is the revival of a past event as a present image related to the present action of which it is the cause?

If someone hears an inner voice directing him to carry out an action, or is under hypnosis, how does this differ from the subject's own "voice" (inner speech) deciding on the same act?[15] In the first condition, agency is assigned to the voice, in relation to which the subject is an object. In the second condition, agency is attributed to the self. Now, an auditory hallucination is still inner speech. It is a product of the self. Normally, the self is active in relation to its propositional content. In hallucination, the self is passive in relation to this content. The difference between passivity (e.g., a command hallucination) and activity (e.g., a willed action) has to do with the action that accompanies the perceptual content. A mental sentence is a perceptual event that can be interpreted as inner speech, hallucination, or perception, depending on the degree of perceptual completion and accompanying (vocal) action. The continuum across these states is documented in clinical case studies.[16] With a shift to action, the process inclines to inner speech or an utterance. The feeling of agency depends on this action component, but does the causal relation between the agent (the self, an hallucinated voice) and the action differ in these conditions?

These considerations raise the following question: If I examine the memory of an event in the recent past, say a thought I had thirty seconds ago, my present self feels itself to be an agent in relation to the image of the past event. The *present* self *remembers* the thought as a content

in memory. But when that event is a memory of the self in a decisional state, or a memory of a decision, e.g., the self deciding to act thirty seconds ago, the memory state becomes agentive in the present. The *prior* self is not just revived in the present state but "causes" the present action to occur. These two memory images do not appear to differ, substantively. In one instance there is a thought, say, a thought about an action in the immediate future. In the second instance, the thought is to implement an action in the (same) immediate future. Why should the fact that the second thought realizes an action serve to distinguish the agency of the two images?

It seems the only way to understand how a self can be either active ***or*** passive in relation to a memory image is to assume that the self in a decisional state, i.e., a decisional or intentional self, persists over the thirty second interval. The concept or intention to lift the finger configures or models the self in each exemplar of a series of states over thirty seconds.

That this must be so is seen by considering a person who is distracted during the thirty-second interval so that he forgets to lift his finger. Suppose, at the end of the thirty seconds, the person (self) recalls that he planned to move his finger, and then carries out the movement. In the normative case, the self intends and wills the act, and thirty seconds later feels an agent to the finger movement. In the case of distraction or forgetting with recall of the intention after thirty seconds, the self remembers the prior state of intention and moves a finger. However, when the intention is a memory, it is no longer agentive. The difference in agency in these two examples suggests that an agentive self requires the "causal" persistence (recurrence) of a mental state within the category (concept) of the original decision. The decision constrains ensuing states of the self, even if, for example, the self is counting seconds and the decisional element is not emphatic in the intervening duration.

In the case of a forgetting or failure to revive the decisional state, or with distraction, there is interference or decay of the original state, so that after thirty seconds the content of the state has lapsed beneath the "floor" of the present. This recession of the original state explains its recurrence as a *memory* rather than a present intention. Just as the rehearsal of a telephone number prevents forgetting and keeps the content fresh, the iteration of the decisional self maintains its presentness up to the moment of action.

These comments illustrate that the delay between decision and action concerns, fundamentally, subjective duration and the awareness of time. There is always a delay between willing and acting of some duration. The length of the delay might be a day, a second, or a fraction of a second, but regardless of the duration, the initial T1 intention is a

memory in T2 at the moment of action. If, at T2, this T1 memory is apprehended as a memory, it is "extrinsic" to the self at T2 and not felt as agentive. If it recurs in the succession of states from T1 to T2, as the self's intention over this duration, it (the self deciding) is felt as agentive at T2, but not as a memory, though there is the awareness that thirty seconds ago a decision was made to act. The agent is the self at T2, not T1, with the duration between T1 *and* T2 (the delay before acting) extrapolated from the decay of T1 in T2. The "causal" relation between intention and action hinges on this duration (delay). The T1 intention seems to impel the action thirty seconds later, but thirty seconds later that intention is a memory ingredient in the present of the T2 state.

In the mental state of *deciding* at T1, the action-to-be is only a possibility. The state at T1 is all that exists. The mental state of *acting* at T2 revives the content of the prior state of deciding, but for that moment only T2 exists.[17] The self at T1 seems to bring T2 about, though it would be more accurate to say that T1 predicts or imagines the then nonexistent T2. The feeling of agency at T2 is not a direct effect of the T1 state. The intention does not cause the action, unless all of the states between the intention (T1) and the action (T2) form an embedded causal series.

On a microgenetic account, the present of T1 does not cause the next present, nor a later present at T2, but is replaced over a series of presents, with the duration from onset at T1 to conclusion at T2 represented in the precedence of phases within T2.[18] The sequence from self to action is not a concatenation but a superimposition of near replications, with the interval between intention and action derived from intrinsic features of the finalmost replicate.

The complex from T1 to T2 is the framework of an intentional act. Given that the self at T1 shapes introspective content, e.g., the selection (resolution) of an idea, a proposition or a decision, the next pulse of mentation would revive this concept in the self of that moment. This continues over thirty seconds of clock time, embracing a sequence of perhaps three hundred mental states, each expressing a portion of the content of the initial state. At T2, the T1 concept exhausts itself in movement. If the planned action is a finger movement, a movement of the finger satisfies the recurrent ideational content. If the planned action is a trip to France, many partial actions over millions of mental states will be required to serially deplete the underlying concept. The states that intervene from T1 to T2 are acts or preparatory acts, and each expresses the self of that moment. Every state after T1 revives the fading T-1 intention to approximate the initial configuration. Like a single note that

slowly changes each time it is replayed, the penultimate "replicate" of the slow change of T1 is eventually derived to an action at T2.

Put differently, we may assume that each becoming is a configurational wave or track over which the next sequence develops. An action, a self, or an object develops on the track of the immediately prior state. This track guides the configuration that follows.[19] If the immediately prior state is the penultimate revival of T1 content, the configural effects that remain of this content contribute to the intentional quality of the action at T2. Specifically, T1 configures T1a, which configures T1b and so on, to T1n which then configures T2. At T2, T1 exists as the residual configuring effect of T1n on T2. If T2 is all that exists at that moment, including the illusory duration back to T1, agency and autonomy must occur within T2, not in the effect of the self at T1 on an action at T2.

NOTES

1. A good discussion of arguments for and against autonomy can be found in P. Van Inwagen, *An Essay on Free Will* (Oxford: Oxford University Press, 1983).
2. For example, see Susan Wolf, *Freedom Within Reason* (Oxford: Oxford University Press, 1990).
3. R. Kane, *Free Will and Values* (Albany, NY: SUNY Press, 1985).
4. H. Frankfurt, *The Importance of What We Care About* (Cambridge, U.K.: Cambridge University Press, 1988).
5. See discussion in Brown, *Self and Process*, 127–146.
6. There is a similarity with Whitehead's account of becoming. See A. N. Whitehead, *Process and Reality* (New York: Macmillan, 1929).
7. *Self and Process*, 189–191.
8. Simultaneous causation should not, in the strict sense, imply a causal relation between simultaneous events since, within a simultaneity, causes and effects are not distinguishable.
9. H. Price, "A Neglected Route to Realism about Quantum Mechanics," *Mind* 103(1994): 303–336.
10. R. E. Hobart, "Free Will as Involving Determinism and Inconceivable Without It." In B. Berofsky, ed., *Free Will and Determinism* (New York, Harper & Row, 1966).
11. Frankfurt, *What We Care About*.
12. D. Davidson, *Actions and Events* (Oxford: Clarendon Press, 1980), 53.
13. B. Russell, *Human Knowledge* (New York: Simon and Schuster, 1948).
14. See arguments in T. Nagel, *Mortal Questions* (Cambridge, U.K.: Cambridge University Press, 1979); Russell's distinction of knowledge by description or by acquaintance, discussed inter alia in G. Madell, *The Identity of the Self* (Edinburgh: Edinburgh University Press, 1981); and J. Searle, *The Rediscovery of the Mind* (Cambridge, MA: MIT Press, 1992).
15. This problem is discussed in *Life of the Mind*, 22–23.
16. *Life of the Mind*, 173–251.
17. In a similar vein, C.S. Peirce wrote, "taking the word 'idea' in the sense of an event in an individual consciousness, it is clear that an idea once past is gone forever, and

any supposed recurrence of it is another idea. These two ideas are not present in the same state of consciousness, and therefore cannot possibly be compared."

18. *Self and Process*, 127–146.
19. *Self and Process*, 45–46.

CHAPTER 7

Intention

ARGUMENT: *A subject is an object that actualizes in an action. The self is a preliminary phase in a subject that corresponds with choice. When this phase actualizes in a behavior directed to a goal, the subject is in a basic (purposeful) intentional state. Awareness of the goal of a purposeful action requires a self and an intentional ideation. An object is selected out of a context or potential of unactualized choices. A self that is conscious of the choice implicit in the selection of a goal is in a volitional state. The progression from simple (purposeful) to complex (conscious) intentions to volitions (choices) involves a growth of perceptual awareness realized in action.*

SUBJECT AND OBJECT

Intentionality describes a state in which there is an aim or direction to an object. The intention is about something thought of or acted on. This something is an object (a content,[1] idea, etc.). Intentionality entails a subject or a self and an object. For example, hope is intentional; something is hoped for and someone is doing the hoping. Even if an intentional state cannot be cleaved from its object, there is more than the object to the intentional state—more even than the intention—and this extra portion is the subject.

What exactly is a subject (p. 169)? A subject is an object in relation to another object. The relation establishes which object is a subject and which is an object. The relation is not between a subject and one of its objects. A subject, unlike a self, does not have objects. A subject can be distinguished from an object but does not have objects of its own.

A subject is unaware of an opposition to other objects (or subjects). Nor is the subject aware of itself as an object. A subject is not a self or an agent.

When an action is generated in an object, the object takes on a direction. The direction entails a commitment, which is for an active object, i.e., a subject. An object becomes a subject through a shift in the axis of its becoming. The derivation of primitive will through drive into action is fundamental to this shift. An action is not an output but an internal state. An object moves, a subject acts. The action of a subject is not its movement. To interpret an action as a movement, i.e., as the output of a subject, is to objectivize the subject that the action generates. A subject does not produce an action but is a product of an action development. In a volition, the self is also an outcome not an instigator.

An object must first become a subject before it can become a self. A self emerges within a subject. The same individual can be a subject one moment and a self at another moment. Subject and self are not properties or conditions that endure but states of transient actuality. A subject is one with an action. A self and its actions are (apprehended as) distinct. A subject becomes a self when an act or object is *selected* (see below). A process of choice or decision or an awareness of an action that is forthcoming, whether the decision is conscious or implicit, is necessary for a (voluntary) intentional state. Intentionality is implicit in purposeful action, but intention with volition (choice) implies the presence of a self.

When a dog digs up a bone, or a bear eats a fruit, that is a subject engaged in a type of purposeful action. A person in a dissociative or fugue state, a transitional state or sleepwalking, perhaps under hypnosis, is a subject that acts in a purposeful manner. We would not term this behavior voluntary. There is no choice, there is no possibility of not acting. There are degrees of purposefulness leading to volition and there are gradations from a subject to a preconscious self to a preconscious self.

The objectivist attribution of "intentionality" to a machine intelligence based on the inability to distinguish an intrinsic intentionality from its simulacrum, and the eliminativist attack on the legitimacy of private content, result from a failure to distinguish the purposeful and the volitional in the description of intentional states. We can describe intentional behavior in others where a subject and a self are not distinguished, but the ascription of intentionality to one's own behavior, or of volition to the behavior of others, involves the consciousness of self in a state of choice.

SUBJECT AND SELF

A deep or unconscious self can be distinguished from a conscious self. The deep self is the core of the mental state (p. 105). A subject includes this core plus the actualized organism. The "person" is all of this plus the unconscious body (schema) and the perceived body (image). The body schema is deeper than the perceived body (p. 63).[2] The unconscious body schema actualizes in the body image which is its conscious derivation, as the deep self actualizes to the self in consciousness. It is necessary to distinguish a self and a subject to account for the origins of the self-concept, and the difference between purposeful and volitional behavior. The transition from subject to self is the transition to agency. The conscious self can reasonably say, "I am not the whole of me." Here, the *I* is the conscious part of the self nested in a subject, while the *me* includes the unconscious portion. The whole is the subject, including the deep or core self; the I is a partial expression of the self, and the conscious self is a configured portion of the wholeness of subjectivity.

A distinction between a subject and an object that involves a direction or intention from the subject's perspective is a distinction between a self and a mental object. The distinction between a self and one of its objects or the self and the object world, or the self in a state of reflective consciousness, or the self and "its self" (core self and/or subjectivity), can change moment to moment depending on the actualization process. How do I know, for example, that I am the subject and not the object in my own intentional state? In the distinction of subject and object where the object has a nexus to the subject, i.e., where there is an apprehended relation between a subject and an object, the subject or self experiences its own objects. The consciousness of this relation is a sign that a self-concept has actualized.

When subjectivity is opposed to an external object, for example, when a bear eats a fruit, the judgment of a purposeful action is made by an outside observor. The bear distinguishes the fruit from its own body—it does not eat its paw—and different types of objects elicit different responses. The behavior of the bear exhibits all the signs of a subjectivity engaged in a purposeful action. Is the bear aware of a self–world distinction; is the opposition of self and world implicit in its behavior? Whether the act of eating a fruit is purposeful or intentional depends not on the directedness of the action, or the having of an object, but on the self-concept of the bear. For the bear to think about an object in its absence, thus about its own actions and those of other animals, including other bears, for the bear to have choices and to make decisions, requires a self. I don't know whether bears have this capacity

but I believe that chimpanzees do. Obviously, it takes a self to even raise such considerations. For pure subjectivity, they do not arise.

A self as a portion of a subject is an ideational content. The self-concept is a *concept*. If the self is an idea, could this idea be thought up by another object? Put differently, could the (mental) object of my thought be thinking me up rather than me thinking up the object? This is not unrelated to the question of whether a self is produced by the action it seems to empower. It is to be decided whether a self intends an act, or the act intends the self. Is the self elaborated by an intentional act, with the direction from self to action a feature of the elaborative process? If the state of hoping for an object generates both the subject (self) that is hoping and the object that is hoped for, what is subject and what is object depends on the state. If I dream that a person plans to kill me, I am the object of an intentional state within an object of my own making. Subject and object are not fixed relations but depend on the sequence of phases and relations between contents in a given actualization.

Action, then, creates a subject that can stand behind its own object formations. The direction of a subject *toward* an object is purposeful. The direction of a self *toward* the idea of a particular object is purposeful and intentional. Intentionality is a self in a subject in relation to the concept of an object. We can speak of the *subject* of an intentional or volitional state recognizing that the subject of this state is really a *self*, and we can speak of the *object* of an intentional or volitional state, recognizing that the object of this state is a *concept*, i.e., a mental object, not a material object in the world.

The Direction of Intention

The intentional relation, from the first person standpoint, merges into volition as the feeling of a direction of the self-in-subject toward an object (image) in the mind. An object is the goal of a purposeful, intentional state. An object that is volitional is a mental object. It has to be if the object is absent but, in truth, the goal is a mental object whether the "real" object is present or not. Both terms of the relation, self and object, are entities in the same mind. That is why it remains to be settled which entity is an agent and which is an object; or rather, which concept is the self-concept and which concept is the concept of the object.

The content of the self-concept and its object-concepts and the extent to which these concepts are shared within the background category, i.e., the deep self and the object category, are decisive in this relation. A limited number of selves, perhaps only one, are derived from the deep self, but there are many potential objects in an object category. For a person with many different selves, say a multiple personality, each

self is like an object in relation to which other selves have a more or less equivalent status. Such patients undergo a type of splitting. Conceivably, the unconscious self partitions to the conscious self, then into conceptual derivations that become relatively encapsulated by virtue of multiple instantiations. A behavior incompatible with the deep self might result from a prominence of one line of individuation.

One personality in a multiple could form an intentional attitude only if it was the entire self for the moment of its appearance. Were the self-concept a manifold of component selves, and were this manifold comparable to the multiple objects in an object category, e.g., instances of the category of furniture, there would be no way to establish a direction from a self to one of its objects. In what way does an object belong to one of many potential selves? The relation of self to object, i.e., the direction of an intention, could as well be a relation between two objects. This direction, *from* self *to* object, establishes which is agent and which is object. These considerations, and the clinical cases to which they refer (see below), make one wonder about the nature of this direction. Is the direction of intentionality or agency reversible? Can a self be a target for one of its own objects?

The direction *toward* an object, for example, when a self pursues an external object, is not a partly extrinsic relation, i.e., a relation with one part (self) in the mind, the other (object) in the world, but is a relation that inheres in the distinction of a subject and an object. The object is realized through the subject (or the self). For a self to have an object, or for a self to have a direction to an object, are probably the same thing, since it is unclear in what sense a self could have (create) an object without a direction to it. A direction *away* from an object (e.g., avoidance) is equivalent to a direction to the object since either direction can suffice for intention and either direction can distinguish a subject and an object.

A direction toward or away from an object is a movement toward the future. A sought after object is not in the future of the self but part of the future self. When I hope to win the lottery, the lottery is not the object I hope to win; what is hoped for is *my* winning the lottery, i.e., the valuation I have given to the lottery. I am involved with the object in the future state. Even with abstract hopes, for example, the hope for human rights in China, the hoped-for state is bound up with *my* valuations, i.e., that human rights in China are desirable.

Temporal Aspects of Intention

There is an appearance of a direction to the past in reminiscence, but a recollected image is a content in the present that attracts the interest

of the self. The self appears to seek or search out or look up an image. This image, however, is not a picture of the past but a configuration in the present with a feeling of pastness. How this feeling is generated as an accompaniment of the image development is a complex story.[3] The direction toward an image is comparable to the direction toward an object. Whether the image feels past or present it is still an image in the present. The feeling of pastness depends on ancillary features of the object formation.

The distinction of a memory and a plan bears on the account of intentionality. A thought about a past action is a memory, while a thought directed to the future is a plan or intention. If I think about what *I ate* last night for dinner, it is a memory. If I think about what *I will eat* tonight, it is a plan or intention. What is the *intrinsic* difference in thought in these examples? The past and future direction could simply be a matter of tense; i.e., I ate, I will eat. The fact that one is an image of a past act and the other an image of a future act seems incidental. Suppose I wish to repeat a past experience. Does this transform a memory to an intention? What is the nature of this alteration? Suppose I am deceived by my recollection. Is a false recollection of last night's dinner a thought to the degree it departs from reproduction. A memory, say for a madeleine, can be the inception of a creative act. What then is the difference between productive and reproductive thought? Inaccuracy in recall is a type of thinking. In cases of memory disorder with fabricated recall (confabulation), such inaccuracies can be quite creative.

The problem of memory is similar to that of a plan for the future. How does the *content* of thought about an action in the future determine whether the thought is intentional? In what way does a thought about a future *impersonal* event or the possibility of various future events, e.g., whether a space craft will land on Venus, where my own action is not involved, differ from a thought about a possible future event involving a personal action, such as whether I will lift my finger, or go to France next year?

Regardless of its object or goal, whether a hope, a reminiscence, or an apple on the table, intentionality is a forward-looking relation. The self is directed *toward* an image or an object. Without this forward relation, the image or object could just as well act on the self (see below). In dream and delusion the self is experienced as a victim of its own images. The forward direction establishes which object is the self of the intentional state, and which object is the object of the intentional state. The object in this instance is the future of a direction that is linked, I believe, to the direction of becoming of the mental state and the asymmetry of subjective time.

The feeling of the self-to-object direction as a forward movement is the direction of becoming from past to present (p. 127). A concept or image can become an object. The self is deposited prior to the concept or image. The subject is an act of the whole person. The self is an accentuation of preliminary segments in the subject, thus felt as prior, guiding the subject. That is why a subject can be an object for its own self. The feeling of the precedence of the self before its own (mental) objects, and the development from the potential of the past to an actuality in the present, are part of the intentional relation between a self and its ideas or objects. They account for the "one-way" direction of agency and intentional feeling.

In sum, a state of purposefulness requires a subject or self and an object. A subject differs from an object in the presence of a constituent action. In purposefulness, there is an active relation to an object. Intention has the *idea* of an object as a goal. In volition, there is choice prior to object selection, with consciousness of a direction toward the idea of one or more objects. An object in the actual present of a subject is in the immediate future of the self. The direction from a self to its objects unfolds from past to present. The precedence of the self before the object concept imparts a direction. Self, object, and direction constitute an atomic unit that cannot be disentangled. The directedness from self to object maps to the becoming of the mind–brain state.

This interpretation leaves unanswered the question of whether an intentional state like hope or fear is an instance in a class of intentional *effects*, or whether intentions instigate actions. Are intentions resultants or what a self brings to an object in a volitional act, namely the intention (will, desire) to act? In other words, is an intention a cause or an effect? In my view, the self brings about intentions (reasons, etc.) as surrogate actions, not causes.

Intention and Pathology[4]

There are nonintentional states in which the self is lost or threatened with loss, e.g., sleep, dream, anxiety, but is there a state of self-awareness that is not intentional, i.e., can a conscious self occur without being intentional? The self needs objects to exist. Having an object probably implies an intentional relation to it. A perceived object is not a coincidence. The self cannot survive without objects. *Ganzfeld* disorientation as in snow blindness or sensory deprivation exposes the vulnerability of the self to a loss of (visual) objects. The paranoia of peripheral deafness or the psychosis of cortical deafness reflect alterations in the self-concept with a loss of auditory objects. The self degrades with pathology

in more than one perceptual modality. Confusion is prominent in the acute stage of cortical blindness, but the self can recover if auditory and somaesthetic objects are spared. A loss of visual *and* auditory objects erodes the self-concept and can lead to a dreamlike consciousness. When object loss is pervasive, an agentive or conscious self cannnot be sustained.

Perceptual disorders can erode into the self-concept, while psychotic disorders can begin with a disturbance of the self-concept and erode into objects. This is the basis of object breakdown in schizophrenia and delusion. A psychosis that begins with an altered self-concept eventuates in derealization, illusion, and/or hallucination. These symptoms reflect a penetration, bottom-up, from the self-concept to (the concepts underlying) perceptual objects. There is uncertainty of the self–world boundary and, in severe cases, a blurring of dream and waking cognition. The feeling of agency is altered with inability to act and eventual catatonia. Symptoms such as "thought control" point to a loss of an active feeling for one's own mental contents.

In sum, a disturbance of the self impacts on objects, a disorder of objects can penetrate the self, and the direction of intention from self to object can be reversed in pathological cases. This is the basis of paranoia and persecutory objects. Self and object are on a continuum with the direction from self to object a dependent feature of the self and object concept. Such observations confirm the indissolubility of the binding of self to object and the fragility of intentional feeling.

Acts and Objects

An intentional action is directed toward a goal. The self feels this direction as an impetus or attraction. Action fills in the direction but is not necessary for the direction to be there. One can hope for something without an action other than the "action" of hoping. One can hope (fear, etc.) without acting. One can act without moving. One can think about an action or refuse to act, so unless an intention to act is itself an action, which it may well be, motoric action is superfluous to intention.

If intending to act but not acting, or intending not to act, is an action without a movement, could one speak of an act of intending[5], as one speaks of an act of perceiving? The latter is not an action so is intending an action? Every cognition has an action; the action minimally is its subjectivity. A subject in action is always in the background. A decision not to act is a continuation of ongoing behavior. If I consciously choose not to lift my hand, I tacitly choose to maintain my present

posture. The choice "not to act" is expressed in acting otherwise, say in standing fast, but in what sense is the refusal to act an action?

I see an apple, want it, and reach for it. What is the difference between seeing, wanting, and reaching, between the directed perception of an object, the intention in wanting and the intention in reaching? Not just the addition of a desire or a movement. In each case, the body is involved in a total adjustment. The action of reaching is a shift from the perception in wanting and the relatively "neutral" seeing of the object. The structure of the *perception* in seeing, the structure of wanting, the structure of the *action* in reaching: the difference is which structure is dominant at a given moment in relation to the sequence as a whole. This determines the intentional quality for a given mental state. The perceptual aspect is important. Perception supports the intention that action instantiates.

Intention is "about" an object. The aboutness in perceiving is the distinctness of self and object that shifts with intention to a direction. The shift to intention is a perceptual adjustment. There is no aboutness in action. Actions are blind to objects. The action generates the subject but its object is perceptual. The object or goal of an action, the "script" from which an action is supposedly "read off," the beliefs by which the action seems to be determined, are forms of *perceptual* intentionality distinct from action itself. The aboutness of an action is in its preparatory phases that are perceptual. Actions occur in body space. They discharge in the person. The object (or concept) toward which they are directed is a perceptual event. There is nothing in the action proper that contains the intentional component.

Perception elaborates a private *and* a public space for a self to live in. The self, consciousness, and its contents are perceptual. Even the action *as known* by the agent, i.e., the awareness of the action, is a product of a (reafferent) perceptual development. The aboutness of perception advances the immediacy of action to an intentional relation, as the "independent external" space of perception advances the action space of the body. Through perception, an action becomes intentional and develops to a space beyond the body.

This is not to say that action is a settled capacity left behind in the growth of perception. Acts and objects unfold together. The growth of an action leads to individuated elements that accompany the outward development of objects. Perceptual intentions, including beliefs and propositions, create the awareness of a directedness that accompanies an action, while certain of the ingredients of intentionality, e.g., the self-concept and its objects, correspond with aspects of percept formation, not the action proper.

An action consists of an action process giving a motor discharge and a feeling of activity, and a perceptual process generated by central and peripheral recurrence that provides (indirect) knowledge of the action that has occurred. The confounding of the various components of an action, i.e., a movement, an action process, secondary perceptual representations, plans, schemas, scripts that are purely perceptual, contaminates much of the theory on intentional acts. The action process makes an intention explicit. It specifies an intention by actualizing the aboutness relation.

Intentionality occurs in purposefulness. In a certain sense, the lack or suspension of a direction to an object in a state of choice or indecision is the mark of volition and agency. Davidson[6] considers intentionality the criterion of agency. He gives an example of Hamlet intentionally killing the man behind the arras, but not intentionally killing Polonius, the same man, noting that an action can be both intentional and nonintentional. In this case, however, whether the action is intentional is a matter of perceptual knowledge. Such examples suggest that the same action can be intentional or nonintentional depending on the state of belief. In my view, there is some difference in the structure of the action according to whether it is purposeful, intentional, or volitional.

In philosophical writings, actions are broadly considered to incorporate perceptual contents, while perceptions are narrowly conceived in terms of sense data and the immediate beliefs and concepts that arise from sensory experience. A more limited portrayal of action is justified by the observation that a pathology restricted to action systems, with few exceptions, does not affect intentionality. Thus, the intentional quality of thought persists in uncomplicated cases of apraxia, paralysis, or Parkinson's disease, though the initiation and/or implementation of the action may be impaired. With prefrontal injury, the planning of an act may be defective, but belief, desire, and the capacity to direct a simple action sequence are relatively undisturbed.

INTENTION AND VOLITION

Intentional action is intentionality in perception instantiated in action. This is the basis on which volitions and intentional ideations develop. The various forms of intention are modes or degrees of *perceptual* willing. Intentional states can be aligned from the archaic to the recent, just as one can distinguish a primitive and an evolved action. The fear of being devoured by a shark is more primitive than the fear of losing one's job. The former is an imminent possibility that is grounded in a

threatening object. The latter is more conceptual, more detached from an object. Even among motor intentions, there is a gradation of complexity. An act such as turning on the light is a simple form of intention that a chimpanzee can master. An action of this type is a prerequisite for more complex intentional *ideations*.

There is a transition from purposeful action without awareness, a dog finding a bone, to a conscious act that is purposeful, turning on a light, and intentional, fear of the dark, to an act that is "freely" willed and voluntary, the choice of a candle over a lamp. The feeling of agency tends to increase over this series, probably developing on the discovery of regularities in acting intentionally (reaching, grasping).[7] From a maturational standpoint, intention is a precondition for (the possibility of) free will, which, however, requires more than an intention; it requires that the decision implicit in an intention become explicit in awareness. Put differently, free will requires an explicit awareness of the choice that is implicit in purposefulness.

Actions realize perceptual intentions and satisfy or instantiate the direction between the self and an object. In purposeful action, an object (goal, plan, etc.) has resolved and the action has been selected. This is the implicit choice in an intentional act. If a choice has not been made the action is not intentional. Indecision is not strictly intentional, either because the act does not have an object (goal) or because there is more than one potential object. In the latter case, i.e., a state of choice, the object is usually indistinct and without direction, rather like a state of anxiety. The choices that matter in the resolution of an object are choices in the mind, i.e., at a pre-object phase, not between actual (final) objects in the world. Otherwise, a rat hesitating in a T-maze would be in a state of volition.

The contemplation of choices and the act of choosing are volitional and also intentional if the choices are clear in consciousness. Suppose a person does not know what to do in a situation. The various options that are verbalized are portions of an underlying state of uncertainty. Suppose the final action is not one of the options previously verbalized. Would such a person have been acting intentionally during the phase of contemplation? This phase is traversed in every act and object of a conscious self. Volition requires this phase to be experienced, not simply traversed. The clarity of a goal surfaces from the deeper struggle of unrealized alternatives. Choice is more than a hesitation filled by thought prior to decision. It is the awareness of the conflict or potential underlying purposefulness. Even when one's options are unclear, hesitation *implies* a choice, including the choice of not acting. For the rat in a T-maze, not choosing is probably not an option. Most likely, the hesitation reflects a competition among stimuli in a tropic behavior. In human

mentation, conscious hesitation points to an awareness of more than one potential outcome. In the ontogeny of choice and free will, the individual is first aware that an action has been decided and then withdraws to the bases of that decision in choice.

Choice, therefore, is not itself an option but a sign that options are obligatory in every conscious act. The choice in a voluntary act and the feeling of agency in deciding are not additions to intention and purposefulness but an uncovering of a usually automatic selection. An action is the product of a decision even if the agent is unaware of a choice. The object that "survives" is specified out of a concept. This object is what remains after other potential objects have been eliminated. All objects have their source in potential, and an implicit decision is inevitable as an object actualizes. There is continuous, if nonconscious, decision as an object resolves. A microgeny is a continuous specification through context to item shifts. The direction or commitment to an object that is essential for an intention is a way of describing the asymmetry of the specification process.

Objects (and subjects) are constantly being specified. I am actualizing with the objects I perceive. The feeling of a direction toward a goal is the intuition of a process through which I and my objects are actualizing. The direction of intention is introspection aligned with becoming, the intuition that process delivers objects.

Every object progresses toward definiteness. This progression takes different forms. Anxiety is a (usually) unpleasant expectation that is nonintentional because an object is not apparent. A preobject is distributed over the perceptual field. The action that accompanies the anxiety discharges in the space of the body, not in an external object. In fear, which is intentional, the object resolves or exteriorizes. The subject knows what he or she is anxious about. With the appearance of an object, a nonintentional state like anxiety becomes an intentional state like fear.

Choice is the narrowing of potential to more than one option. Options can be subsurface, in which case there may be conflict, weakness of will (akrasia), or indecision. Or options may become conscious, in which case the anxiety associated with preliminary phases dissipates and choice is less conflictual, even arbitrary. If choices are conscious and there is still conflict, the conceptual underpinnings of the options have not been fully delivered in their conscious representations. The unrealized or subconscious content generates the affective tonality.

A hope without an object is a state without a name. A state of "objectless" hope may be something like a pleasant feeling of expectation. The transition from expectation to hope is like the transition from anxiety to fear. A subject can be expectant but only a self can hope. A dog

that awaits its master with expectation cannot be said to hope the master will appear. The resolution of an object or the idea of an object, i.e., the goal of the hoping, is the transition to intention as the object (belief, goal) clarifies. The further shift from an intention to a volition involves the choice that underlies the object that has clarified.

Awareness develops with the directedness prior to the object in perception and increases with a choice of possibilities in development. Awareness and choice reflect segments in becoming as an object is replaced by its conceptual precursors. This taking-the-place is a prerequisite for introspection.

In sum, there is a continuum from immediate action to purposefulness without awareness, to awareness of a goal or purpose, to actions with hesitation over options. Awareness progresses from the automatic to the intentional (purposeful) to the volitional. The shift from purposefulness to volition is a reclamation of earlier phases in the mental state. Consciousness is the relation *across* moments in becoming as preliminary segments of the self, imagery, and choice are enhanced in the context of a complete actualization.

More precisely, intentionality or purposefulness is the awareness of directedness, the relation of a self to an object. Intentional action is an action in the context of perceptual intentionality. The subject is aware of the goal prior to its realization. Awareness of choice gives the feeling of free will or volition. In a situation of choice, a goal is incompletely specified. The action is delayed and the failure to select an outcome is apprehended as a menu of options with the freedom to choose among them. An occasion of *choice* entails the feeling one can choose, and the feeling of *free will in choosing* is the experience of the self as an agent.[8]

Decision requires the possibility of more than one direction, not the awareness of different directions; i.e., not an awareness of more than one intention. An awareness of two intentions is an awareness of something other than the direction of an intention. An awareness of a state of choice is centered on the concept of the object. If I am deciding whether to have cake or pie, I do not have two intentions but one intention with two objects (see below). In other cases, there is a potential for separate intentional paths, with the object of each path not clearly before the agent. The coming-to-the-fore of preliminary configurations exposes the potential embedded in the goal.

Searle[9] argues that intentionality is independent of language; some animals have basic intentional states (e.g., fear). Animals do show purposeful, goal-directed behavior, but presumably, most animals do not have the ideational equivalents of their intentional acts; i.e., complex perceptual intentions such as hope or guilt or explicit beliefs. Such

states are transitional from intentions to volitions. Through choice, volition extends the range of intentions.

Basic intentions are purposeful behaviors that expand the community of action to private introspective space. The directedness in a basic intention becomes the aboutness of a *perceptual* intention. The aboutness devolves to the choices available to the self in a volitional state. Choice is only superficially about objects. Choices have their sources in values that are part of the self. Animal cognition should approximate human volition according to the evolutionary status of the animal's self-concept. Thus, it is likely that a chimpanzee has a primitive sense of self, and should have some capacity for volition and choice.

Voluntary action presumes the desired event would not occur without the action or inaction that is the object of the willing. With choice, I know the implications of not acting; e.g., that I will burn my finger if I do not blow out the match. The counterfactual, that an effect would not have occurred without its cause, inheres in the implication of alternative courses of action. Intentional action is committed to an outcome. I can abort an intentional action, but if I do, choice is involved; i.e., the option not to act. If the action can be aborted, the direction was not determined in the first place. If the option of not acting, or acting otherwise, does not occur to the agent, the person is not in a state of choice and the action is not volitional[10] When choice intervenes, the action becomes voluntary. The counterfactual for an intentional action implies the choice not to act. This is not an element of intentionality but is derived secondarily from the experience of willing freely.

CHOICE

There are always choices, no matter how restricted or dire their consequences. A person in a police state is still "free" to agitate for change. He can step before a firing squad as a voluntary willed act. The will is engaged regardless of whether he is resigned to his fate, resists, or welcomes execution to demonstrate courage in the face of death, or martyrdom. In each situation, there are conflicts. The question is, should the jumble of external pressures and internal conflicts have a separate accounting, e.g. (trivial) internal and (decisive) external, or the reverse, or are all choices and conflicts ultimately private?

When there is a decision to be made, whether to have cake or pie, or a career or a marriage, the choice tends to be visualized as a fork in the road, a path up a decision tree to a future that *diverges* from the point of the present. If a path in the tree is obstructed, e.g., if a gun is put to one's head, options appear to be limited and the will

appears to be unfree. This ordinary concept of choice differs from choice in the mind, which is a lack of resolution of one path from the potential of the many. This is choice as a *convergence* into the present.

From the standpoint of the mental, choice is for latent content that is surfacing into (with) awareness. Choice actualizes partly realized concepts. In the world, choice is an opportunity for action. In the mind, the decision that follows a state of choice is the resolution of the one from the many. Is the future divergent? Is the present convergent? The concept of free will hinges on the answer.

Intentionality and Choice

If free will depends on choice and intentionality is the direction to an object that is implicitly chosen, intentional action may prefigure a volition but is not a critical aspect of free will. An action needs a goal to be purposeful, but this is not enough for the act to be volitional. An announcement of an intention is closer to volition because what is left unsaid implies alternatives. The self that states or is aware of an intention is capable of stating or being aware of an intention to do otherwise. The "I will" is the bridge to volition.

The problem of free will rests on the state of a self in the act of choosing. A decisive self is not preoccupied with choice. Behavior may seem automatic or preset and driven by personal motives or the circumstances in which it occurs. Decisiveness does not exclude a prior state of choice; it does exclude conscious choice at the moment of the action. An individual is decisive when little weight is given to alternatives. Such a person is an example of G. K. Chesterton's advice to keep an open mind, but after having made it up, to close it! Conversely, indecision can paralyse a will fixated on choices. Such differences reflect the segment of becoming with the major impact on personality.

We think of choice as the entertaining and reflecting on more than one goal. But there is a difference between a list of options and a lack of resolution of one option with clarity. When several possible goals, or several means to the same goal, are surveyed in the mind as ideas, the advantages and disadvantages of each course of action can be considered without the necessity of choice. If choice is required, we assume that rational decision is possible. The ideal of philosophical and scientific thought is decision from the standpoint of detachment. This ideal, however, seems possible only if the individual is disinterested in the outcome. Every decision, even to have cake or pie, traces back to a value, and values are not arbitrary attitudes of the self but its fabric.

A decision is not necessarily a commitment to action. When a commitment is required, we often find we are not ready to act on the facts

as known but wait for an 'instinctive' feeling to guide us to a goal. Freud remarked that important decisions come from the heart, not the head. The affect is needed to motivate the intention. We withdraw from rational choice to the affect-laden conceptual antecedents of objects that have already actualized. Without this content, choice with conviction is impossible.[11]

Choice and Context

An object is derived from a manifold to the object it is. The derivation is from the potential of the many to the actuality of the one. Potential is a manifold of *possible* objects. The many are not objects but possibilities. The one is not a singularity but an actual fact. The process is shaped to what it is as other configurations remain unactualized or are eliminated. The subsequent reinstantiation of the object is its persistence and stability, a regularity that obscures the iterated selection among competing forms.

An item is derived from a context. We can say the item is *selected* from a context if by selection we mean the actualization of the item. Derivation suggests a nonconscious process of transformation, selection a conscious process of choice. But one does not choose among items in a context. The context is the potential for an item, not a set of entities from which a selection is made. The shift from derivation to selection involves a heightening of awareness for a process that proceeds more or less automatically. The shift from selection to choice is the awareness that the potential of a preobject precedes its selection as an object in the world. The shift is essential for the feeling that options are available midway between idea and fact; i.e., that choices are grounded in the mind prior to their instantiations in the world.

Choice becomes selection as the derivation "zeros in" on final (actual) items. The range of alternative objects at each moment in the derivation is constrained by the proximity to the actual. There is delimitation of the multitude of symbolic dream objects through the constraints of experience to the members of an abstract class. The restriction of choice to a small set of forming items—choice is usually between two alternatives—may correspond with the number of disparate contents (object concepts) that can be contemplated at a given moment. The capacity to focus on forms competing at the conceptual base of the object-to-be is the essence of introspective awareness and volition.

Microprocess of Choice

The specificity to category is important in choice, which tends to be a conflict between contents in the same category. One chooses between

different ties or deserts (preferences) but not ordinarily between a tie and a desert. There could be a choice between a tie and a dessert if, for example, one's budget were limited, but then it is a question of an overlap between disparate categories (clothing, food) where the categories become member items to an external condition. The overlapping of categories in the binding of disparate concepts by shared attributes, is fundamental in ordinary thinking as well as in creative throught (p. 233).

In cases where separate categories are in conflict, for example to buy a tie or a dessert, or to have dinner or go to a concert, contents (categories) participate in the valuation attached to a given object or period of time. Highly disparate categories can be brought into relation. Thus, I do not ordinarily have to choose whether I will lift my hand or go to a concert, not because it is possible to do both but because these are unrelated actions. Should they be related by another condition, for example, were I required to lift my hand as a request to leave a room in order to go to a concert, and reluctant to do so, the categories would compete as contents in the momentary ideation.

When values are prominent, a choice may devolve on different moral positions, say on abortion or pornography, but not ordinarily on a position for abortion *and* pornography. A choice between values tends to be more coherent than a choice between preferences. A value is an indication of character and a marker for other areas of local valuation, including preferences. We expect values to be consistent across categories. A position on abortion tends to predict a position on issues as diverse as pornography, gun control, vagrancy and divorce. This is because the weight assigned to the balance between the interests of the self and those of the society affects these issues in common ways. This is less true of preferences. An individual feels greater freedom for a decision between preferences than a decision between values. A decision on a value may seem impossible; one could not do otherwise. One's whole being inclines in that direction. The value is bound up with a sense of personal identity. Values have a stake in, indeed, constitute, the self-concept. Free will requires that action is driven by values but says nothing of what those values should be.

The fact that choice tends to be for instances or attitudes within a category, or for different categories related by a common predicate as a linking principle, suggests that in choice the originating concept of an action-to-be realizes a portion of the multiplicity of the context of a category, whether a conventional category like clothing, or an invention of the moment, though a specific content has not individuated.

Categories (contexts) are left behind as contents (items) are specified. A decision to have cake or pie, or skip dessert, is an exercise in free will. I am as free to choose between these objects as between lifting

or not lifting my hand. While I am deciding whether or not to have dessert, my behavior is volitional. Once I decide to have desert, there is no longer indecision and my behavior becomes purposeful. Should I then hesitate over cake or pie, the volitional quality once more comes to the fore. In the feeling of a freedom to decide, to act, or to remain inactive, in the passage of this feeling into decision and its transformation to intention, in the reappearance of indecision, in the shift from one momentary attitude to the next, each state identifies a phase in the resolution of an object out of a conceptual (categorical) background to the threshold of actuality.

NOTES

1. J. Searle, *Intentionality* (Cambridge, U.K.: Cambridge University Press, 1983), distinguishes the (propositional) content of a concept from an object. I take the object (e.g., speech) to be a later stage in the derivation of a concept (mental sentence or proposition). A concept and an object, or statement, are segments in a common process.
2. S. Gallagher, "Body Image and Body Schema: A Conceptual Clarification," *Journal of Mind and Behavior* (1986): 541–554.
3. Brown, *Life of the Mind*, on memory, 335; Brown, *Self and Process*, 127–146, on incrementation and duration.
4. See also *Self and Process*, 113–126; *Life of the Mind*, 316–321.
5. Wittgenstein argued that the description of an intention is a description of an intended action, so that intention and intentional action are not logically distinct events. Moreover, the argument that an intention is not an entity like an image or a thought is consistent with the intention being in the relation between such entities; see N. Malcolm, *Wittgenstein: A Religious Point of View* (Ithaca, NY: Cornell University Press, 1994).
6. D. Davidson, *Actions and Events* (Oxford: Clarendon Press, 1980), 46.
7. J. Piaget, *The Mechanisms of Perception* (London: Routledge and Kegan Paul, 1969).
8. It does not follow that an increase in choice corresponds with an increase in free will. Indeed, the agent can be overcome by an excess of choices. It is the *presence* of choice, not the nature or number of choices available that is decisive in free will.
9. J. Searle, *Intentionality*.
10. In this respect, the inability to perform a morally reprehensible act is a weak example of free will, since inhibitions (values) overpower choice. See P. van Inwagen, "When Is the Will Free?" in *Philosophical Perspectives* vol. 3, ed. J. Tomberlin (Atascadero, CA: Ridgeview, 1989).
11. For Hume, reason determines what is true and what is false but is dumb on the question of what to do: G. Watson, "Free Agency," in *Free Will* (Oxford: Oxford University Press, 1982).

CHAPTER 8

Desire and Value

ARGUMENT: *Drive transforms to desire through the constraints of experience. This parallels the growth of a self in a subject. As drives infiltrate implicit beliefs to create core valuations, the desires are configured by the explicit beliefs to create values that are learned. The belief in a desire is its goal. The desire in a belief is its conviction and realness. Free will requires a mitigation of drive. Agency is not in the service of reason but in choice between competing options.*

FROM DRIVE TO DESIRE

Desire[1] extends willing in a direction toward ideas. The extension is the *feeling* in desire. The feeling is a derivation of drive as the "must" of drive becomes the "wish" of desire. The object of this feeling is its conceptual *content*. The content or goal of the desire is extracted from the self-concept. This content and its accompanying feeling constitute desire.

Every mental state begins with the articulation of the deep self to concepts (implicit beliefs) and drives. As the process continues, a drive that is penetrated by implicit beliefs becomes a desire that incorporates explicit beliefs. Implicit beliefs become conscious, i.e., explicit, in the transition to desire. Beliefs, whether conscious or nonconscious, ground the desires in reasons, goals, and justifications (p. 86). The conceptual element derived through belief from the self-concept provides a goal for the desire. Conversely, desires also penetrate beliefs to create values and, through the valuation of concepts, they provide the affective tonality of internal and external objects.

Drives express *implicit* beliefs while desires are the affective equivalent of *explicit* beliefs. As the drives transform to the desires, implicit belief or tacit knowledge gives rise to concepts, images, and propositions. An intense, brief latency drive that discharges on an object that brings it satiation, shifts to an interior surge of image and feeling. A drive is depleted for a time when its target is confronted. In contrast, a desire is not exhausted even when the object is acquired, and can be sustained in the absence of the object. A desire can be realized over time by a kind of incremental satisfaction. One can say the object that satisfies a drive only "whets the appetite" of a desire.

Beliefs are generated with object categories, so the object of a desire begins as an implicit belief in the self-concept. The infusion of belief into desire, or the converse, creates value. Feelings are a sign of the valuation of concepts. Concepts deposit those objects toward which feelings are directed. The bond of concept (belief) and feeling (desire) is valuation. Desire is an intentional relation to concepts that matter, i.e., are valued. The strength or depth of a value is determined by the extent to which the constituent beliefs and desires share in the self-concept.

A delay in the actualization of a drive allows the drive to undergo an interior development. This development is also a withdrawal from a discharge on objects to the concepts or images behind them. In drive, an object is a target for satiation. In desire, the object is an internal image for fulfillment. The "mental" phase of an object is enhanced in a concept that would otherwise have been buried in an actual object. The discharge of a drive in an object becomes, in desire, the *pursuit* after an absent object or its idea. Desire requires the enhancement of preliminary phases in the object or, what amounts to the same thing, the withdrawal from the final object to its prefigurative forms. Through this enhancement, or withdrawal, anticipatory contents substitute for real objects as imagined goals.

The prominence of antecedent phases in the microgeny of the object occasions a greater selectivity of targets. Drive objects are less specific than those of desire. The gestaltlike targets of drive, e.g., prey, are perceptual *categories*. The instance is accidental. In contrast, desire is for *exemplars* in drive categories, i.e., for a particular object or member of a class. For example, a sexual *drive* is for an available partner while a sexual *desire* is for a particular person. The categories or concepts that are satisfied in drive constitute primitive valuations. These categories individuate in desire to a valuation of specific objects or object features, such as beauty or intelligence. Individuality is unimportant to drive but everything to desire. A drive directed to (fight) or from (flight) a category of objects transforms to the pursuit (anger) or avoidance (fear) of a *specific* object or idea.

The drives actualize a subject. They define behaviors that are valued by a *subject*. These categories of drive-valuation are really value-creating drives that elaborate the subject. The desires and beliefs actualize a *self*. In drive there is no residual personality. A subject is not distinguished from other subjects with respect to its changing drive states. Desire promotes agency by providing objects to be sought after. The affective wish for the object is the seed of intention.

VALUE AND DESIRE

The drive-specific components of the deep self (flight, fight, hunger, etc.) are precursors of valuation. For the deep self, survival is the paramount value. Drives are means to self-preservation, charging conceptual primitives (implicit beliefs) with affect, and transforming them into core valuations. The articulation of the deep self into implicit beliefs, and then the conscious self with explicit beliefs, and the fractionation of drive to desire through learning, are the basis of the more complex values of maturity. For example, the value of a self-preservative drive is the antecedent of a linguistically "infected" value (concept) such as freedom.

The relation between value and desire is complex. An object of desire should be valued but valuation does not guarantee desire. Attributes can be valued but not desired or emulated. We value in others, e.g., humility, what we may not seek for ourselves. We can even value contradictory objects, such as an expansive confidence and a quiet serenity. One can desire a "devalued" object, as in sexual predation or compulsion, a dangerous sexual partner or an addictive drug, knowing such objects are valueless or harmful. Desire and value dissociate. A positive value may be associated with a weak desire, a strong desire with a neutral or negative value.

This bifurcation of value and desire occurs because value is defined as a good determined by reason, i.e., what reason determines "should" be valued, while desire is (treated as) a base emotion or selfish pursuit. If a desire is held to a moral standard, the value that determines the desire becomes a type of judgment. Thus conceived, values are abstract and generic, related to moral norms, while desires are concrete and specific, unique to an individual.

But a value is an expressive content, not a judgment based on an evaluation. The opposition of value and desire is an artifact of defining value as an external good, not a constituent of the self, and of defining desire as directed toward an object, not a feeling bound up with a given concept. Still, value seems to encircle the objects pursued in desire so

a desire that is felt or expressed by an individual is an indication of the self's underlying "value-system."

The fact that desire occurs for acts or objects of little value is important. Desire can engage an action that does not appeal to a value, such as a desire to take a walk or lift the hand. In such instances, the value in the act is inapparent, while the desire to perform the action is like a sudden impulse. Perhaps desire is not an accurate description of this feeling, which is more like a pure volition stripped of feeling and value, resembling very little the depiction of desire as an intense feeling or a passionate longing. The desire for an action of indifferent value, however, is a paradigmatic case for free will. The desire is not overtly motivated by an emotion or valuation, and the action appears to have no cause or incitment other than the intention of an agent.

Still, a sudden impulse to walk or lift the arm traces to an idea in the self. The context of the impulse is the ideational background of the desire. This context might, for example, be a wish to demonstrate that an act is free—I lift my hand simply to show that I can—in which case the act arises from a concept of freedom or agency. Or, the desire to walk might result from a feeling of boredom or discomfort grounded in the body image which, by way of intermediate concepts, derives to statements about the self such as: "I feel like a walk."

In such cases, when the emotional sources of a desire (say, to walk) are inapparent, a delay or impediment to the action can revive this emotive quality. If an individual is prevented from walking or lifting the arm, he or she may become agitated and the desire to carry out the action will intensify. A restraint on the action exposes its affective root (in desire). We then see that the desire to take a walk stems from, say, the value given to an active life, to physical conditioning, or to the concept of oneself as a free agent.

The desire for what is valueless or harmful, whether a high cholesterol diet or a hopeless love affair, is construed as an appetite when it is drivelike, or a habit or obsession when it is coercive or fixed. All desires, whether selfish (drive-expressive) or sacrificial, reflect the outcome of competing values, even if they are subconscious or unacknowledged. We all have a tendency to "take credit" for our unselfish or morally approbational values and relegate our less positive attitudes to outside influences; for example, to traumatic experiences in childhood, for which we disclaim responsibility. In a sense, this is justifiable. The values composing a self are not the self's to decide, nor are they the desire's to decide. It is a question of which set of values (instincts, drives, beliefs, desires, etc.) is dominant at a given moment, what is the content of the set, and how successfully the selfish values have been

replaced by the unselfish ones, i.e., the extent to which drive is tempered by learning or the extent to which desire is penetrated by belief.

From these considerations the dissociation between value and desire can be reassessed. We desire what we unconsciously value. The goal is a subjective benefit even if, objectively, the desire is vain or harmful. A harmful desire is usually motivated by the immediate needs of the self: An urge to commit suicide is, inter alia, the desire for a release from pain; sexual desires can overcome rational fears; a desire that is indifferent to the values of another person, i.e., a victim, such as in rape, is unopposed by these values in the self. A sociopath who is unmindful of the desires of others, an individual whose values and desires are defined by the drives, is the product of an incomplete education of the self-will.

Reason, Desire, and Free Will

One definition of free will is the pursuit after value, or the choice of what is valued by reason when value is defined as what reason determines to be good. Desire appears antagonistic to reason and thus to free will. We say, one is in the grip of an irrational desire, as if free will is canceled by intense desire or irrationality. Strong desires would seem to diminish free will by compelling action or eliminating rational choice. Irrational desires, the impotence of reason as a goad to action, and the opposition of rationality and emotion, are some of the arguments for the independence of reason and desire, and the association of free will with rational action, a distinction that goes back to Plato's division of the soul into appetite and reason. But if values are attitudes, not moral standards, i.e., if values are perspectival like objects, the rationale for a desired action will express what is desirable from the standpoint of the self. A strong or weak desire for an object reflects the self that is mirrored in that need.

Reason or the pursuit of the good are objective criteria imposed on subjective states. They are a measure of the conformity of behavior to a standard, not an explanation of the process through which the behavior is deposited. The belief that guides a free action is not necessarily rational. Indeed, reason can expose the tenuous basis of most every belief and its interpenetration by feeling. The admixture of belief and desire usually presumes that beliefs are rational constructs that are distorted by desire. Yet belief penetrates desire, or the reverse, irrespective of rationality. How else can one explain the justification of irrational desires, the emotions that guide creative thinking, or the way a belief can induce an emotion, e.g., the emotion stirred by the belief one has

been betrayed.[2] Can one say an object is desired because it is believed to be desirable, e.g., beautiful or rare, or is the belief motivated by the desire?

The interpenetration of belief and desire is evident in the resistance of desire or belief to contradiction. It is difficult to convince a lover, or an enthusiast of any sort, that the object of his or her desire is unworthy. This is because desire is driven by personal beliefs, and beliefs are not deduced from facts. The beliefs incorporate core valuations so that to surrender a belief is to sacrifice part of the self. Indeed, the deeper the belief the less rational it is. The "laws" of deep cognition are closer to dreamwork than waking mentation. Reason is achieved as explicit beliefs are shaped by the world. The more natural condition has reason in the service of feeling.

More precisely, reason is an arrangement of statements or propositions. These propositions are preceded and framed by concepts. The propositions discharge the contents of these underlying concepts. The affective tonality of concepts is graded, the more intense the affect, the closer to the core. As concepts fractionate to their linguistic (and other) derivations, the continuity with desire or value is blurred to the point where a proposition can be experienced as affect-free. There is a successive mitigation of the affective link to desire. But explicit beliefs and desires articulate the self to become the progenitors of the concepts from which propositions issue.

The ability to act according to reason implies a decision as to what is good or bad, right or wrong. Actions based on what is good are ostensibly for extrapersonal (objective) reasons. They satisfy a consensus as to what constitutes a good or reasonable act. A consensus is independent of the claims of an individual. Reason dictates that free will should be impartial, i.e., involves a denial of self-interest. Personal desires should be sublimated for acts on behalf of others. The extreme case is altruism, where the interests of others come before one's own survival.

Whether a choice is on the basis of reason or desire does not determine whether it is free. If reason dictates suicide in spite of a desire to live, would a choice to live or to die be an act of free will? What of cases where the choice is between equally rational, irrational, or unpalatable options. I would say that free will depends on the presence of choice *in the mind* and a conscious self that "feels" it is choosing, not on the rationality of the choice that is made. These requirements are not independent since conscious choice obtains only with the feeling of agency.

It is not whether a thought is fantastic or rational that establishes the freedom of an action but the potential for alternate outcomes, in-

cluding the possibility of not acting. A consciousness of different outcomes or the different actions that can achieve a given outcome, where neither the action nor the outcome is of overwhelming strength, *is a state in which reason tends to dominate*. Put differently, a state of choice tends to obtain when the contents of that state are relatively close to actuality, i.e., "realistic."

Delusional or obsessive thoughts express forthcoming acts and not deliberations with options of equivalent strength. In a state of irrationality, I doubt one can entertain choices that are equally irrational or that one can consider two opposing concepts simultaneously. If this is incorrect, i.e., if irrational choices have equivalent weights, free will would apply to irrational as well as rational acts. An individual might freely choose one of two irrational options, i.e., decide between two acts independent of their rationality.

If reason is to determine the correctness of a choice that is made, i.e., the most reasonable alternative, the further judgment as to whether that determination was correct, i.e., whether the judgment of what is reasonable is a reasonable judgment, raises the specter of a regress of decisional states as to correctness. Free will depends on the microprocess of choice, not the typology of choices available in a given circumstance or the acceptabilty of the choices that are made.

Desire as Shared Affect

Desire is a feeling that is grounded in subjectivity but shared between a subject (self) and an object. To say an affect is shared between a self and an object is to say the affect is present in the becoming of two forming objects, one of which is a self. The affect deposits at a depth common to both objects. Or, feeling in the self continues into the object. An intensity of feeling that is usually intrapersonal fills an extrapersonal object. This is the affective intensity of the desired (or feared, etc.) object.

What precisely does it mean to say the affect in an object is shared? A self is a layered concept, as is an object. The interior segment of the self is enlarged by drive, belief, and their valuations. Every object actualizes *through* a subject or a self. The distal segment of this process becomes the object, the proximal segment, the self. Because of this common origin, every object, even as it exteriorizes, retains a portion of subjective feeling. This is the source of value, meaning, or beauty in an object. In everyday experience, these early phases are largely given up in the actualization of a final content. The meaning in the object is implicit in its recognition. What remains of feeling in an actual object

is the reality of its appearance. The final tributary of will is the feeling an object is real. This feeling binds the object to its source in subjective valuation. Valuation is the investment of affect and signification in the reality of a world of indifferent facts, the realness of which is the affective link to subjective becoming.

Self and object have different microgenetic fates and consequently different qualities of feeling. An object of desire is imbued with feeling, for example the love for a person who is perceived to be beautiful. The beauty of the object and the love that is received are, as we say, "in the eye of the beholder." This means the love and beauty, the feeling and meaning in the object, are attributed to an external source but arise in the self of the observer. When an object that was loved and perceived to be beautiful becomes trivial or ugly, the distribution of feeling between self and object changes as a unit. The quickening of desire increases the object's appeal. The waning of desire saps the object of its interest (value).

In sexual *desire*, object knowledge awakens valuation to become explicit as a wish to possess the object and a conviction of its beauty, generosity, etc. To say the beauty or generosity of the beloved is less an attribute of the object than the observer is to say that the self seeks or desires the beloved to replicate or complement its own conceptual valuations. A beloved is a composite of attributes in the self-concept. We value in another what we need or value in ourselves. The features that articulate a desired object also articulate the self that desires that object. That is why the selection of the object reveals what sort of self would select an object of that type.

Anger and Guilt

The inward path of drive creates a subject; the outward path creates an object. The quality of a drive differs according to the direction. This is also true for desire, which is the inward residue of drive. As discussed, a drive is delayed in expression, mitigates and is inwardly elaborated to desire. The microgeny of this transition can be examined in relation to a specific drive such as aggression. In aggression, affect can discharge in relation to an object (fight), or turn inward to desire (anger) through a subjective development. Anger is the narrowing down of the (external) field of aggression to a single (mental) object. Aggression becomes anger with a delineation of the implicit belief that drives the aggression. Anger is desire with hurtful intent, a focused aggression toward a particular object or the idea of the object. Anger can occur in the absence of an

object. It can be replaced by justifications or reasons. Anger can be expressed in a complaint about a wrong the object has inflicted.

Anger that is incompletely eliminated is directed inward. When this occurs, the anger tends to be altered to another affect. One such outcome is the feeling of guilt.[3] Anger toward an object that injures the self can lead outward as anger or inward as guilt. Guilt is one form of self-destructiveness. In anger, an object is perceived to commit a wrong. The concept of the object is defined in part by its aggressiveness toward the subject. In guilt, a personal object or concept commits a wrong and is defined by this injury. In both anger and guilt, a wrong can be real or imaginary. The self is less emphatic in anger than guilt. One could say the interioricity of guilt articulates the self to a greater extent than anger which leads outward. This is another way of saying that anger is more drivelike than guilt, or that guilt, like desire and belief, departs from drive in having a further subjective development. A subject expresses aggression (drive) but a self feels anger (desire), especially when the feeling persists without the object.

Guilt is anger in relation to values. Guilt arises in the opposition to values or desires, even though both have a share in the self. The anger directed to an opposing value results in guilt, which is the negative value assigned to a desire.

Although one can have guilt without responsibility, say for advantages due to birth, guilt implies a feeling of responsibility or at least a feeling that one value and/or desire is not "true" to another. The feeling of responsibility is the allegiance to one of the values that creates the conflict. Values constitute a portion of the self. The self cannot avoid responsibility for its own constitution. Guilt occurs when *values* in the self-concept compete for dominance. Guilt arises in the opposition and the relative success or failure of a value in shaping an action.

Since the same thought or action, e.g., infidelity, can arouse guilt, joy, or pride depending on its "fit" with the prevailing self-concept, or the insufficiency of negating value, guilt is less a specific emotion than a dissonance between values that are incompatible. This dissonance is a sign of irresolution even if the action has already transpired. The desire for the object confronts the anger from an opposing value that would otherwise objectify in aggression toward the object. The absence of "causal" responsibility does not relieve guilt because guilt is not punishment for doing (or thinking) something wrong. It is anger discharging in the self when the development of feeling toward an object (the internalized value) is hindered or attenuated through conflict. The depth and content of the conflict determine the quality of the experience; e.g., guilt, shame, humiliation, embarrassment, etc.

The interpretation of guilt as undischarged anger explains the transfer of anger to the object (e.g., a parent) that instilled the value from which the conflict arose. As anger in guilt can exteriorize in a conflict between the self and another object, say a guilt-inducing mother, so values can "objectify" as cultural norms. The conflict between values can be intrapsychic, as a competition for supremacy of one value over another, or the conflict can be "located" at the interface of the individual with the society.

In an oppressive society, a value such as submission to authority has objectified as a fact or law. Submission to authority is no longer perceived as a personal value but a social necessity. Whether the submission is accepted as an obligation, i.e., a social norm, or protested as a compulsion, depends on the degree to which the value has infiltrated the self-concept. A value that objectifies as a social norm still traverses the self-concept of an individual reared in that culture. For such a person there is always a potential for intrapsychic conflict. There may be guilt or anxiety over the submission to authority or its opposite, the pursuit of individual freedom. Perhaps the latter is less likely to induce guilt than the former because self-expression is the more natural (instinct-driven) behavior. In any event, the healthier, i.e., "truer," condition would seem to be one in which conflicts are apprehended as intra-personal rather than between the individual and the society. This is because the competition for primacy as values specify to acts and objects promotes a contextual richness that is the basis of an informed moral conscience.

VALUE AND BELIEF

Valuation is established by the drives. The shift to desire accompanies a growth of concepts in the self. This growth is the penetration of desire by knowledge (belief). Every desire incorporates a belief, or concept, that actualizes as a goal. The self cannot desire an object it does not know, and the object that is desired traces back through explicit belief to tacit knowledge. Concepts in the self lead to objects that satisfy in desire their initial valuation.

The knowledge that counts in desire is based on the valuations of implicit beliefs in the self-concept. Desire may increase with an increase in object knowledge if what is learned accords with the beliefs that instigated the desire. Knowledge can also extinguish desire with the discovery of facts that contravert beliefs, or through a diffusion of desire into the object and away from the driving concept, as when a passion becomes a hobby or a friendship. In such cases, the "true" nature of the desired object is forced upon the self. The object may be rejected

to safeguard the self-concept. Otherwise, the self-concept will undergo gradual change to accommodate and retain the object. We say, "love is blind," because desire is not based on object knowledge, and we say, "familiarity breeds contempt," because desire can be eroded by a knowledge of the object that conflicts with the needs or values of the self.

Linguistic objects do not have the same factual status as external events. Propositions are derivations of implicit beliefs. They are not the contents of such beliefs, which are broader than propositions. Because of the openness and inventiveness of propositions and their proximity to the self, i.e., unlike an external object, a proposition is felt to be a personal object that is created by the self. The affective tonality in conviction tends to increase with increasing knowledge, for the knowledge accumulates in such as way as to validate the original belief.

Take for example the concepts of freedom and privacy. These are mature values that define a portion of the self. Such concepts have an affective charge, which is value. These concepts can surface as explicit beliefs. A belief with its complement of affect can lead to a statement (proposition), for example that abortion (or pornography, firearms, etc.) should be legal. This statement is an object that grows out of the general concept of free self-expression. Knowledge reinforces such a belief, that by virtue of its valuation guides the search for confirmatory fact. Such knowledge, developed in relation to personal beliefs and desires, is forced upon events. The self shapes the objects it encounters. That is why beliefs can be "unshakeable." Conversely, a "disinterested" self that seeks knowledge in a humble "unbiased" way may have difficulty taking a stand on one side of an issue since every argument has some validity.

This process is independent of the value content. The shift from a value based on drive to a value realized in desire, and the role of belief in the targeting or withering of desire, is revived in every mature value. By mature value is meant the articulation of desire by a knowledge that is sufficient to provide options for action. For example, the desire to obtain, contemplate, support or protest an abortion, when it is an outcome of choice, is a manifestation of a mature value or the supremacy of one value over another in the self-concept.

There is an indistinct boundary between beliefs, values, and desires.[4] Every desire, including fears, needs, wants, and wishes, expresses a value and a belief. The value is partly the personal value to the subject, i.e., why an object is important or necessary, and partly an "extrapersonal" value that is the signification or meaning of the object. Implicit beliefs provide a context for desire, e.g., the belief that the object exists, can be obtained, will satisfy the desire, etc. Explicit beliefs constrain the plans and goals of the desire. For example, the desire for an apple contains a personal value; the apple's value to the agent as an object to

be eaten, painted, or juggled, and the intrinsic value of the apple; its dictionary meaning or function "independent" of the agent. Belief surrounds desire as a framework for action. The nature of the apple, its edibility and categories of belonging, are tied to the value of the apple and the agent's desire for it.

Another way to characterize the relation of knowledge or belief to desire is to say that desires are infiltrated by beliefs about objects that matter, have value, while beliefs are penetrated by desires to give feeling (value) to knowledge. Values express interest according to what is guiding behavior at the moment. A value announces a belief. A belief orients a value. Value is what is largely personal in belief; belief is what is partly impersonal in value.

An intensification of feeling transforms belief to desire. The *desire* to turn on a light dominates belief when it derives, say from an excessive fear of the dark. The affective transformation of "scientific" concepts or abstract beliefs is the basis of magical thinking and pathological states. Abraham's belief in a punitive God explains (becomes) a desire to sacrifice his son. If the conceptual element in desire exteriorizes, it becomes a belief. The desire for immortality, or the fear of death, becomes the belief in an after-life.

The notion of values as *objective* is a form of acquisition of beliefs. A value is a "should." Does this exist in nature? A value is a ground for behavior. For the value to be objective it has to be grounded in nature, as a type of fact. Facts are derived from beliefs that contain value as the realness of their derivations. One can say that beliefs are values assigned to what we know about the world. Values are beliefs assigned to what we know about ourselves.

The contents of belief are concepts. A hierarchy of realness (experiential truth) is assigned to concepts. Value is covertly present in the commitment to a belief. The subjective value in belief is conviction, while the objective value is the realness of the object the belief is about. The feeling of reality is part of the belief that apples are red, not blue. The apple and its redness have to be real. If the reality of an apple is in question, the belief (conviction, certainty) that an apple or its redness exist cannot be sustained. One cannot support a conviction about the truth of a particular belief if the belief, or the object of the belief, is *felt* to be unreal.

Similarly, the truth of a proposition can be examined only if the reality of its terms is assumed. The terms of a proposition are not just the words that constitute the propositional content but the entities to which the words refer or what they signify. The truth of a statement is its reality. The *determination* of truth is wedded to a system of relational concepts—how well a conception of the world "hangs together"—but the feel-

ing of reality or the *conviction* of truth is an affective limb of value distributed into objects. Indeed, values are the kernels of propositions.

VALUE AND LEARNING

J. S. Mill said, "the object of moral education is to educate the will."[5] The will has to be subdued, the drives diverted, values instilled, and beliefs learned. All of this must occur for desires to have goals that are worthwhile. How does the education of the will occur?

The aim of an education should be to infiltrate the drives with shared values and inculcate a set of desires that balance the needs of the individual with those of the group. Values are instilled by instruction. A society must decide the values to be taught its young. Without a proper moral education, the learning of values will be guided by the precursor drives. It is essential to instill good values in children so they will act responsibly as adults. The concept of individual moral responsibility for one's actions serves as a value in itself even if this concept has limited application to the problem of free will.

Take the development of infantile sucking to erotic desire. The sucking instinct of the infant is triggered by an object (finger, nipple) in the appropriate body region, and discharges in rhythmic axial movements organized, at least in part, in brainstem (sucking is preserved in anencephalic infants). A specific object (breast) gradually exteriorizes as a target. Instinct transforms to drive, then to desire for an object that can later be wished for. The sexuality nascent in appetitive behavior transforms to a desire for erotic targets. This is one basis for the later preference for certain acts or partners. Sucking in the infant is an example of instinctual feeding and a forecast of primitive sexual activity. Hunger transforms to a desire for nourishment, then a preference for certain foods. The class of edible objects differentiates within the releasers of the nursing instinct. Every specification is a distinction, every distinction is a categorization.

Primitive values in the drive categories of a subject are concepts for the deep self. The instinct for the breast represents a configuration of infant and breast as a unit. Act and object are one. In the derivation to drive, this configuration fractionates a subject and an object. The breast exteriorizes as the infant gains autonomy, i.e., is distinct from the objects of its drive states. The separation of the object (breast) anticipates its pursuit as a goal, and the planning and incrementation of acts.

In the learning of values, experience sculpts innate valuations. The infant explores its linguistic community. The word *no* dictates which object or action is maladaptive. The mind of the child seeks a niche in

the surround. A "no," a scowl, is a path cut off; a "yes," a smile, a confirmation of a direction. The child's action is not just blocked by a "no" but diverted to less disruptive or more congruent avenues of discharge. Inhibition enhances covert phases in mentation. Self and space articulate in conformity with the world.

Valuation is the making of categories of feeling in the context of an assertion of self. The category expresses the self guided by the drives. The microstructure of value is the process of concept formation in relation to personal memory and feeling in the partition of a drive-based *subject* to a desire-based *self*.

Learned beliefs can mitigate the effects of drive on value, but there can be excess in this direction as well. An intense desire that is driven by values that are unselfish and affirming, e.g. the love of God or a patriotic fervor, may lead an individual to become a zealot for a cause. In this instance, the value draws its intensity not from the drives but from a usurpation of the self by a consuming belief. The concept in the belief replaces a portion of the self-concept until it (the self) finally identifies with the value.

Consciousness and Value

A desire is a downstream fragment that explicates the deep or unconscious self. A value is marked when a distinction is made by feeling. Similarly, the feeling in every distinction points to a valuation. A process of unconscious valuation is constantly going on. The conscious values one proclaims are not necessarily the values that matter. Conflict is a sign of the subconscious root of valuation. The individual who "says one thing and does another" has an inconsistency between opposing values that actualize through different paths (language, action). When a desire conflicts with a value, e.g. a personal want and a moral obligation, the conflict stems from competing values that may be nonconscious, e.g., between the expressed desire in a subconscious value and the expressed value in a subconscious desire. Thus, a sexual desire that overcomes an opposing value expresses another covert value. An individual may communicate in statements a portion of a valuation that is not exhaustive of core values in that category. We all know people who are selfish yet proclaim their self-sacrifice. A profession of compassion can conceal a deep-seated misanthropy. If language did not disguise more than it reveals, politicians and psychoanalysts would be looking for other means of employment.

The learning of values is decisive in whether an act is "free" in the sense of rational, but not with regard to the content of the act. The

content does not aide in the determination of free will. Acts of moral correctness are not in some sense more free. It is the process leading to the content, not the content, that is instrumental. Double[6] put it nicely, "What increases freedom of the will is the rationality of the *process* of decision that the agent performs; but the rationality of the process does not imply any particular moral *content* of the decision."

Society should promote the individuation of behaviors that are mutually supportive, but not to the point where individuality is sacrificed for the sake of order. Since innate dispositions and experiential memories in the self are the source of all concepts, the learning of values is, arguably, the most vital part of an education. Values decide what concepts are important. The education of values, and the responsibility of each individual to participate in the educational process, are themselves values that society must instill. From a theoretical standpoint, the attainment of morally sanctioned values is an educational concern, not a personal accomplishment. An individual deserves neither credit nor blame for being the outcome of a learning process.

NOTES

1. *Desire* is a generic term that includes wants, hopes and wishes, likes and needs, as well as such adversive states as dislikes, fears, etc.
2. C. Calhoun, "Cognitive Emotions?" in *What Is an Emotion?* ed. C. Calhoun and R. Solomon (Oxford: Oxford University Press, 1984).
3. Some of the issues in this section are discussed in P. Greenspan, "Subjective Guilt and Responsibility," *Mind* 101(1991): 287–303.
4. The explanatory linkage of belief and desire as argued in the text is essential to process theory, yet it is oddly critiqued as a "technical" objection to behaviorism by J. Searle, *The Rediscovery of the Mind* (Cambridge, MA: MIT Press, 1992), 34.
5. J. S. Mill, *On Liberty* (New York: Liberal Arts Press, 1956).
6. R. Double, *The Non-Reality of Free Will* (Oxford: Oxford University Press, 1991), 68.

CHAPTER 9

Belief and Conviction

ARGUMENT: *Beliefs develop out of the deep self. The conceptual element is emphatic in belief, the affective element (will) in desire. Knowledge is shaped by* core *beliefs and valuations. Action is structured by implicit beliefs, which include experiential and world knowledge. Explicit beliefs are action equivalents of knowledge when truth judgments are required. Conviction develops in the derivation of value to a feeling of reality that accompanies the actualization of objects as facts.*

THE GROUND OF INTENTION

The strong interpretation of intentional actions is that they *emanate* from beliefs or desires. I move my hand because I want an apple. I press on the light switch because I *believe* the action will turn on the light. When desires, beliefs, or reasons are associated with ensuing actions, not just contiguous with them but grounded in a single concept, the precedence is construed as a causal relation. Beliefs are commonly held to play a causal role in action, even if the detail and microstructure of this relation has not been the focus of philosophical study. According to Bogdan,[1] a critical failure of philosophy is that "the very essence of belief, that of being a function from mental representation to causal role in cognition and behavior, is left totally unexplained" Nor, he goes on to say, is the failure much noticed.

Less strongly, actions *contain* beliefs and desires which exert, at best, a constraining effect. A belief that is an accessory to an action is weakly causal in delimiting its scope. To the extent a belief constitutes a portion of an act, or is an act, it is not causal *as a belief*, but is an

effect of a prior cause that is not a prior belief. This is implied if beliefs and desires are intentional acts that are replaced by other intentional acts that entail movements. From this point of view, a belief might only announce a concept, the ensuing portion of which constitutes an action. Whether they are interpreted as causes, contents, presentiments, or adjuncts of actions, belief and desire dominate the philosophical landscape on intentionality, so it is necessary to consider their status in relation to microgenetic theory.

Actions express concepts and feelings. At least a concept and a feeling can be found in every (self-generated) action. It is important to clarify the concepts that constitute beliefs and desires and the feelings that permeate them. An account of the microprocess of belief and desire is essential if their intentional structure is to be mapped to the initiation and structure of an action. Is intentionality a secondary or derived phenomenon? Are beliefs and desires types of (incipient) actions? Do they motivate or instigate actions? Are they instantiations or realizations of actions, or justifications of actions generated in some other way?

BELIEF

An action can appear to be driven by a belief, such as the belief that pressing a switch will turn on a light. An action can appear to be driven by a desire, such as the desire to illuminate a room. A desire is usually conscious, though there may be subconscious ("repressed") desires. It is not a simple matter to access a desire that is unconscious, whereas unconscious beliefs reach awareness with greater facility.[2]

A belief can be conscious (explicit) or unconscious (implicit). Unconscious (core) beliefs are not copies of conscious beliefs. They are not identical to conscious beliefs, but nonconscious, and then accessed into awareness. There is a qualitative difference between beliefs that are conscious and those that are nonconscious. This is also true for the distinction between drives and desires.

It is questionable whether the *content* of the belief is effective in action. What matters is the process leading to the content, not the content that the process deposits. This is counter to everyday experience. One is *conscious* of the content of a belief, not of believing as a process.[3] But from the standpoint of microgenetic theory, the process delivers the content that, as it becomes conscious, informs the agent of the belief. A belief-content is a partial and explicit product of a deeper implicit belief with a scope that is wider and in some sense more basic. The explicit belief that a light will go on when a switch is pressed is an expression of the implicit belief in a world of real objects, and real

actions that affect those objects. The implicit belief shapes the process that leads to the explicit content. What, then, shapes the implicit belief?

Every object contains a set of core (implicit) beliefs: that it is an object, that it exists, is real, perhaps that it is animate. There are also secondary implicit beliefs that are specific to the object—how it behaves, is made, used, and so on. The implicit beliefs constitute a core of object knowledge that is foundational. Other beliefs are secondary or contingent. Foundational beliefs are coordinate with the drives. The secondary beliefs, both implicit and explicit, are coordinate with the desires. The foundational beliefs are essential to life, the secondary beliefs to occasions of living. Both types of object knowledge can be separately disrupted by pathology. The foundational beliefs are disrupted in psychosis, those specific to objects are disrupted in agnosia. Both sets of beliefs participate in the selection of an action but are presumably not instigatory. If a portion of the knowledge of an object served to instigate an action, the question would arise, what is selecting that portion?

If belief is crucial in prompting, driving, or guiding a willed action, it is not the *abstract* knowledge of the object that is primary, i.e., facts about the object in a system of world knowledge, but the *functional* knowledge that incorporates a value and a goal, such as the belief that turning on a light will brighten up a room. This belief is part of the general store of knowledge concerning the nature, class, behavior, location, etc., of the object grounded in personal experience. At other times, a different knowledge base in the same object concept, or a different locus of overlap with other concepts, can be accessed, as when a light switch is used as a hook. The different attributes of an object describe its possible uses or functions. These attributes structure the selection of the action but it is unclear how they are engaged in the initiation or selection process.

Belief and Desire

The affective component in belief includes the commitment to the belief and a feeling for the reality of the content of the belief state, i.e., the feeling of the truth or reality (existence) of an object. We know that the reality of a belief has a component of feeling from the study of psychotic cases, where a withdrawal of affect from objects—actually a withdrawal of the object—undermines the feeling of certainty for the object experience. This is the basis of derealization and depersonalization. The withdrawal of feeling from (with) the object corresponds with a change in the belief, e.g., as when an ordinary object takes on menacing features.

In such cases objects reclaim those properties embedded in concepts that preceded their realization. The regression from an object in the world to a concept in the mind accesses phases "in the mind" through which the object developed. The cost of this reclamation is the loss of the feeling of reality. The individual doubts or is uncertain of the realness, i.e., solid existence, of animate and inanimate forms. At the same time, reasoning and object knowledge may seem unaffected or minimally disturbed. This affective tonality is the terminus of valuation. Realness is the gift of value to a concept that achieves actuality as an object in the world.

Every object is shaped by foundational beliefs and contains a portion of the valuation in the self. Conversely, the self, the drives, and their valuations are the seeds of belief and desire. Belief develops out of personal knowledge to object concepts that require a commitment as to truth. This commitment actualizes in a real object. The object is no longer felt as *my* object but as an independent entity. It is no longer a felt image but has an affective life of its own. This affective component includes the feeling of the reality of beliefs and belief statements. Belief and desire terminate in object valuation, in realness or personal meaning depending on whether objectivity or subjectivity is the primary mode of development. In both instances, a common thread leads from the object back to an origin in the self.

Explicit Belief and Truth

A belief is a concept that can be understood in terms of truth and falseness. Truth concerns what is real, what is actually the case or exists. The concept of a blue apple is based, inter alia, on the concepts of apple (fruit) and blueness (color). This concept can be entertained without a commitment as to whether blue apples exist other than as ideas. When this concept is posed in relation to whether there are such things as blue apples, it becomes a belief. Knowledge in the form of statements becomes belief when its basis in truth or reality is questioned. Belief conveys the assurance, or certainty or conviction about the truth or falseness of knowledge; i.e., its basis in the real world. One does not ordinarily act on unreal (known to be false) beliefs. The feeling of realness is essential to conviction and action.

Beliefs are concepts that require a judgment of truth. This judgment, in my view, is not achieved through reason but through the experiential context of the belief, and/or the *feeling* of reality for the concept in question. The context is the coherence of a given belief within the belief system and the adaptation or fitness of the belief to

the world. The relation of one belief to another, and the relation of the system of personal beliefs to perceptual experience, are the only guarantors of truth. The coherence between belief and experience results from sensory constraints on object formation that provide a continuous lifespan barrage of environmental trials. A correspondence to reality is the condition of best fit. The apprehension of this fit, i.e., the awareness and feeling that one's beliefs correspond with conditions in the world, is our closest approximation to reality and the scientific standard of truth.[4]

Belief is an active or assertive form of knowledge, knowledge a tacit or undeclared form of belief. Knowledge is wider than belief because it incorporates subconscious skills that remain underived to consciousness, e.g., the skill of driving a car, the "rules" of a grammar, the sequence of movements in putting on a necktie. Most knowledge is unconscious. We know more than we can possibly express, and we believe the truth of what we know even if the bases of these beliefs are inaccessible.

Implicit belief provides a ground for the continuous elaboration of personal and object knowledge. Actual or explicit belief is knowledge with conviction where the reality of the content is at stake. Belief can take the form of a statement, e.g., "Blue apples do not exist," but there are other kinds of beliefs apart from statements. If an animal declines to eat a blue apple, could one not say the animal does not believe the apple is natural (of nature) or real?

When tacit knowledge[5] becomes explicit, it may pass to uncertainty, e.g., on saying, I believe that blue apples do not exist but I'm not sure. In such instances, the belief includes the uncertainty as part of its domain of knowledge. The uncertainty in this example arises through insufficient knowledge, but conviction of the truth of a belief can occur where knowledge is just as insufficient; e.g., "I am convinced she loves me" or "I am certain there is a God." Faith is belief without evidence, a leap beyond reason to conviction. But aren't most beliefs accepted on faith? Certainly, belief is not just a function of how much one knows.

We have the idea that belief follows knowledge as a kind of deduction from what is known. However, some beliefs appear to be prior to knowledge. By this I do not mean instinctual or biological knowledge, but categories such as personal existence, the present moment, and the boundary between the self and the world. Perhaps these are what have been termed *animal beliefs*[6] in contrast to language-infected opinions. Such core beliefs are a guide to experience. They constrain the knowledge that is acquired, so that learning will be in conformity with the preexisting belief state.

Regarding Moore's statement about his hands,[7] Wittgenstein says that to doubt I have two hands is unimaginable. With this belief, he goes on, "I have arrived at the rock bottom of my convictions." This does not mean this conviction is more certain than others. There are cases where a hand has been severed, e.g., in the case of Wittgenstein's brother, a pianist, and the individual still feels the limb is intact. With eyes closed the person may believe he still has two hands. In some cases the phantom hand can overpower the visual evidence of its absence (p. 64). The individual has to weigh the feeling of the hand against the visual fact. It is a question of which perceptual modality will dominate. Incidentally this raises the question of whether some beliefs are specific to a modality, e.g., to visual or auditory perception.

People with brain damage may deny their hand is their own. The hand can be lifted before their eyes, yet they insist that it belongs to someone else. A patient of mine was asked to touch his shoulder and pointed to the door. There was no misunderstanding of the request. He accurately defined the meaning of the word *shoulder*. One could say that the body image was disrupted. Still, the limbs were used appropriately. Such cases show the vulnerability of "rock bottom" beliefs.

There is nothing more fundamental than a belief in the existence and stability of objects. This belief is implicit at every moment in waking behavior. The belief that apples exist is independent of what I know about apples since everything I know about apples only reinforces this belief. If a particular apple were demonstrated to be, say, a holographic image, my belief in its existence could only be shaken if the image of the apple was not fully objectlike in its appearance, or if its relation to other objects was unnatural. If the holographic apple was perfectly contextualized in the external world, I would have to concede that the entire world, including my own body, could also be (nothing more than) a holographic image. Since I could not believe such a thing and remain sane, my belief in the existence of objects, myself, and the world, takes precedence over the knowledge I have acquired about them. Belief is not a deduction on the basis of knowledge, rather, knowledge accumulates as a confirmation of certain foundational beliefs.

KNOWLEDGE, BELIEF, AND CONVICTION

Is knowledge a form of belief with conviction? To know something is not to doubt it. A great deal of what we know is unquestioned belief. Such knowledge is implicit belief. I am convinced I know what an apple is until I am confronted with a blue apple. Suddenly, my knowledge of apples, or at least an attribute of apples, their coloration, is thrown into

doubt. In what sense is knowledge with uncertainty still knowledge, or does the uncertainty change knowledge into (dis)belief? Wittgenstein[8] says, "One can say, "He believes it, but it isn't so," but not "He knows it, but it isn't so." The questioning of what we know turns knowledge into belief, and belief can be true or false, or certain, or lack conviction.

Belief without conviction, therefore, is a type of uncertain knowledge. Knowledge that is uncertain is decontextualized from an existing knowledge base, either because of insufficient information or unprovability. (This presumes that conviction follows on a proof or on sufficient information. When is information sufficient for conviction? Is a proof something other than additional knowledge? In what sense do proofs confirm?) To know something "for sure" is for belief to pass into knowledge. I learn that a scientist has created a blue apple. Now, I do not just believe there may be blue apples, I know that blue apples exist. I could be misinformed of this fact. My knowledge could be incorrect, but it would not be disbelieved.

The "I believe" is weaker than the "I know." In explicit belief, the certainty of knowledge is in danger. A category of knowledge that requires a conviction is a belief that becomes explicit. Or, it is an occasion when a probability enters an expectation that an instance may deviate from the norm; e.g., I *know* that lights generally go on when a switch is pressed, but in this instance, I do not know, am not certain, the light will go on but I *believe* that it will. Knowledge is an average over instances of encounters with an object. For any instance, knowledge becomes equivalent to a belief; or, belief is an assertion of the correctness of the knowledge. Explicit belief is the selection or instantiation of a sector of knowlege en route to actuality where uncertainty reflects the incomplete (or inconclusive) specification of a concept out of rival potentialities.

CERTAINTY

Certainty and uncertainty apply to belief, not to knowledge, and in relation to belief they have a kind of equivalence. They are different attitudes toward the (same) belief content. A given content, e.g., the existence of blue apples, can be accompanied by certainty or uncertainty. But can one ever be certain that blue apples do not exist somewhere in the universe? There is room for uncertainty in every belief. Doubt can enter any thought or action. Even implicit beliefs about the reality of the self and the world can be shaken. Yet life goes on with conviction. This is because conviction is not a matter of the most plausible alternative given the knowledge at one's disposal; rather, conviction is a

product of the emotional bases of implicit beliefs (tacit knowledge) arising in the core of the personality. Any intelligent person could be pursuaded that most any belief is true, false or indeterminate; but the *conviction* of the truth or falseness or indeterminacy of a belief is not wholly a function of its truth or falseness. Conviction does not follow a judgment as to truth, but is a *feeling* based on the definiteness or actuality the belief achieves.

Belief with certainty carries with it conviction, but there are individuals who are convinced of the truth of false beliefs. Patients with a language disorder may name objects incorrectly yet be convinced the misnaming is correct, even when they are informed of the lack of correspondence between the word and the object. Patients with a memory disorder may be convinced of the accuracy of their misrecollections or fabricated accounts of events. This is not due to a loss of memory for the events, since future or non-occurrent events can also be affected. The potential for recovery shows that the knowledge was not lost.[9]

Hallucinations (and dreams) may be accepted as real events. The subject believes in the reality of the hallucination or dream image because there is no object world for comparison. This accounts for the experience of patients with cortical "blindness" who believe they can see. They have lost external visual objects and the internal precursors (images) of those objects are the only visual objects they have. Patients with delusions have false beliefs with an unshakeable conviction as to their truth. The fact that delusions coexist with normal beliefs suggests that conviction is not applied to a proposition (why would it be selectively misapplied?) but is bound up with the actualization of the propositional (or perceptual) content. More generally, such cases indicate that the truth of a belief, and a conviction as to its truth or falsity, are not necessarily correlated. Conviction is more a matter of what seems real to the subject at a given moment.

Certainty for false beliefs is common in pathology; uncertainty is rare. Brain-damaged patients usually do not fail a task by saying, "I don't know," even for information beyond the normal range of knowledge. This suggests that certainty and uncertainty are not polar states, that the opposite of a true belief with conviction is probably a false belief with conviction, not a lack of conviction for a true belief. The polarity is between true and false with conviction invariant, not between certainty and uncertainty or conviction and lack of conviction.

More likely, uncertainty is a form of indecision where a concept has not developed with the clarity and feeling needed for conviction. Uncertainty is a state of choice in relation to objects, indecision a state of choice in relation to feelings. The lack of uncertainty in pathological cases reflects the intolerance of patients for a state of indecision. This

may point to an incapacity for choice. Since choice is fundamental to free will, such patients, it could be argued, are not truly free.

What is the relation between uncertainty as a lack of conviction, and the doubt that characterizes a tendency to disbelief? Doubt is the feeling that a concept is false. Uncertainty over the existence of blue apples reflects the possibility that somewhere blue apples may exist. Doubt cancels this possibility as feeling tends toward commitment. In uncertainty, no single option resolves or is selected with sufficient distinctness on which to base a truth statement. In doubt, there are grounds to disbelieve. But there is more to uncertainty than indecision and more to doubt than disbelief, or a lack of conviction. Conviction concerns, vitally, the role of feeling in rational thought.

Doubt seems to follow belief because it supposes a belief in an opposing concept. I doubt there are blue apples because I believe all apples are red or green. Could doubt precede belief without contradictory beliefs on which to base the doubt? Children tend to believe what they are told. The system of relations that is the ground of knowledge is maturing and can incorporate almost any fact except those that would contradict the most fundamental core beliefs. A child will believe that fairies exist but not (have such questions been put to a child?) that he or she does not exist. Conviction is the natural condition, doubt develops later. Implicit beliefs structure knowledge to create a system inherently certain because it is the only system the child has. Once this system matures, dissonant facts can generate doubt, but the whole system, or events coherent within the system, cannot be doubted.

One might object that certain facts, say mathematical proofs, achieve truth through logic rather than a feeling of realness. The word *real*, in this context, however, does not mean objectively real, but what is real to the observer. What is real is relational. The relational world can be perceptual, linguistic, or mathematic. A logical demonstration of a mathematical proposition can give the truth of the proposition but not a conviction of its truth. The conviction is in the feeling that accompanies the demonstration. Mathematicians have commented that a proof *feels* right before it is logically demonstrated.[10]

In a state of belief the self is committed to a direction. The direction is the definiteness a concept achieves as it resolves out of tacit knowledge into the field of possible beliefs (concepts) related to the actualizing state. The judgment of reality on which perception depends becomes explicit when a truth (reality) is to be decided. The "becoming explicit" is the assumption of a direction in the context of an implicit decision about the realness of a statement.

KNOWLEDGE AND VALUE

In relativity theory, the relations between objects are their perspectivity from the standpoint of a reference. An object is also a perspective from the standpoint of an observer. Science assumes that different perspectives can render an object objectively, i.e., that the perspectives are independent of the object, or that an object can be subtracted from the redundancy of its perspectives.

The subjectivist position is that all knowledge is perspectival; knowledge approaches but never achieves objectivity. From a deluge of perspectives a representative (contextualized) object is elicited. An objectivity centered in the context of an object leads to a richer account of the object, even if the perspectival basis of knowledge is an argument for the relativism of all objectivity. The contextual account does not gather up perspectives to construct an object by consensus, since all perspectives are ultimately private. Nor does the account sever the perspective from the object after it has done its work. The knowledge of an object is an exploration within the perspective, since any perspective followed deeply enough into the subjective is a path to the source of the object.

What is the relation of value to an account of knowledge that is supposedly value-free? If knowledge and the objects of perception are derived from conceptual valuations, the difference between an objective fact, a value-free idea, and a subjective value depends on the extent to which the relevant attributes of the self actualize. Value is latent in every "objective" fact. A chair can become a sacred object. An indifferent event in the remote past can be infused with personal meaning. This infiltration by value of "neutral fact" occurs because values generate concepts that exteriorize as objects. The "assignment" of a value to an object is an uncovering of the value leading to the object. The "distortion" of objectivity by feeling is a resurgence of feeling in the revival of concepts that prefigure the object. Objects barren of feeling actualize from layers rich in subjective value.

Objects and values have a locus in personal and public space. Rules of conduct are objectified values. The feeling of reality is a subjectivized object. A value shapes an object, an object realizes a value. An object is a type of valuation, a sampling of what reality has to offer. We live in a world of other objects that seem invariant from one observer to another. To *understand* that objects are perspectives is a leap of the imagination. To *feel* they are perspectives is a lapse into psychosis.

DO BELIEFS CAUSE ACTIONS?

An explicit belief can explain an action but is the action caused by the belief? If I press a switch because I believe the action will cause the light to go on, what is the role of the belief and the "because" in generating the action? If the belief (reason) is causal, so is the because. The because is a nexus from the belief to the act. It is the connecting element in explaining the *reason* for the act. But a because can apply to voluntary and involuntary acts. I can say, "I ranted and raved because I was out of control." In this case, the because is explanatory, yet not causal. Being out of control is not a cause of ranting and raving but a description. A because can be a causal link to action, e.g., I ate the apple *because* I was hungry; or the cause can be the equivalent of an action, e.g., I raved because I was angry, where raving is the display of the anger. Moreover, reasons and becauses usually follow actions. One does not often say, "*I will eat an apple*" because I am hungry, and then eat it; rather, one says, "*I ate the apple* because I was hungry" in response to an inquiry after the action is completed. These observations suggest that the because which links an explicit or conscious belief to an action is interpretive, not explanatory in a deeply causal sense.

What then of the belief? If a belief (or desire) is a cause, what is its cause? Is there a regress from act to (conscious to implicit) belief to self to whatever is prior to the self? If a belief is causal, what is its relation to the agent that has the belief? Does the agent produce the belief which then causes an action or is the belief the causative component of the agent? Can the agent (or belief) cause an action contrary to an (effective) belief (or agent); i.e., are there two potential causes for any action, the agent and the agent's belief, Or do agents and their beliefs have different causal roles and/or histories? Can a belief that is ingredient in the self, one of many beliefs composing the self-concept, instigate an action that is realized as a global expression of the person? For a belief to be causal entails that a part-content is capable of inducing a total response of the organism. This raises the question of how an "element" in mentation, i.e. one of the self's beliefs, can engage the entire action system. Conversely, if it is not a belief but the self that causes an action, how does belief participate in causation? These are but a few of the questions that need to be answered for the proposition that beliefs *cause* actions to be seriously entertained. These problems, incidentally, pertain to implicit as well as explicit beliefs.

However, with explicit or conscious belief there is the additional problem of an earlier subconscious phase. This is because the causal sequence is not *from* a state of belief *to* a state of action, i.e. not across two states which are separate moments of time, but from depth to sur-

face in the mental state, i.e., from the antecedents of the belief to its conscious realization in acts and percepts. A conscious belief is an outcome of a becoming, thus a product, like an action, and comparable to an action in being generated through nonconscious phases. If explicit beliefs are products, they are action equivalents. If so, the problem of causation of belief is identical to that of agent causation, i.e., agents produce *both* beliefs and actions. An explicit belief is a verbal act that expresses part of an underlying concept. This concept underlies the belief and the motor act, whether the belief actualizes before or after the action. If after, the belief could be a rationalization, a justification or an explanation of the action. If before, the belief is not transparently causal but could be an expression of the same concept from which the action is generated.

That beliefs may not be causal in behavior is demonstrated in the phenomenon of confabulation. This is the fabrication of reasons for behaviors when the memory of their "causes" is impaired. For example, an amnestic patient asked where he is might respond that he was kidnapped and is being held for ransom. A confabulation is an attempt to provide a reason or explanation for a behavior for which a cause needs to be found. If the patient is confronted with the inaccuracy of his account, e.g., if he is told he is in a hospital, he may adjust the account to incorporate the inconsistencies, e.g., he might say he was kidnapped by a group of doctors. The content of the confabulation can be manipulated by the examiner. It is not fixed, and may change spontaneously on requestioning. Although it can occur for future events, confabulation is generally retrospective and invented "on the spur of the moment." Such phenomena argue that beliefs do not have a causal role.

A confabulation can be considered a false belief with varying degrees of conviction. Since the confabulation is not contextualized in a "belief system," the individual can be persuaded to modify its content. In the paramnesia of psychotic patients, confabulation transforms to delusion and is embedded in a system of false beliefs. In such cases the belief, such as the delusion of being kidnapped, is unshakeable and resists attempts at contradiction. The link from confabulation to delusion, and the probable relation of both to limbic pathology, suggests that reasons in normal people might resemble delusional contents integrated in a coherent system of beliefs that are shared, thus confirmed, by other members of a group to satisfy a common need. The delusion is no longer encapsulated as an island of abnormal cognition but invades and usurps the entire belief system.

The transition from a confabulation, to a delusion, to a fully delusional cognition implies that normative belief systems may not be easily differentiated from delusional ones. There are many examples of this in

literary and philosophical works. One of my favorites is Swedenborg's detailed and systematic account of heaven that was dictated to him by an angel.[11] Indeed, every commonsense belief is an adaptation in relation to other commonsense beliefs. If the entire system of belief was false, how would we ever know?

Belief and Free Will

Is the belief that one is acting freely essential to free will? Again, the pathological is a guide to the normal. There are cases of false belief in the rationality and purposefulness of an action. A paranoid believes that his action is freely chosen and not motivated by the paranoia. We would say his belief is false because of his diagnosis and because his acts and choices do not meet the usual criteria for free will. They are not a product of reason and deliberation. Normally, we believe ourselves to be free even if the basis of our free will is disputed or undecided. Those people who believe their acts are determined or deny free will act as if their acts are freely decided. They behave *as if* their choices matter. In a sense, the belief in free will inheres in free actions, even in those who deny it.

The idea that a belief in free will is a necessary accompaniment of free will derives from the subjectivist position that the existence of objects is conditioned on the belief they exist. That is, the belief an object exists is central to its existence.[12] As with free will, a solipsist may believe the world to be a mental image but continues to act in a "real" world. There is a difference between an explicit belief and an experience relating to that belief. The feeling of the reality of objects and free acts is more fundamental than any statement we can make about them.

Certainly, there are beliefs without objective correlates. A soldier may believe he is invulnerable. A mediocre novelist may believe he is greater than Tolstoy. A belief in a thing does not guarantee the thing exists. A belief in a thing that exists, including the belief in its existence, may not be a belief that is true. I can believe in the existence of an animal that has, unknown to me, recently become extinct. Beliefs are not copies of the "out there" but mechanisms of coping that, with respect to core beliefs, evolved over a long prehistory. The belief in a real world of real objects, even if it is a false belief, is a necessary condition of survival.

Beliefs are shaped as a fitness to the environment so the adaptation of an organism is its version of truth. Nietzsche said, beliefs are irrefutable errors. This does not mean all beliefs are false, but that the truth of a belief is its functional utility and the test of utility is survival. Leon

has written, "beliefs are the implicit guides on which survival depends. They have an interest in representing the real."[13] If, then, a true belief promotes survival, those beliefs that survive are "true" beliefs. The beliefs we share are true because they are the only beliefs we (as a group) could have and still prevail. An individual in the group could entertain a belief or a system of beliefs that deviates from the norm, but the price of this deviation, in proportion to its magnitude, is that he will be "eliminated." This is the case, for example, with delusional schizophrenics in institutions or sleeping on the streets. It is also the case with philosophers propounding radical new theories who are marginalized by their colleagues.

NOTES

1. R. Bogdan, "The Manufacture of Belief," in *Belief*, ed. R. Bogdan (Oxford: Clarendon Press, 1986), 149.
2. Consistent with the principle that feelings are less readily "retrieved" than ideas, Brown, *Self and Process*, 155.
3. N. Nelkin, "Propositional Attitudes and Consciousness," *Journal of Philosophy and Phenomenological Research* 49(1989): 413–430.
4. See F. Bradley, *Essays on Truth and Reality* (Oxford: Oxford University Press, 1914).
5. M. Polanyi, *Personal Knowledge* (Chicago: University of Chicago Press, 1958).
6. G. Santayana, *Skepticism and Animal Faith* (New York: Scribner, 1923).
7. G. Moore, "Proof of an External World," *Proceedings of the British Academy* 25(1939).
8. L. Wittgenstein, *On Certainty* (London: Basil Blackwell, 1969), 42.
9. See N. Malcolm, *Memory and Mind* (Ithaca, NY: Cornell University Press, 1977), on how we know what we remember is accurate without a standard in memory for comparison. The pathological material argues against such a standard, suggesting that the feeling of conviction (or lack of awareness of error) is generated with the content and is not an extrinsic function of correspondence in awareness.
10. Henri Poincaré described the "feeling of absolute certitude" preceding a proof in "Le Raisonnement mathématique," 1908; reprinted in *The Creative Process*, ed. B. Ghiselin, (Berkeley: University of California Press, 1952).
11. E. Swedenborg, *Heaven and Its Wonders and Hell*, trans. J. C. Agen (New York: Swedenborg Foundation, 1940). Originally published in Latin in 1758. Regarding philosophical delusions, Schiller has written: "Those unfortunate enough to have acquired and retained an exclusive view of truth are usually secluded in prisons or asylums, unless their truth is so harmlessly abstruse as not to lead to action, when they are sometimes allowed to be philosophers!" F. Schiller, *Humanism* (New York: Macmillan, 1903).
12. G. Strawsen, *Freedom and Belief* (Oxford: Oxford University Press, 1986).
13. M. Leon, *Philosophical Papers* 21(1992): 299–314.

CHAPTER 10

Action

ARGUMENT: *Subjectivity and agency are generated from primitive will with engagement of the core in action development. This feeling distinguishes an action from a passive movement. Reflex is extrinsic to mind. Automatic acts are low-level cognitions that express the subjective directly. Voluntary acts express the self after a delay. There is a transition from reflex to automatism to volition. Automatic actions become volitional as part acts issue from concepts. Volitions become automatic (skilled) as acts that require partial concepts pass to part-acts that express concepts. An action incorporates distributed evolutionary systems in a phased sequence of oscillators. The sequence progresses from an axial core to a proximate grasping space to an effectuation on external objects. Orientation and initiation are prior to awareness. Time constraints of readiness potential studies conform to expectations of the model for simple volitional acts.*

ACTION AND SUBJECTIVITY

Action is mental process, movement its (inferred) physical effect. Action, not movement, turns an object into a subject. Conversely, an organism that moves but does not act is an object. A subject is not created by adding motion to an object. Insects have intricate movements but are not subjects; tropisms and reflexes are not actions. A subject exists for its own sake, an object for an observer. A person is not a subject *of necessity*. A comatose person can be a subject or an object depending on the presence of automatic action, since automaticity, even at a primitive level, can generate a subject. A person in a coma can have a sub-

jectivity that is not possible in a computer. A computer might be able to move and have a mind and still not be a subject. To destroy a computer is not to commit a murder. What is lacking in a computer, even one that moves, is an action. This is independent of whether or not the software or network of the computer is imputed to be conscious.

Fouillée[1] wrote that the sense of activity is linked to our existence as thinking entities. This implies that action entails a state of thinking. When thought goes on in dreamless sleep, a sleeping man is a subject, perhaps even a rudimentary self. Sleep is low level cognition. Wordsworth wrote of "the whole creative powers of a man asleep." Sleep engages the action core. Dream involves neocortex, possibly the more recently evolved divisions.[2] In dream, actions are imagined but then, so are all actions. A preliminary action makes an image possible by creating the subject having the dream.

Subject and action are preconditions of self and agency. Activity creates a subject, agency requires a self in opposition to its own actions. Agency is the relation of a self to its own mental contents of which actions are one type. Or, agency is the relation between the neural correlates of these entities. The mind–brain relation is the relation of subjectivity to neural process. The self–brain relation is the relation of self in subjectivity to neural process. Velleman wrote, "Just as the mind–body problem is that of finding a mind at work amid the workings of the body, so the problem of agency is that of finding an agent at work amid the workings of the mind."[3]

The ascription of subjectivity to an object, a dog, human, or machine, is patterned on the knowledge of our own mental states. An implicit belief in the personal subjective precedes the belief in the subjectivity of others. This belief even precedes the scrutiny of one's own mental content.[4] A child's personal self-awareness is not inferred from the observation of others, since this could not give the *conviction* of self-knowledge, only the possibility of such knowledge. One could never be certain from an observation of other minds, or the behavior from which the existence of another mind is inferred, of the existence of one's own subjective states.

Moreover, self-awareness and awareness of subjectivity are different modes of awareness. The subjective is *within* the instinctive or automatic. The subjective is not the agent in a state of agency but the field from which the agentive state develops. Actually, the self is a barrier to the experience of the subjective which is grasped only when the self lets go. The subjective is not apprehended as an object. One succumbs to subjectivity.

PASSIVE AND ACTIVE

Passive movements are not actions. If the sensory nerves of an arm are severed and the eyes are covered, and the arm is passively displaced, the person is unaware the arm has moved. The lack of awareness is due to the absence of self-initiation and central or peripheral feedback. If peripheral feedback alone is preserved, e.g., with passive displacement, the movement is judged to be passive. With an action, even if movement and kinesthesis are lacking, e.g., a phantom or paralyzed limb, there is a feeling of activity that arises from the action discharge or its central reafference (p. 63).

The concept of a central sensation arising from an action goes back to Wundt's description of a feeling of effort or innervation, a concept criticized by James[5] who attributed the feeling of innervation to peripheral sensation. Phantom limb phenomena indicate that the feeling of activity and/or effort is a central (cerebral) phenomenon, though feedback from a limb is essential for movement. Central and peripheral reafference, and the action process, contribute to the awareness of the action, while a loss of peripheral kinesthesis disrupts initiation and the awareness of limb position.[6]

The psychologist H-L Teuber was fond of contrasting the displacement of objects by pressure on the eyeball with the lack of object motion when the gaze is diverted. Voluntary movements maintain object stability. The voluntary initiation of the action cancels the displacement of the retinal image so the object remains stationary. With weakness of the eye muscles, an effort to divert the gaze may result in illusory displacement. Since displacement occurs with passive and active ***but weak*** eye movements, the feeling of activity is in the action, not the displacement, and occurs through central reafference, not peripheral kinesthesis.

Nocturnal jerks of the legs, tics, or movements on stimulation of motor cortex are experienced with uncertainty as to whether they are self-initiated. Or, they may be felt as imposed, i.e., as something that ***happens to*** the person. There is a passive relation to these acts, even though they are generated as endogenous products. Such phenomena occur when the ***action core*** is not engaged. The activation of the core is the minimal condition for agency.[7]

Action and the feeling of agency do not entail movement. Inner speech is an action without a movement. Are the verbal images or propositions of inner speech actions? The behaviorists postulated incipient articulations during inner speech, and defined inner speech as an utterance minus its articulation. Whether inner speech can be characterized in this way or whether subvocal articulations accompany inner speech is unclear. We do know that brain damage disrupts inner speech

when the production *or* perception of language is involved. Inner speech depends on a combined activation of (speech) action and perception systems.

A schizophrenic may confuse inner speech with a hallucinated voice. A hallucination can be mistaken for a perception. From the perceptual standpoint, there is a transition from inner speech (active) to auditory hallucination (passive) to auditory perception (external, detached). The shift from agent control to object causation is linked to the nature of the perceptual content and a mitigation of the corresponding action development. The same sequence occurs in the progression of active to passive action.

Whether an image is a personal action or a hallucination is at stake in a phantom limb. An actual movement cannot occur if a limb is amputated but the self can will the limb to move and feel a phantom movement in conjunction with the willing. What is the basis of this feeling? Does it arise from the action? Is it a "false" perception? The feeling of movement could be perceptual if generated by the body schema or central reafference, but this would not explain the feeling of agency with the central phantom. Presumably, the core activation gives the feeling of agency, and *central* reafference gives the perception of a moving limb and the knowledge of body segment, displacement, etc. In a phantom limb, central reafference can overcome the lack of peripheral kinesthesis. If the eyes are closed, the missing limb feels present. Inner speech and phantom limbs are actions without movements but the lack of movement does not alter their volitional character.

To sum up, a subject is an object that is active in relation to other objects. An action generates a subject as an expression of primitive will. An active subject evolves to an agentive self. Collaterals of the action discharge or its peripheral implementation give a perception which the self experiences as an action. The main components of an action consist of the actualization of primitive will creating a subject through the primary or direct *feeling* of activity, and the "feedback" of this discharge (central and peripheral) giving the secondary or derived *knowledge* of what action is occurring. These components comprise the action in awareness.

In philosophical writings the causal nature of an action does not hinge on the distinction of active and passive. A simple passive act is having one's arm lifted by someone else. Another is having the arm elevate through an epileptic discharge in the brain. Still another is being killed by lightning. To what extent are these events *caused* by the self? Suppose I give a loaded gun to someone who then kills me. In a sense I am responsible for my murder without causing it. Can agent causation be passive? Attempts to attribute agent causation to such acts, e.g., by

participating in a situation that leads to the passive act, placing oneself in danger, not taking medications, etc., confound the concept of agency by chaining it to the external context. In these examples, the agent may be responsible for a series of events which precede and even precipitate the final event, but is not responsible for (does not cause) the final passive event which intersects with the agent through its own line of "causation."

AUTOMATIC AND AUTOMATIZED

There is a gradation from the automatic to the volitional.[8] Many automatic acts (e.g., breathing) can become volitional, and volitional acts (e.g., playing a new piece on the piano) can become automatic. The transformation of the automatic to the volitional seems to involve the incorporation of the automatic in the structure of the voluntary. How does this compare with the shift from the voluntary to the automatic in learning? Presumably, a voluntary task that becomes automatic differs from one that is automatic from the beginning but in what way? How does the organization of a later-acquired skill, say typing or an athletic skill, that passes through a voluntary phase, differ from that of early acquired actions like walking that can become voluntary, but are relatively automatic ab origo and do not become more automatic with practice?

A skill that becomes "modularized" in the course of learning fulfills most of the criteria for a putatively innate (genetically determined) module. To say a skill is innate, however, avoids having to specify its nature. What is the meaning of innate? Even the genes are context-sensitive. The continuum from the automatic to the voluntary, or the relation between volition, automatism, and skill, not the encapsulation or genetic determination of the automatic, is the more challenging phenomenon to explain.

For example, we can ask whether the voluntary or automatic quality of the "same" action involves a change in the action, e.g., a shift to another neural substrate, a loss of constituents or the addition of new content. Based on the work of Evarts, Lieberman[9] suggests that the automatization of complex actions such as speech results in subroutines that are stored in the motor speech cortex; e.g., articulation patterns in Broca's area. Automatization carves out functional assemblies comparable to modules. But to say an automatized action is a subroutine and locate it in the brain is not to provide an explanation. Articulation patterns are presumably related to Broca's area before and after automatization. The finger movements used in playing an old and a new piece on the piano must engage the motor cortex. Do naive configurations get "wired in" with training? What does "wired-in" mean? If a pattern is

a subroutine, one might expect it to be the same each time it is realized like the read-out of a computer. This does not occur. A pattern that is wired-in should not be susceptible to voluntary control, yet articulatory actions can be volitionally altered. The automatic and unconscious, including articulation, can become volitional or deautomatized.

An articulation pattern or a skilled finger movement is a terminal or endstage process. These actions are embedded in actions of the body, such as locomotion, posture, gesture, and facial expression, not to mention the cognitive and affective context in which they occur. Indeed, it is the context around a skill such as playing the piano that gives technique its conceptual and expressive force. There is no reason to conclude that automatized part-acts are impenetrable.

What then does it mean for an action to be automatic or automatized? In a sense, all actions are automatic in that their underlying process is nonconscious and independent of the self. Though the self may influence the process through which an action is realized, the self is as much a product of this process as is the action. The description of actions as more or less automatic, or more or less voluntary, depends less on their automaticity or voluntariness as on their phase of discharge and perceptual correlates. Actions differ in degree of awareness, choice, effort, skill, and in delay, latency, and speed, but all actions, all cognitions, run on "automatically."

A voluntary action that becomes automatic, such as playing a new piece of music on the piano, becomes more skillful. Skill is the facilitation of a learned behavior through repetition. When an action becomes skillful, the phase of indecision for a given item is replaced by the rapid selection of many items. This appears as a facilitation of action patterns with learning, or an increasing "consolidation" of the neural traces that constitute the target movements. Supposedly, the consolidation (wiring?) of the pattern results in an accelerated traversal of neural activity over the patterned elements.

In playing a new piece on the piano, each action of a finger, or a group of such actions, comprises an individual volition. An action plan or concept incorporates one or several such actions. The indecision (hesitation, choice, uncertainty) arises because one act or small bundle of acts constitutes one volition or concept. Increased fluency in the shift from concept to act, i.e., in the derivation of part-acts from their conceptual underpinnings, accompanies a more rapid turnover of individual volitions. In effect, a naive act which required a plan for itself becomes, with skill, a part-act embedded in a common plan with other acts.

With skill, the series of part-acts and their guiding concepts aggregate within parent concepts. The voluntary act that consisted of the action of a single finger now consists of a series of finger actions. This is

a specification of the conceptual base of the original volition. The concept that generated the action of a single finger is parsed to embrace an action series. The parent concept or category is a simultaneity that prefigures the not yet actualized sequence of part-acts. Because the series is generated by the same concept, it preserves its configurative or rhythmic frame. Facility in the implementation of one or several part-acts occurs because the part-acts are generated by a more deeply organized concept that embraces the series of individual volitions.

Skill, therefore, is partition at a depth of cognitive wholes, not the combining of separate cognitive or motor elements. With this partition, there is an accelerated transition to definiteness. In the entrainment of deeper sources in the progression from naivete to skill, the indecision over a note becomes the indecision for a measure or a theme. Eventually, one concept or plan steers an entire performance.

As a performance becomes skillful, the potential residual in the behavior as a remnant of indecision is realized directly into action. The automaticity, the technique, can lead to sterile repetition, or it can express the creative potential of underlying concepts. When creativity is on the wane, the potential (indecision) at the depths of the original act must be revived to invigorate the conceptual base of what has become automatic. This renewal is accompanied by a resurgence of the self in the depths of the action and the awareness of further possibility.

The difference between the naive and the skillful is not (just) the rapid processing of a terminal phase (e.g., finger action, articulation) but the parcellation of concepts that give rise to this phase. The parcellation discharges the potential as a rapid transit through choice. The rapidity owes to the multiplicity of part-acts that issue from the same concept where previously only one act was possible. Increasing multiplicity out of unity is the basis of skill.

AUTOMATIC AND REFLEX

There is a graded transition from reflex to automatism. For the purpose of this discussion, a reflex can be defined as a sensorimotor arc that is impervious to agent control. An example is a knee jerk reflex with stretching of the patellar tendon. An automatic action is a synergy of patterned movements which to a variable extent is susceptible to agent control. This may involve initiation, e.g., swallow, or initiation plus rate and amplitude, e.g., cough, respiration, or automatic actions that can become fully volitional, such as walking.

Some automatic actions like swallow require (micro)environmental triggers. Others such as walking are fine-tuned through interaction with

the (macro)environment. Early acquired acts are "etched" into process and become the configural ground of later experience. This may explain the inability to become conscious of early acquired events, including the grammatical rules of language. Generally, precocity in development is a mark of later automaticity.

There is a continuum from a reflex to automatism to automatic actions that become purposeful, to fully volitional acts. Spinal reflexes are physical events outside the mental state. In the brainstem there is a transition from reflex to automatism. An automatism can participate in a volition. There is voluntary control of automatisms as low as the medulla. An example would be the volitional control of the onset of swallowing. How is this explained?

Neurologists interpret the voluntary control of the brainstem "reflexes" in terms of separate neural centers for reflex and voluntary action. They postulate a downstream influence of "higher" cortical centers on "lower" brainstem mechanisms that mediate the reflex. This account leaves unexplained the demarcation of higher and lower or how the voluntary finds and activates the reflex. An alternative approach is that automatisms such as cough or swallow assimilate with the core and are derived through the self-concept to quasi-voluntary acts. This might occur through recruitment to neurons that will generate the core self. The recruitment could occur through overlapping networks[10] or virtual oscillators,[11] with subsequent generation through the full microgenetic hierarchy (Figure 10.1).

An absence of overlap with core units accounts for the insularity of a reflex. With partial overlap, the core incorporates low level automatisms such as swallow or cough so that initiation is under volitional control. More extensive overlap, and more rostral locus in the brainstem, permits further engagement and volitional transformation, e.g., walking. The increased control is not "top-down" but the actualization of the base automatism through the action structure.

An automatism becomes voluntary with a delay in discharge. In the shift from automatic to volitional walking, a locomotor apparatus in the brainstem generates an internal action rather than immediate discharge. The progression is through the basal ganglia, orbital and mesial frontal limbic system, to the neocortical convexity. Through this development, and with a visual or somesthetic component, walking becomes voluntary. To perceive oneself walking alters its subjectivity and evokes perceptual intentions. These attenuated percepts give rise to beliefs, plans, goals, and other ideational or propositional contents. Perceived (observed, deliberate) walking is a shift of a subjective automatism to the self's own object. The action not only contributes to subjectivity, it develops to an external object.

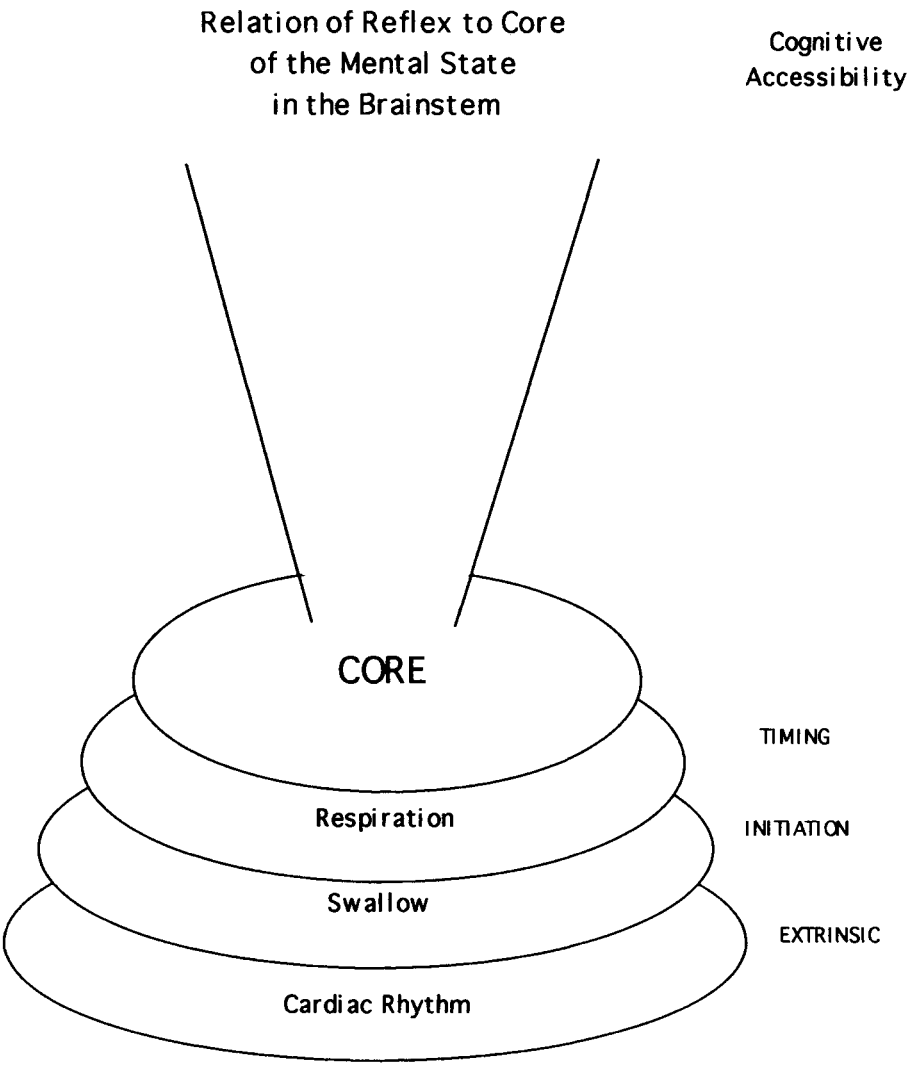

FIGURE 10.1. Schematic illustration of the progressive incorporation of reflex in the core of the mental state in the upper brainstem. The cardiac reflex is extrinsic to the mental state. The swallowing reflex can be initiated but not otherwise controlled. Respiration can be initiated and its timing controlled.

Voluntary walking is like deliberate stepping as the rhythmic nature of automatic walking is lost. An action of a leg is the content of a single volition. The traversal of a conceptual phase accounts for hesitation (choice) since each part-act represents a distinct (partial) concept. The

gaining of volition by an automatism is to some extent a reversal of automatization in skill.

In the monkey, a lesion of the motor cortex disrupts fine movement but ambulation is regained. The impairment of walking with similar damage in humans suggests an occlusion of the automatic by the engagement of damaged volitional systems. In pathological cases, the automatic can be lost and the voluntary preserved, or the reverse. Such dissociations reflect the phase of engagement in a common system. Automatic walking is a theoretic possibility with cortical injury. Automatisms occur in paralytics, e.g., yawning and stretching in hemiplegia. The impairment with cortical damage tends to reflect the degree of individuation, e.g., distal more than proximal, and involves the upper more than lower limbs.[12]

A dissociation between the automatic and the voluntary was described by Luria and Vygotsky in a Parkinsonian unable to walk but able to climb stairs. Vygotsky tore a sheet of paper into strips and placed them a foot apart on the floor. The patient could step over the strips without difficulty but "froze" after the last strip; i.e., he could step but not walk. This observation became the basis of a hierarchic approach to treatment in cases of brain damage. Other Parkinsonians show preservation of dancing or coital movements though unable to walk. Within the automatic there are degrees of vulnerability. In other conditions (apraxia) the converse occurs, impaired volition with preserved automatic activity.

In cases of dissociation, one would not postulate a different system for every automatic or volitional act. The basal ganglia, involved in Parkinson's disease, are not bypassed in stepping. Rather, the action of automatic walking, which is mediated partly by extrapyramidal structures, including the brainstem, is transformed to a neocortical phase. The dissociation reflects a bias toward deep or superficial planes in a common process.

A person can talk and walk at the same time because a self can engage in volitions while a subject is elaborated by automatic actions. The self observes its own automaticity but does not participate. I can perceive my automatic actions and sense they are part of the generation of subjectivity but I (my self) am ingredient only in the act of observation.

Automatic actions in a subject differ from mechanical movements in an object. Pure mechanism can be complex, as is the case with insects, but the deciding factor is not complexity, it is the presence of subjectivity. The difference between the automatic and the mechanical, or automatism and reflex, is the subjectivity that announces a mind where otherwise there is physical reaction. This activity can be "felt" by others through an empathy modeled on a personal subjectivity, an intuition of the personal subjective that taps a common source of organism.

TRANCE STATE AUTOMATISMS

A trance state such as hypnosis, fugue, or sleepwalking is a complex automatism that is purposeful but involuntary and so marks the transition to agency. Trance behavior is notable for its compulsory or induced character and the poor recall when the state is over. Probably, as with multiple personality, there is a failure of the core to generate an authentic self-in-awareness. A partial concept derived from the self establishes a transient autonomy. The trance state is more than subjectivity, yet not a state of agency. Choice and decision are lacking.

In hypnosis, the action is guided by a plan (concept) that replaces concepts of intrinsic origin. The self-concept is configured by the hypnotic suggestion which usurps competing concepts to shape the (dictated) act. The configured self prescribed in the suggestion is inauthentic; it is not contextualized with current and past life experience. Endogenous trance states like sleepwalking are a form of autosuggestion in which acts are driven by concepts occasioned by the self-concept of the moment. These concepts fail to individuate through sensory constraints to mature objects, resulting in purposeful acts incongruous with the objective world.

Trance state automatisms are goal-directed and driven by concepts. The passivity and lack of agency of the subject under hypnotic control are similar to the receptivity of the self in dream. A lack of agency also occurs in psychosis, e.g., in command hallucination, where endogenous "voices" express partial concepts that become autonomous and dictate the action. There is a common mechanism in hypnosis, command hallucination, multiple personality and related states. Whether a hallucinated voice is perceived to derive from an external or internal source, i.e., a hypnotic suggestion, a command hallucination, a multiple personality, or dream, depends less on its "real" source than on the phase of accentuation in its becoming. This determines whether the "voice" is apprehended as an intra- or extrapersonal content.

Complex skills such as playing the piano, even a well-rehearsed lecture, resemble trance states in their relation to the autonomy of their guiding concepts and their reduced self-monitoring. They differ with respect to depth (level of wakefulness), persistence or brevity and the conceptual "locus" of the autonomy, i.e., where the configuring has its main impact. In all these cases, behavior proceeds with the self not *inwardly* aware of the action. In skill, the self can be occupied with other thoughts. A pianist may converse while playing. A behaviorist once joked he could lecture while asleep through a stimulus–response chain. The absence of deliberation in highly skilled acts gives the immediacy and even a higher level of fluency to a performance. Part-acts distribute into

behavior without the detachment, uncertainty, or concept by concept replacement that is characteristic of voluntary action.

In sum, a common thread links normal, pathological, and induced trance states to complex skills. In each case, a concept guides an ensuing act as a relatively autonomous partition of the self-concept. Hypnosis entrains a part of the self-concept. In auditory hallucination, an endogenous partition is illusorily exteriorized. Sleep-walking, fugue, and command hallucinations, arise through intrinsic constraints that individuate the momentary self to a concept that actualizes a behavior unmodulated by the environment.

Both endogenous and extrinsic constraints can lead to acts that seem obsessive or compelled. Obsession is compulsion as an idea rather than an act, the idea being a surrogate action. Compulsion is the other side of choice, though even in compulsion choice is always a possibility. The essential feature of compulsion is its automaticity. Compulsion acts out a concept that does not have a major share in the self. A compulsion can be self-destructive. Compulsions in psychiatric disorder or drug addiction differ from hypnosis in being islands of subconscious cognition insulated from choice. Choice is obviated by the autonomy of the partition. The self is aware of the compulsion, and the alternatives to compulsive behavior, but lacks the feeling of control over its own acts. In a sense, the self is not truly *in* the behavior. A compulsion is an automatism imbued with subjectivity that displaces or supersedes volitions that the self, through propositions fractionating from the core, acknowledges as preferable. The propositional content describes a preferred course of action, and participates in and helps to elaborate a feeling of agency, even if this feeling is not for the object of the compulsion.

The Action Process

An action is not a *product* or *output,* not what an organism does, but is the mind (will) of the organism. Put differently, an action actualizes an organism. This actualization develops as a rhythmic process[13] of nested configurations (oscillators) over successive planes in evolutionary growth. For example, the basal ganglia and motor cortex are constituents of such growth planes. An action is a rhythmic structure composed of the full set of phases from brainstem, *through* the limbic formation, *to* the motor cortex.

This process is *serial* and reiterative, but develops in *parallel* with perception and language. From the actor's perspective, the seriality is not a step by step sequence of phases *in time*; the process is time-creating. A microgeny is a whole (irreducible) unit of psychological time

(p. 19). The realization of successive part-acts out of an underlying spatial or conceptual whole generates temporal relations out of an atemporal or simultaneous core. The temporal relations follow on the actualization. The sequence of actions at the surface of the mental state, e.g., the perceived motion of my hand from one position to another, develops from a simultaneity that "contains" as potential the full set of acts and positions (worlds) to be generated. Subjective time is fabricated *within* a microgenetic unit and displaced to a surface progression across successive replications.[14]

The *primary* action process is the intrinsic set of oscillators that is read off into physical movement. The *secondary* perception of this process anchors the action in content. The perceptual "chunking" of the primary action creates a sequence of stable objects, body parts or the body image, that persist as "solid" entities over a certain duration. By means of the perceptual revival the self looks on at its own acts. The self's perception of its actions, such as seeing or feeling an arm move, is a form of perceptual knowledge that grows out of the self-concept. The active feeling of the primary action, and the contribution of the secondary perception of the action, the relation between the self and its own actions as entities that endure, are essential components of the experience of agency.

The seminal work on brain and action was Yakovlev's[15] description of three concentric planes of motility that evolve over the neuraxis: the oldest in relation to autonomic or vegetative motility within the body; a surrounding layer, mediated primarily by basal ganglia, for action in body space ("body on body" action); and an outermost layer (motor cortex) for actions in the world. While Yakovlev's work has not received the attention it deserves, it did have a strong influence on Denny-Brown, Luria, Teuber, and other neuropsychologists, especially with regard to the concept of early and late stages in action realization.[16]

Yakovlev's paper is the starting point for a microgenetic theory of action based on pathological case study. The clinical material documenting the theory has been published elsewhere and is only summarized here. The objective is not to describe the anatomy and physiology of an action but to provide an overview in which more detailed accounts of physiology and localization can find a place.

The theory assumes that an action is initiated in the upper brainstem. This construct is organized about the body midline and distributes into postural and orienting systems within the body axis. This phase develops to an oscillator for the axial and proximal musculature sensitive to "internal context," then to asymmetric actions with the distal musculature that are "goal oriented," "context-free," and directed to external objects. The progression is from a space of body-on-body action, then

a proximate manipulation space of the arm's reach, finally to discrete actions on external objects. The action proceeds "bottom up" as a propagated wave from the upper brainstem, through the basal ganglia, mesial frontal lobe and convexity, to motor cortex. The core envelope of the action incorporates the to be realized postural, vestibular, axial, locomotor, and respiratory movements, as well as incipient vocal gestures in a matrix of primitive motility. Each phase in this structure is ingredient in the final act. Moreover, each phase in an action corresponds with a representation of perceptual space. Act and object develop out of a common ground and diverge or individuate in parallel over the microgenetic sequence.

Neural processes combine as oscillators linked to rhythmic patterns of respiration and locomotion, perhaps circadian rhythms.[17] For example, a kinetic rhythm supporting postural tone might transform to swimming, rocking or walking motions, then to selective limb and digital action or to the individuated vocal movements of speech.[18] Speech, of course, can be viewed as a type of rhythmic activity.

Thus, the upper brainstem and basal ganglia discharge in axial or midline motor systems; e.g., crying or sucking actions of the newborn, swimming, crawling, walking. This phase can be a terminus or persist as a background of transformation to extrapyramidal and limbic components linked to the proximal musculature. Similarly, this phase can discharge, or it can persist and develop to an action in an extrapersonal but not fully independent space, a manipulation space that is revealed in cases with damage to the parietal lobes. This is comparable to the grasping space of the infant or the action perimeter of the congenitally blind. Finally, the action specifies into the fine digital or articulatory musculature coincident with a perceptual phase of featural analysis on fully exteriorized objects.

In sum, the action core contains incipient vocal and somatic elements in a space centered on the body axis linked to rhythmic automatisms and drive like states. This complex specifies through progressive demarcation of limb, body and vocal motility by way of frontal paralimbic formations, including anterior cingulate gyrus and supplementary motor area. The paralimbic configuration specifies through premotor and motor cortices to a space apprehended as real and independent. The final derivation is from a space of object manipulation to one of discrete, partial actions on independent objects.

This phased parcellation transforms a core (fundamental?) frequency through a sequence of rhythmic oscillators (harmonics?). For example, an oscillator for the respiratory rhythm frames an oscillator for the speech melody or prosodic contour. This complex is derived to the programation of phonetic units in speech. An oscillator for axial and

proximal motility fractionates to a frequency for the fine-tuned sequence of digital and orofacial action. One can envision a stratified system of kinetic melodies discharging into motor keyboards at successive evolutionary planes, creating a dynamic pyramid of rhythmic, vibratory levels. This, then, is the presumed microgenetic "structure" of the primary action process.

Implications for Agency

The microgenetic account of action has many implications for the concept of agency, some of which include:

1. The initiation and ballistic of the action are deposited prior to awareness, not only prior in the sequence of the action over time, and prior in the iteration of mind–brain states, but prior in the precedence within a single microgeny.
2. The conscious self cannot directly cause an action to begin, i.e., determine its onset and preparatory phases, since the conscious self follows initiation. The conscious self might still control the development of the action, so long as the initial phase does not establish the "plan" of the action or if the plan issues from the self. This would be the case, for example, if the "plan" or concept of the action is mediated by limbic formation, while upper brainstem systems are responsible for initiation and orientation.
3. Graded or staged unfolding entails that a copy of the final action does not exist prior to its actualization. The final action is not selected or chosen from a menu of options but is elicited by a narrowing of the possibilities of expression.
4. The primary action precedes its instantiation in movement. The secondary (perceptual) awareness of an action, i.e. the knowledge of the action, follows the primary action process and is either simultaneous with or successive to the actual physical movement.

These topics are discussed in turn below.

Initiation

When I decide to lift my arm, and then do so, I am aware of the moment of decision, the onset of the perceived action as a whole and the lifting of the arm, but I am not aware of the precise moment of activation or the effectuation of the action at serial points. My awareness is not coincident with the phase of activation, and there is no awareness of the

various stages as the act unfolds. For example, in lifting the arm, I am not aware of the postural set or the motion of different joints and muscles. I can become aware of these stages at the expense of the smoothness of the act and its conception as a whole, but then each phase becomes a distinct action. These observations are consistent with studies of physiological correlates of initiation prior to conscious implementation. The work of Libet and Kornhuber is often cited in this regard. This work has received some criticism, perhaps because its implications may be difficult to assimilate with conventional models, but the findings are robust and in accord with microgenetic principles.

In the initial studies, a slow negative (readiness) potential (RP) was found to precede self-paced finger flexion by about 1 second. The RP was vertex and bilateral, maximum over supplementary motor area (SMA), which is limbic-derived (meso)cortex, for all movements sampled, with sequestration over the opposite motor cortex coincident with finger flexion. Subsequently, a variety of studies concluded that the RP occurred about 0.6 seconds prior to planned movement. Libet found awareness of the urge to move 0.2 seconds prior to muscle activation while the RP preceded awareness of this urge by 0.35 seconds. Of note, excitation in SMA also occurs in monkey 0.1 to 0.2 seconds prior to a "voluntary" task.[19]

The microgenetic interpretation of preconscious activation has the orientation, postural, and ballistic setting of the action originating in upper brainstem and/or basal ganglia. The concept (plan, schema) that organizes the action arises in limbic formation. This is not yet the conscious plan or content of the act but a preparatory phase. The knowledge of an incipient action, the *conscious* plan, goal, or strategy of the act, and the conscious decision to act, are conceptual or propositional contents generated with the action out of this phase.

Moreover, the urge (want, desire) to "move my finger now!" is also an act, comparable to the action of the finger that follows. The experience of an urge to act before an action, or a decision before a volition, are two distinct and successive acts (mental states). An urge to act announces a decision or a behavior. When the urge becomes conscious as an intention, a desire, or a decision, it constitutes the action of that moment. A subsequent (motor) action is anticipated by yet another (subconscious) urge. The subconscious urge (e.g., orientation) precedes the conscious urge (desire, decision) by 0.35 seconds.

The timing relations between the RP, the awareness of an impulse or decision to act, and the conscious action, even for a simple deliberate action such as lifting the finger, reflect the series of mental states that obtain between two boundary events, such as the RP and the conscious impulse. This series gives the act in awareness. The act incorporates a

sequence of (absolute) mental states each having a duration of about 100 milliseconds. The absolute present or the current mental state contains a series of embedded presents from the immediate past. The onset of the action sequence (RP) defines a limit within the present, i.e., the conscious impulse, anchored to the surface of the mental state. The discrepancy between this limit and the surface creates the (minimal) duration of 0.5 to 1.0 seconds for the phenomenal (specious) present.

In the RP studies of an impulse to move the finger, the duration between the RP and a movement incorporates a sequence of perhaps four to six mental states, sufficient for a conscious present to frame the performance. No conscious act is so brief or basic as to require only a single (absolute) mental state for its enactment; a state decontextualized from its predecessors. At least four, probably more, successive mental states occur in the RP paradigm for the consciousness of a simple purposeful action:

1. The subconscious urge or impulse to act. This occurs in the core of the present of a preceding state of consciousness, as the onset of the ensuing (i.e., replacing) mental state. The onset of the act corresponding with the RP is buried in the occurrent state which is a resultant of mental states over the preceding 0.5 to 1.0 second interval.
2. The awareness of the impulse as a decision, plan, goal, etc. A very simple concept ("I will move my finger now!") appears to require about 3 to 4 mental states. If the source concept is more complex it will require a longer sequence of mental states to serially realize (exhaust) its content.
3. The action process. This too may require multiple states depending on the complexity of the action and its rate.
4. The awareness of the (motor) action. This follows the action process, since the knowledge of the act, not its plan but its enactment, i.e., the awareness of a finger movement, arises through the perceptual reafference of the action discharge.

These states can be dissected in the one action under scrutiny. If each state, e.g., the spontaneous "urge to act," the awareness generated by the urge, the motor action, and the awareness of the action (from the "feedback" of the action process or motor discharge), requires about 100 milliseconds, the 0.35-second interval between the RP and the awareness of the urge to act, or the 0.2-second interval between the urge and motor initiation, would reflect the duration for the becoming of two or three mental states, and the sequence of the states would account for the 0.5 to 1.0 second interval between the RP and the final conscious action.

These studies have been interpreted in different ways. Libet tried to retain a guiding role for consciousness, linking it to a veto (or trigger) reminiscent of the fiat of William James. Eccles interpreted the dissociation of consciousness and action as the intercession of consciousness in the physiology of action. He argued that the self scans a cycle of activations (RP's) before choosing to implement an act, i.e., that the (physiological) RP announces a (self)-initiated action.[20] However, it is implausible to maintain the primacy of the self in the initiation of action when physiological activation consistently precedes the first conscious urge. This finding provides a serious difficulty for a (strong) dualist position on willed action since it refutes the thesis that a voluntary action, or the timing of a voluntary act, is ***initiated*** by a conscious self.

If initiation precedes conscious decision, the determination of the content of an act, or the type of action that is forthcoming, may precede the awareness of what type of act will occur. This follows if the content of a volition is not established independent of its onset, since content and timing are interdependent. The RP studies have been interpreted as relating to initiation, not action type, but initiation must be linked with content. The form of the act must be appropriate to its onset. If I throw a ball to a person on the run, the form of the action must be precisely adjusted to its timing. This is just as true for a voluntary flexion of the finger. However, given that activation is prior to awareness, for the choice of a volition to be freely decided, initiation must be indecisive in the selection process. The conscious self must have some control over the menu of possible acts, or at least a choice of acts within the menu, i.e., what is possible given the form of the act thus far, even if the menu is initiated prior to the awareness of choice.

In sum, how one interprets the lack of conscious initiation depends on the degree to which initiation ordains the action content, i.e., is activation motor arousal or orientation or does it include or bias the conceptual antecedents of the action. Orientation is decisional. The readiness to act might cancel the option of not acting. The readiness establishes the scope of possible acts (the "menu"). If there are degrees of activation, the orientation could range from a nonspecific arousal to all but the distal implementation, e.g., whether the act involves the upper or lower limbs, whether it is conceptual or motoric. The activation (RP) might entail a concept, a plan or code specifying the components of the action, rather than a purely motor preparation. If the ballistic is included in the code, then the coordinates of the action are preset.

Self as Causal

If an action is not initiated by the self, and if certain actions are excluded once the action is initiated, the self cannot activate all (any?) possible actions. Activation prior to awareness excludes not only conscious activation, but the conscious selection of the set of actions that are possible within the framework of the activation. Preactivation may not exclude selection within the menu or conscious guidance or control of the action after it is activated. Conversely, if the activation is for the potential to act and does not prescribe the action, i.e., if the potential is uncommitted and preliminary to any possible action, the degree of resolution of the potential will establish the "menu" that enters consciousness. Depending on the degree of commitment in preactivation, the conscious self could still promote (cause?) or delimit (constrain) the choice or final resolution of an action that is predetermined.

What elements of the action are forecast in the activation that the RP samples? The multiplicity of options available to the self after initiation (RP) must be reconciled with the fact that the RP and related slow potentials preceding the urge to act correlate with activity in mesial frontal cortex. Observations in cases of pathology and electrical stimulation,[21] indicate that mesial frontal cortex mediates a phase of segmental commitment involving the proximal musculature. The association of motor activation with limbic neocortex suggests that the action correlated with the RP has already specified to category and body segment. If this is the case, the specification may have proceeded too far in the action sequence for consciousness to exercise more than a steering effect on the distal resolution.

Presumably, the RP is present (recurs) as a continuous background activity that is accentuated in deliberate or purposeful acts. The decisional aspect delineates an *occurrent* RP from the *recurrent* background. A continuous baseline activation seems probable, since activation also has to precede willed *inaction*, to which volition has an equal claim. Conscious activation would be crucial only if background activation did not precede inaction, i.e., if the procedure did not sample or elevate the threshold of the background state. There is evidence that a "last moment" decision *not to act* consumes a duration equivalent to one or two intervening mental states; i.e., that voluntary inaction or an inhibition (veto) of an action is itself an action.

The RP is clearly the product of an artificial situation. Brain process is continuous; there is no zero-state preceding activation. At what point in ongoing brain activity does initiation develop? Presumably, the initiation of an action reduces to the initiation of the brain process that initiates the action. If the conscious decision (or impulse) to act is an

action, in the sense that every mental state involves an action, even if it does not terminate in movement, the RP for the decision to act would be the RP for a mental state, not the RP for the intrusion of conscious will in brain process. The RP should precede any mental state regardless of whether a movement is involved. The decision to act enhances the baseline RP through a mesial frontal development. It is a way of slicing the continuity of experience to mark off a discrete episode.

In sum, self and consciousness do not initiate action but develop in the replacement of prior states of activation. For this reason, except for the possibility of backward causation (p. 27), the self is not causal in initiation. The RP studies demonstrate that the activation of a mental state precedes awareness of the content of that state. If one assumes that activation precedes awareness but that activation is nonspecific, the activation (for an action, a concept, a proposition) could incline the subject in a certain direction. How far in this direction is uncertain. The RP studies leave open the possibility that the conscious self can select an action from a state of activated potential.

In microgenetic theory, the resolution of an action out of the self could enable the self to configure, guide, or constrain the action even if it does not select the content that is developing. If the urge to act is preset and if the category, trajectory, and segment of the body are given prior to awareness,[22] this would reduce the scope of free will or agency to a choice within a limited set of options, i.e potential for those options, or to the fine-tuning of the distal phase of an action the greater part of which is pre-selected.[23]

The self is generated over a part of the depth that the action process traverses. The deep self is cotemporal with early phases in action generation but not the inception of the action, which begins earlier. Motor initiation precedes the deep self, which emerges from the developing action. Not only is the action not initiated by the self, the self is a product of the initiation, which establishes the subjectivity of the organism. The action begins in the upper brainstem, too caudal a system to support the deep self, which more likely depends on limbic structures.

An action has a depth beneath awareness and a surface in the world. A self has only a depth. The act goes outward to the world, the self remains intrapsychic. Otherwise, the self would exteriorize like an act or an object. The conscious self could influence that segment of the action distal to its phase of origin, whereas the deep or subconscious self could affect more basal phases in the action. The problem of causation applies equally to conscious and nonconscious process, but agency requires consciousness. For agency only the more distal phases in an action could be under conscious guidance.

Parenthetically, this account of the initiation of an action, or cognition, prior to conscious control, and the configuring effects of conceptual phases, recalls Whitehead's description[24] of the transition from the physical to the mental pole in the actualization of momentary objects. An actuality begins in the physical domain and acquires integration as it is guided to self-completion through its own conceptual operations.

Finally, if initiation constrains but does not prescribe an action, preactivation might be comparable to a predisposition to act in a certain way. If free will is reduced by a predisposition, does this apply to states analogous to the RP studies? For example, are habits, preferences, and compulsions comparable to preactivation as constraints on freedom? With preactivation, initiation in the core anticipates, possibly configures, the deep self. With habit or compulsion, the deep self anticipates and probably configures the conscious self, which may configure the surface portion of the action. The deep self revives the experiential history of the individual including habits, preferences, tendencies, and compulsions. These are psychological descriptors later in becoming of the inclination that preactivation samples and the self guides to actuality.

Graded Unfolding

Graded unfolding is the continuous actualization of a configuration. The figural transform is a sequence of implicit phases embedded in an actual object. The final content, whether an act or an object, is not a copy that is "retrieved or accessed" into consciousness. A "menu" is not a list of options but a range of potential outcomes. The potential for an action is the potential for a derivation to an ensuing phase. A word, an act, an object is the result of a progressive delimitation in which the multiplicity of potential outcomes fractionates to the potential for a narrower set of outcomes that are qualitatively different from those of antecedent phases.

In the actualization there are opportunities for novelty or derailment at successive points. The statistical probability, however small, that the atoms of an object, say a chair, might fly apart at any moment or transform into another object cannot be ignored. In microgenesis, there is a possibility of a departure from replication or causal prediction, i.e., novelty, at every phase in the generation of a content. The departure is not because the potential for novelty is overwhelmed by constraints on replication. The high degree of predictability gives the impression that novelty is rare and pertains at best to a restricted part of any action. From a microgenetic standpoint, the stability of acts and objects is not the solid outcome of causal expectation but a categorical perspective within which novelty is rampant.

Secondary Awareness of Action

There is general agreement that the awareness of a (motor) action depends on central and peripheral "feedback." Is this true for the conscious urge to act? If the urge is an action, the awareness of the urge should be governed by the same conditions as any action, i.e., awareness should follow the urge as a secondary effect. This is required even if the urge or disposition preceding an act does not involve the excitation of motor systems. To the extent the urge is an action, its awareness depends on the primary action discharge. In contrast, direct awareness would occur if the urge is perceptual, e.g., a goal or reason. Awareness is direct for perception, indirect (recurrent) for action. The occurrence of the RP 0.2 seconds prior to the conscious urge to act suggests that the urge to act is a preliminary action for which there is indirect awareness; i.e., the *awareness of the urge*, like the awareness of any action, is a secondary phenomenon.

For every mental state there is direct awareness of the perceptual content and indirect (perceptual) awareness of the action. The lag for the awareness of an action inheres in every perceptual content. Act and percept are synchronous, but the perception of a given part-act develops in the perceptual content of the ensuing mental state. The RP studies do not measure *perceptual* activation. If a decision to act is perceptual, the subject would feel the decision in advance of the action, the action constituting the content of the next state of awareness.

These speculations are provisional, since the difference in the precedence of object and action awareness, at best about 0.1 seconds, is probably blurred in the phenomenal present. After all, a subject is unaware of the 0.35 second interval between the RP and the urge to act. Still, the precedence of phases within and across mental states, and the reafferent nature of action awareness, may give rise to subtle influences on agency and the distinction of self and world.

NOTES

1. Fouillée, "Le sentiment de l'effort, et la conscience de l'action," *Revue philosophique* 28(1889):561; cited in James (1890, 570).
2. I have had dreams in which this could be inferred, for example, a courtroom scene in which a person overheard to say "docket money" in an attempt at a bribe insisted he said "pocket money." There is reason to think that Wernicke and Broca cortex are required for rhymes and rhyme-judgments so that such a dream should involve neocortical language zones.
3. J. Velleman, "What Happens When Someone Acts?" *Mind* 101(1992):461–81.
4. J. Garfield, "The Myth of Jones and the Mirror of Nature," *Philosophy and Phenomenological Research* 1(1989):1–26.

5. W. James, *Principles of Psychology* (New York: Henry Holt, 1890).
6. For a case study of the effects on movement of extensive denervation, see J. Cole, and Sedgewick, "Perceptions of Force and of Movement in a Man Without Large Myelinated Sensory Afferents Below the Neck, *Journal of Physiology* 449(1992):503–15.
7. Consistent with the argument by Frankfurt that the way to determine if a movement is an action is not by considering prior causes, but rather conditions during the movements. "The problem of action." In H. Frankfurt, *The Importance of What We Care About* (Cambridge, Cambridge University Press, 1988), 72.
8. The term *automatic* rather than reflex was preferred by Hughlings Jackson, who argued for a continuum between the automatic and the volitional. Jackson's concepts were endorsed by E. Evarts, Y. Shinoda, and S. Wise, *Neurophysiological Approaches to Higher Brain Functions* (New York: Wiley, 1984).
9. E. Evarts, "Motor Cortex Reflexes Associated with Learned Movement," *Science* 179(1973):501–503; P. Lieberman, *Uniquely Human* (Cambridge, MA: Harvard University Press, 1991).
10. Such networks have been demonstrated, for example, by J. Siegel, "Behavioral Functions of the Reticular Formation," *Brain Research Review* 1(1979):69–105.
11. E. Dewan, "Consciousness as an Emergent Causal Agent in the Context of Control Systems Theory," in *Consciousness and the Brain*, ed. G. Globus, G. Maxwell, and I. Savodnik (New York: Plenum, 1976).
12. Based on such studies, Evarts et al., *Higher Brain Functions*, conclude that, "one of the roles of the cerebral cortex in movement is to enable the animal to make *fractionated voluntary* movements of any motor organ" (emphasis added).
13. N. Bernstein, *The Coordination and Regulation of Movements* (London: Pergamon, 1967); J. Fuster, "Commentary." *Behavioral Brain Sciences* 15(1992):741–743; Brown, *Life of the Mind*, 302–321.
14. A. N. Whitehead, *Process and Reality* (New York: Macmillan, 1929), on precedence without temporality in the actualization of abstract entities.
15. P. Yakovlev, "Motility, Behavior and the Brain," *Journal of Nervous and Mental Disease* 107(1948):313–335.
16. For example, in different ways, Derek Denny-Brown, A.R. Luria, and Hans-Lukas Teuber interpreted right hemisphere motor deficits as disruptions in a proximate or intrapersonal space with an impact on actions close to the body, left hemisphere deficits as disruptions of distal actions directed outward to extrapersonal space: D. Denny-Brown, *The Basal Ganglia* (Oxford: Oxford University Press, 1962); A. R. Luria, *Higher Cortical Functions in Man* (New York: Basic Books, 1966); H.-L. Teuber, "Perception," in *Handbook of Physiology*, Sec. I, Neurophysiology, vol. 3, ed. J. Fields (Washington DC: American Physiological Society, 1960).
17. *Life of the Mind*, 303. There is also a major role of the cerebellum in oscillatory phenomena.
18. The concept of multiple oscillators supporting different aspects of movement originated with Bernstein. See M. Turvey, "Preliminaries to a Theory of Action with Reference to Vision," in *Perceiving, Acting and Knowing*, ed. R. Shaw and J. Bransford, (Hillsdale, NJ: Erlbaum, 1977); re time perception, see R. Miall, "Oscillators, Predictions and Time," in *Time, Action and Cognition*, ed. F. Macar et al. (Dordrecht, Kluwer: 1992), 215–77.
19. H. Kornhuber, "Cerebral Cortex, Cerebellum and Basal Ganglia," in *The Neuro-sciences: Third Study Program*, ed. F. Schmitt and F. Worden (Cambridge, MA: MIT Press, 1974); B. Libet, "Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action," *Behavioral Brain Sciences* 8(1985):529–566; G. Goldberg, "Supplementary Motor Area Structure and Function: Review and Hypotheses," *Behavioral Brain Sci-*

ences 8(1985):567–616; C. Brinkman, and R. Porter, "Supplementary Motor Area in Monkey," *Journal of Neurophysiology* 42(1974):681–709.

20. *Behavioral Brain Sciences* 8(1985):542–543; personal communication (1988), 324.
21. For discussion: G. Goldberg, "Premotor Systems, Attention to Action and Behavioural Choice," in *Neurobiology of Motor Programme Selection*, ed. J. Kien, C. McCrohan, and W. Winlow (Oxford: Pergamon Press, 1992), 225–249. Metabolic studies of frontal lobe and willed action are reviewed in C. Frith, K. Friston, P. Liddle, and R. Frachowiak, "Willed Action and the Prefrontal Cortex in Man: A Study with PET." *Proceedings of the Royal Society, London* 244(1991):241–246.
22. The pathological literature suggests that actions can be initiated and implemented prior to, or without, awareness. For example, hemiplegic writing in severe aphasia may occur with little or no awareness. See J. W. Brown, B. Leader, and C. Blum, "Hemiplegic Writing in Severe Aphasia," *Brain and Language* 19(1983):204–215; J. W. Brown, "Discussion," *Aphasiology* 1995:9, 277–279. With bilateral occipital damage and severe deficits in object recognition, patients may show preshaping and rotation of the hand when reaching to grasp an object. See M. Goodale, A. Milner, L. Jakobson, and D. Carey, "A Neurological Dissociation between Perceiving Objects and Grasping Them," *Nature* 349(1991):154–56.
23. In a personal communication, Gary Goldberg (1993) notes that physiological studies of the distribution of the RP are consistent with predictions of microgenetic theory, namely, greater "limbic" involvement for actions sensitive to "internal context," and frontal convexity involvement for actions that are sensitive to "external context."
24. A. N. Whitehead, *Process and Reality*, 524.

CHAPTER 11

Morphogenesis and Mental Process

ARGUMENT: *Parcellation and heterochrony (neoteny) reflect the pattern and rate of a growth mechanism in morphogenesis. Structure (morphology) and function (behavior) are staged realizations of morphogenetic process. This process continues in adult cognition as the actualization of the mind–brain state. Parcellation obtains in the pruning of cells and connections in early growth, inhibition in a relatively stable morphology and constraints on context-item transforms in microgeny. Selective retardation in process (neoteny) leads to growth at earlier (juvenile) phases. This accounts for the specification of the language areas and elaboration at preliminary phases in mind, in dominance, introspection, and creativity.*

> We create distinctions, then
> Deem that our puny boundaries are things
> Which we perceive, and not which we have made.
> WORDSWORTH, Prelude II,221

Ontogeny is a bridge from the genetic code to the phenotype, phylogeny the sum, or average, of the ontogenies of a given line. Ontogeny covers a period from conception to adult structure with no clear boundary except death at the end of a developmental sequence. At some relatively stable point in this sequence an organ is judged to express a phenotype. The phenotype is usually assumed to be directly generated by the genome. For example in brain organization there is assumed to be a correlation between the genetic code and the specialized adaptations of

the mature organism. Ontogeny is the process through which this correlation is achieved.

The extent to which morphology is specified in the genome is uncertain. Presumably, the genetic code contains a set of instructions for the processes—or events that lead to the processes[1]—that will realize the code in structure. Such processes as are specified by the genetic code are still a long way from phenotypic structure. Neurobiologists recognize this problem and focus, therefore, on the epigenetic functional relations or algorithms that translate code to structure, not the manner in which genes produce nerve cells or circuits.[2] However, our knowledge of relations underlying the morphogenetic process, the transition from the genetic code to the developed organism, is still very incomplete.

From the standpoint of mental development, the onset of sensation and learning at birth creates a natural testing ground for the study of genetic specificity. The newborn is prepared for many complex behaviors through a genetic endowment that is relatively uncontaminated by learning. Yet the innate capacities of the mind–brain of the newborn are difficult to specify and continue to be a source of lively debate. Speculations range from innate rules, ideas, and/or mental processes, to the more conservative notion of constraints on action and perception.

The problem with studies of innateness that begin with the newborn is the focus on behavior as an interaction of a delivered morphology with experience. Circuits in the brain are taken to be the outcome of a prenatal process of growth when, in fact, growth continues on into late life. Similarly, brain function is assumed to be determined by the pattern of electrical activity that maturing circuits generate, i.e., the output of the structure, when the pattern of activity, as behavior, is an expression of sustained growth. The cleavage between growth and function, or structure and process, results in a neglect of the formative history of the mind–brain, in evolution or in utero. The goal of development is a machine that can be instructed or realize a functional program. On this view, the fetal segment of ontogeny is less informative than the confrontation of the innate and the acquired in the earliest period of postnatal life.

In contrast, suppose that development ***is*** morphology, i.e., that morphology is an artificial slice through development with behavior its four-dimensional structure.[3] From this standpoint, the newborn is not a starting point to study the innate determinants of language and behavior, but is a phase in a lifespan process. Onto-(morpho)-genetic process leading to the mind of the newborn, also lays down function after birth. The dichotomy between the innate and the learned is orthogonal to the **nature** of mental activity, and this activity is independent of birth as a pivotal event.

On this perspective, an early phase in development deposits morphology, a late stage deposits function. A common process elaborates both morphology and function, function being the iteration of growth through the morphology. There have been prior speculations along these lines, e.g., the ideas of Loeb and Goldscheider, that "configurations experienced in perception might derive from excitation in the brain resembling the 'force lines' that determine form during embryogenesis."[4] A relationship has been suggested between "ontogenetic sculpting" and mechanisms of learning and information representation in the brain.[5] Tucker[6] commented that physical growth in the brain is psychological growth. The problem is to specify the growth process and determine whether this process is related to processes underlying cognition in the adult.

There have been attempts to define the morphogenetic process with greater precision. Goodwin[7] argued that developing organisms have "an extensive range of morphological potential, describable in terms of probabalistic fields which collapse . . . into specific morphologies reflecting the particular conditions, internal and external, which act upon them." According to Goodwin, the generative principles that account for the progression from whole to part recur so that a taxonomy of biological form can be achieved through a hierarchic ordering of the transformations. Katz[8] proposed that "ontogenetic buffer mechanisms" mediate the transition to functional brain architecture. These mechanisms include: exuberant growth with specificity through parcellation or pruning of connections, possibly by competitive interaction[9] and heterochrony, a variation in the timing of developmental process.

It is a thesis of this chapter that processes in the development of the brain recur as processes in cognition; that development and cognition—ontogeny and microgeny—are different ways of looking at the same process. Two lines of evidence for this thesis are explored, the relation of parcellation in development to specification in cognitive processing, and heterochrony as a theory of developmental abnormality in relation to pathological symptoms in mature organisms.

PARCELLATION

Morphology

Parcellation is the pruning of exuberant connections in the growth of the brain as a way of achieving specificity in mature brain structure. According to parcellation theory, the connectivity of the brain is accomplished, at least in part, through a loss of connections. Indeed, most

structures in the vertebrate brain have a larger number of neurons during development than adulthood. The decrease in synapses is even more striking. In the macaque neocortex there is a loss of over 2 trillion synapses by the fifth year of life.[10]

Ebbeson[11] has written that "most, if not all, systems go through phases of diffuse projections that later become more restricted, presumably by the degeneration of selected axonal branches or the loss of selected neurons." The finding of initial proliferation and later elimination in the progression from the general to the specific has been described mainly in the study of sensory systems, where exuberant growth in juveniles with loss of cortical connections (neurons and synapses) in adults, is a characteristic feature. There is also evidence for parcellation in the growth of callosal fibers;[12] connections are initially diffuse and abundant and become specified through elimination. In studies of cerebral dominance, hemispheric asymmetries are related to callosal thickness and may reflect pruning rather than accentuated growth.[13]

Even cytoarchitectonic specificity has been attributed to the gradual connectivity of initially homogeneous neocortex.[14] Innately driven process determines the connectivity but so does experience. Early visual deprivation in animals can prevent foveal specificity and lead to a more ancestral state of diffuse or ambient perception.[15] Experience enhances the specification through constraints on emerging form.

At some point in morphogenesis, presumably after most anatomical connections are established, the parcellation effects that produced the structure of the brain give way to parcellation-like effects that characterize processing within this structure. The transition is from the ***elimination*** of connections in development to the ***inhibition*** of connections in maturity. The question is, is specification in neuronal development only analogous to later processes of differentiation through inhibition or are development and behavior manifestations of a common process?

Physiology

There are numerous examples of ***physiological*** parcellation. Many years ago, Coghill[16] proposed that partial patterns, e.g,. local graded reflexes, individuate out of global patterns, such as mass reflexes, and that the earlier state can reappear in pathological conditions. Thus, spasticity with CNS injury represents a recurrence of the excitatory patterns of earlier stages.

Another example of physiological specification is the diffuse to focal gradient in dominance establishment, e.g., regional specification of the language zones.[17] Evoked potentials show a gradual restriction to the cortical site of the stimulated modality with recurrence of the global

pattern in cases of brain damage. The progressive specification could be attributed to synaptic pruning, but the recurrence of the generalized pattern with pathology implies disinhibition. Wall[18] suggested that some forms of recovery after brain injury may reflect disinhibition of latent synapses and reenlargement of receptive fields.

Such instances of progressive specificity in postnatal development can occur either through the active inhibition of synapses or their elimination, or both. There is a transition from a predominantly morphological pattern to one that is predominantly physiological. The specificity achieved through elimination in the development of the brain continues as the selectivity achieved through inhibition in the further development of functional systems. Lateral or surround inhibition (and figural contrast) accomplish the specificity in process that parcellation accomplishes in growth. The implication (see below) is that the putative modular organization of the mind–brain *results* from a process in which constraints are applied at successive phases in the differentiation of form,[19] whether in morphology or behavior.

Growth and Process

Our concepts of morphology and process are shaped by implicit beliefs on the relation of structure to function; i.e., that function is the output of structure. However, in a temporal context, organic structure is neither the source nor residue of process but the momentary appearance that process takes on. In growth, process generates morphology. In a developed morphology, process generates function. Is there a common process that elaborates structure and function depending on the stage in growth? Put differently, is the process of nerve growth and connectivity the basis of the process that the connectivity instantiates in function? If the meaning of growth is restricted, say, to alterations in synaptic protein, and if the concept of mental process is confined to the output of a population of cells, there will be a gulf between the two activities, even if synaptic growth is a determinant of the activity pattern. However, if the concept of growth is expanded from a cellular event to a population dynamic, and if mental processes are conceived as the configural properties of this population, it becomes possible to map one process to the other.

Cognition

Parallels in the development of structure and process extend to the activity underlying cognition in the adult. The emergence of elements through parcellation and the development of specificity through inhibi-

tion correspond with the analysis of cognitive wholes through context–item transformations.

The description of the cognitive process as a treelike series of nested context–item shifts is a core feature of microgenetic theory. Selection of items occurs through the actualization of elements within manifolds with each tier serving as a ground for another specification. The specification occurs through constraints on emerging content. The constraints are both intrinsic, on the unfolding mental representation, and extrinsic, in the shaping by sensation of perceptual representations. The account of the process of actualization as one that is driven by intrinsic and extrinsic constraints corresponds with the idea of physiological inhibition guiding the extraction of featural detail. Essentially, elements are contrasts. On this theory, details are neither modeled nor assembled, but exposed, sculpted, or realized through a suppression of alternative routes.

A few examples are given from studies of language, action, and perception in adults and children:

There is a line of thinking in which features are not the elements of object construction but "emergent characteristics of form."[20] In children, phonemes are not concatenated into words but develop as emergents.[21] Gestalt recognition precedes feature analysis in studies of object (face) identification. Word categories are established early in lexical acquisition with derivation of specific words out of categorial representations. The acquisition of word-meaning in children shifts from representations based on characteristic features to those based on defining features,[22] a process similar to the analysis of wholes or the articulation of concepts from the primitive to the scientific.

In aphasia, disorders of word meaning show a zeroing-in on lexical targets in the process of word-finding. Error patterns confirm a specification from wide to narrow semantic relatedness as the word is finally selected. The emergence of the skeletal frame of the word evokes a specification of phonemes out of phonological gestalts[23] as word sounds fill in the "slots" of abstract lexical representations.

In action, infants show progressive differentiation beginning with proximal and axial movements and continuing to the specification of distal grasp and manipulation.[24] Inhibition plays a prominent role in the development of reaching in human and monkey infants.[25] In frame–content theory, a postural to analytic shift underlies the evolution of manual asymmetry.[26] Digital movement is elicited out of background axial and proximal motor systems. The proximal setting provides the context in which the distal movement develops.[27]

"Blindsight" phenomena[28] reveal substantial perceptual ability in extrastriate regions and suggest that patterns are analyzed into their con-

stituent features. The many instances of implicit perception[29] reflect holistic or contextual properties of stimuli subsequently analyzed into conscious or explicit perception. Holistic and analytic are in a relation of precedence, a shift from context to item, or whole to part, not simply the product of asymmetric or parallel brain systems.

Vision in the newborn is controlled by the subcortex with cortically mediated behavior coming into play during the first few postnatal months. This shift coincides with increasing selectivity and awareness of spatial detail, discrimination and orientation. In some respects such as, fixation, the behavior of a one-month-old is similar to an adult with striate lesions.[30] The process corresponding to this shift has been described as an emergence of adult patterns of connectivity through the refinement of an initially diffuse set of connections.

Summary

The generative principles of developmental growth include a process of proliferation and elimination, parcellation, in the establishment of connectivity in the brain. This principle is akin to the evolutionary process of adaptation in which unfit exemplars are pruned by the environment. A process similar to that of parcellation is described in physiological studies of neuronal populations. This process is a shift from a diffuse to a focal organization in which inhibition plays the role in function that elimination played in growth. This process recurs in context-to-item shifts in cognition.[31]

According to microgenetic theory, the multiple levels (phases) of mind are generated through a progressive context:item shift that retraces patterns in phylo-ontogeny; i.e., one operation lays down multiple levels (phases) rather than multiple operations acting on separate contents. Content is specified through a phased transformation at successive moments in the mind–brain state, which itself is cyclic and recurrent. There are points of contact with evolutionary models, e.g., Goodwin's[32] description of limb formation in which elements are specified along a proximodistal gradient as solutions in a single periodic generative process. A context-item shift is an actualization process, not a sequence of self-duplications. Still, there is a resemblance between this process and fractal geometry.[33]

HETEROCHRONY

Heterochrony is a temporal disparity in the development of organ systems. The timing of development is uneven. The change in rate can

affect a process that is focal and delimited or one that is widespread and pervasive, with the result that organs develop at different rates. Both retardation and acceleration occur. The result is that adult and juvenile features are loosely bound in covariant sets.[34] In addition to the rate of process, the timing of onset and offset is important.

Neoteny is a form of heterochrony in which a retardation of development prolongs the duration of a juvenile stage. Neoteny can be of adaptive value, when juvenile features survive in the adult as an escape from the more rigid specializations of adult structure. Gould[35] has argued that neoteny is an important mechanism in human evolution. Some human features reminiscent of a juvenile primate include a more upright posture and flat face, high brain to body weight, absence of brow ridges, thin skull bones, central position of the foramen magnum, and reduced body hair. The juvenilization of morphology accompanies a retention of physiological features associated with juvenile ancestral stages. For example, increasing brain size over the primates is associated with prolongation of sexual immaturity, from 2 years in the lemur, to 7 in the great apes to about 14 in man.[36]

Heterochrony can lead to successful adaptations, but it can also lead to developmental abnormalities. Serres[37] argued that when different parts of the fetus develop at different rates "monstrosities" (*monstres par défaut ou excès*) can arise if certain parts lag behind and retain, at birth, the character of a lower animal. In his account, Serres combined a recapitulation argument with differences in the timing of development of different organs. The recapitulation is not for "lower animals" but embryonic stages, while the maladaptations that result from altered timing may be subtle and qualitative, not just omission and excess.

One implication of the theory of heterochrony is that a rate change early in development should have more generalized effects than one late in development. Another implication is that an altered timing of brief duration, that leads to anomaly rather than truncation, should not exclude a normal *subsequent* development, with the resultant deformation carried into future stages as a signature of the altered phase. For example, delayed closure of the cranial sutures is a neotenous feature that permits the rapid expansion of a similarly neotenous brain. Closure occurs but at a later phase in development. In this example, the delay is associated with other features that identify the phase in ontogeny where the retardation began. The focal change, and the context of the change, identify the phase. The question is whether an error from a disparity in the timing of developmental process is comparable to a symptom of brain damage; i.e., whether early (developmental) and late (cognitive) errors can be explained by heterochronic change.

Efforts to establish a correspondence between the maturational sequence and the pattern of errors in adult decomposition have floundered on the misconception that disorders of adult cognition should recapture stages in acquisition, the "regression hypothesis." This approach, at least in the work of Hughlings Jackson[38] or Roman Jakobsen,[39] has been disconfirmed by the finding of weak correlations between, say, grammatical errors in aphasics and stages in the acquisition of grammar. It is true that pathology does not unpeel the acquisitional sequence, but this obscures the deeper truth that mental process reiterates developmental **trends**, not facts; i.e., the commonality of development and cognition—morphogenesis and microgenesis—is in process, not content.

HETEROCHRONIC PRINCIPLES OF ERROR ANALYSIS

Lesions and Errors

The principle effect of a brain lesion is to retard process, not destroy function. The effect is not an ablation of elements in a circuit board, but a change in a configuration that is a type of traveling wave. There is evidence for such "waves" in studies of brain development.[40] The lesion is comparable to an obstruction in a river (Figure 11.1). The obstruction impedes and delays flow but does not interrupt it. The retarded segment persists as a local disparity or dyssynchrony in relation to the cross-section of the stream (process) in which it is embedded. Unlike a river that leaves the obstruction behind, mental process is recurrent, so the obstruction is encountered anew in each traversal.

The nature of the error is determined by the stream (component) involved, e.g., action, perception, language, and the extent to which that segment of the stream has developed prior to the disruption. The error depends on the dynamic of flow and differs slightly for each traversal. The segment corresponding to the error is only approximated by the error type. The error category or pattern identifies the damaged segment. Conversely, error pattern, together with the location of damage, provides a basis for reconstructing the normal sequence.

Errors reflect perturbations as well as "obstructions." In a river, eddy currents develop around an obstruction. A lag in mental process may produce local eddies around the damage point. Such eddies and whirlpools have been described mathematically in Hopfield simulations[41] which link errors to normal processes. There might be analogous conditions in development, where a local delay impacts on regional systems

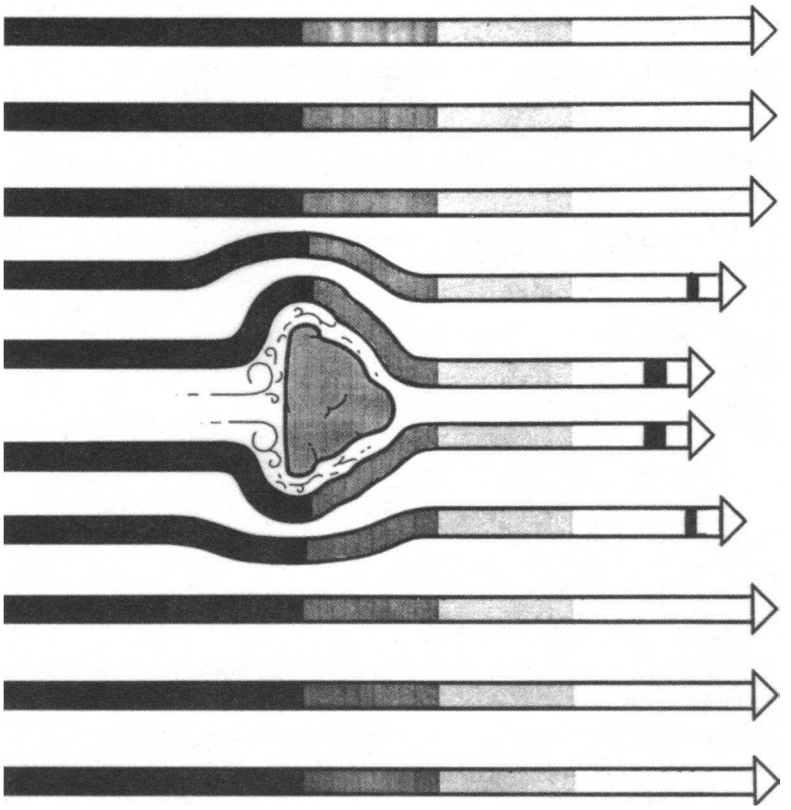

FIGURE 11.1. The lesion delays flow in a segment of process. The retarded (neotenous) segment is the symptom; the unobstructed flow parallel to the segment is the normal context in which the symptom occurs.

not directly involved by the defect. Still, a perturbation effect should be comparable to a lesion effect; i.e., retardation (or acceleration) of process.

A disparity in the timing of a process that advances like a wave front should lead to a local delay (or acceleration) that is "out of synch" with concurrent streams of process. The local delay is the error and the concurrent streams establish the normal context within which the error occurs. An (adult) error is abnormal because it is out of context with ongoing behavior. This is the meaning of abnormality. If a person says *chair* when he should say *table*, the error involves word choice but also the fit between the word and the context in which it is used. Hallucination during sleep is normal, but pathological during wakefulness. The dislocation between a focal deviance and a normal surround is the basis for the pathological symptom. A few examples of error types illustrate this concept.

Aphasic errors are preliminary normal stages. For example, a substitution such as *chair* for *table* reflects a phase of equivalence for either item. The error reflects the semantic category prior to the individuation of the correct item. The deviation reflects an attenuation (juvenilization) of the process of lexical realization. In this example, parcellation and neoteny combine in the symptomatology. Incomplete fractionation is a result of a local neoteny in process. The delay produces an incomplete specification with the naming error revealing the context prior to the intended item. Even a correct name ("table") can be an error, when a holophrastic noun does not achieve the referential or denotative specificity of a fully individuated word. The name is used to label objects that are not usually incorporated within its semantic field. There are similarities with early language learning, where nouns are used for objects in a class, e.g., "dog" for many different animals.

In amnestic cases, the inability to revive an event reflects the retarded activation of that event in the stream of ongoing mentation. The retardation leads to omission or substitution errors according to the depth of the delay. An amnestic gap points to the survival of a preliminary (unrealized) segment into mature (end-stage) cognition. That the event is still active we know from intrusions (delayed recurrence), implicit memory, cueing and contextual facilitation, affective correlates, personality growth, and the potential for recovery.

Patients with perceptual disorders can have errors of object meaning, e.g. the (perceptual) misidentification of a knife as a fork, but the object (knife) is seen (or drawn) properly. The error in such cases reveals the underlying object concept, while subsequent analysis of form is unaffected. In other cases, form is involved, e.g., a knife perceived as a stick, with preserved object meaning. Object detail does not individuate and objects are perceived on the basis of size, shape relations, etc. In both instances, an error of object concepts or object form, a background phase is carried into endstage cognition. Put differently, a focal "juvenilization" of process persists as a symptom of the retarded segment.

In sum, pathological symptoms are focal attenuations in action, perception, or their derivations into language, at various depths in the mental state. The focal delay is a brief neoteny of a covert phase that survives in behavior. The preservation (normal rate) of process parallel to the focus establishes the normalcy of behavior apart from the altered segment. Intact processing prior to the delay, i.e., the depth of the lesion, establishes the error type within the component. Since the process laying down the mental state is reiterated, and the direction of processing is obligatory, the involved segment is traversed in every iteration.

Errors: Developmental and Acquired

The relation between developmental deficits in children and acquired deficits in adults is complex. As mentioned, in adults the deviance is in relation to the context. The delay occurs at a segment in a mature system. The maturity influences the error since the error samples the context. In contrast, errors in young children reflect a still-forming context. The shift from context to item occurs—context is critical in learning—but since the context is limited the distribution of errors is reduced. This is why errors are more predictable in children than adults. Since the context is impoverished, the deviance has to be sought for in the relation to what is age-appropriate in other children, or to better performance in another (intact) domain of function. An error will respect the structure of a system, or a primary component of a system, at the point of the delay. In young children, the juvenilization tends to affect a system or component as a whole.

Consider, for example, the developmental language disorders. In such cases, the disturbance is usually generalized at first, and resolves into a disorder of production or comprehension. This can be interpreted as delayed maturation of language generally, i.e., the linguistic derivation of action and perception, with a residual deficit in one of these domains (action or perception), whichever is most delayed.[42] There is considerable evidence that motor or perceptual deficits play an important role in the etiology of developmental language disorders, with little support for a disturbance in a language acquisition device.[43] A disparity in the rate of development could affect either (or both) action and perception or different phases or epochs in the development of these components. This could give the heterogeneity that has been associated inter alia with age and severity; i.e., the rate of language development or the timing of the onset of the delay. In any event, symptoms still reflect a general or local alteration in the **rate** of development.

Typically, errors in developmental language disorders are abnormal only by virtue of being inappropriate for a child of that age. Simplifications are the rule. Performance tends to be characterized in terms of norms; e.g., a 3-year-old at the stage of an 18-month-old. Errors that violate rules of normal development, other than timing, are rare and controversial. Contextual cues are less effective in children than adults. These observations suggest that the difference between the developmental and the acquired reflects neotenous change in the initial formation of a system as opposed to a segmental delay in the reinstantiation of a system already formed.

For many, the acquisition and disturbance of syntax is a critical issue, even if perceptual deficits explain most syntactic errors.[44] Syntax

involves the extraction of relations between levels in mental structure. This raises the question whether context–item shifts at successive moments in the unfolding of the mental state—as transitions across nested or embedded units—are process-equivalents of the core rules of syntax.

Double Dissociation

A process can be retarded at one segment without an effect on a subsequent segment. For example, an alteration of timing in development gives syndactyly or webbing of the digits but spares the ensuing stage of nailbed formation. The defect labels a segment of process but the process continues to develop. Naturally, a severe disruption can abort development (or mentation). But with focal alterations, ensuing phases may be normal though postponed.

The specificity to a segment in process, in development or cognition, relates to the concept of "double dissociation." This concept is important in (adult) neuropsychological study because it seems to show that functions are interactive and modular. An example would be a separate impairment of phonetic and lexical reading. If a lesion separately disrupts phonetic and lexical reading, these performances are presumably independent; i.e., one operation is not contingent on the other. Obviously, a double dissociation pertains to functions that are cognitively "close." Limb paralysis and reading dissociate, but this is not of interest. The dissociation suggests interactive systems and seems to refute the idea of serial processing for the involved symptoms; i.e., if A goes to B, damage to A should prevent the occurrence of B. However, the concept of a symptom as a local delay is compatible with a dissociation of serial elements.

Thus, in the example of lexical and phonetic reading, if phonetic reading derives or develops from lexical reading, disruption of the prior segment would give errors through local retardation. The patient reads the word horse as, say, ***zebra***. Lexical realization is attenuated; the selected item falls within the word category. The symptom is a persistent "juvenile" (preliminary) feature. The successive phase of phonetic encoding would then occur on a deviant lexical form. In this example, the disruption (lag) displays a segment of process as an incomplete transform without affecting a subsequent phase. The subsequent phase is spared because the deviant content on which it develops is premature, not "abnormal."

In children with developmental language disorders, the dissociation is less emphatic; e.g., lexical–semantic disorders tend to be accompanied by disorders of phonology, though isolated phonological disorders occur. This is due to the lack of sufficient context for errors of derailment to

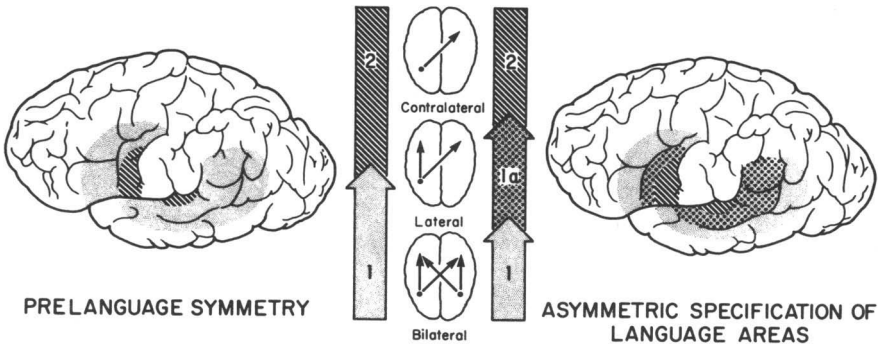

FIGURE 11.2. The emergence of the Wernicke and Broca zones within the "integration" neocortex represents neotenous expansion in brain development corresponding with the development of cerebral asymmetry through parcellation. Neoteny and parcellation combine in the emergence of asymmetric language representation. This phase is interposed between bilateral (limbic) and contralateral (motor cortex) representation.

provide a substrate for ensuing transformations, and the graded nature of development, so that preliminary phases must be accomplished before subsequent stages can begin. A prolonged delay puts ensuing stages at risk. Indeed, the linkage between deficits in children is greater evidence for their continuity than is the specificity of deficits in adults for their autonomy.

Heterochrony and Creativity

Heterochrony has been discussed in relation to pathology but is an important feature of normal cognition. On the microgenetic theory, a reminiscence (memory image) is an attenuated perception; inner speech is a preliminary utterance. Metaphor, imagery, and creative thinking (p. 241) are elaborations at submerged phases in mentation. A preliminary interior phase of perception or language survives into consciousness. More precisely, self-awareness and mental content, i.e., the "space" of introspection, owes to a juvenilization in microgenesis with anticipatory stages coming to the fore as final forms. The brief delay permits elaboration of segments that are normally analyzed into behavior.

Neoteny is an important mechanism in the evolution of the human mind–brain. The persistence of earlier features accounts not only for the dramatic expansion of the brain but the emergence of regions unique to man, such as the language areas, within the "integration" neocortex (Figure 11.2), a "growth plane" preliminary to the sensorimotor cortex in forebrain evolution.[45] The language areas arise as preliminary phases to mediate an interior cognition that is also preliminary. Self and

introspection are not appended to the repertoire of cognition, but are accentuations of preliminary phases in mental process. Neoteny is an evolutionary and cognitive mechanism that generates novelty through delay at earlier form-building stages.

Summary

The principle of heterochrony (neoteny) can be applied to the analysis of "lesion" effects in the developing and mature brain.

Symptoms reflect a local or generalized delay in process. The major change appears to be retardation though acceleration is theoretically possible. The delay results in an incomplete transition through the damaged segment with preliminary (juvenile) features carried to subsequent phases. Such phases are postponed, but may occur normally on deviant (attenuated) forms. This occurs more commonly with acquired disorders in adults than developmental disorders in children. On this basis, the significance of double dissociation for the imputation of functional autonomy is rejected.

The onset of a microgenetic series (mental state) is presumed to be driven by a pulselike activation (pacemaker) that could vary as to pulse frequency. Offset can be truncated, as in sleep and dream or in coma and hallucination. Acquired errors in adults represent a dyssynchrony between the local retardation and the contextual surround. Errors reflect context at normal preliminary phases that unduly persist. In developmental language disorders, there is a retardation of the fine derivation of perception and/or action into the complementary language components. Contextual frames for errors are weakly established and opportunities for derailment reduced, so error "tokens" show less variance than in adults.

Retardation at "juvenile" phases in mental process permits expansion at subsurface levels. The elaboration of preliminary phases in the development of acts and objects populates the mind with intrapersonal content, creating an intrapersonal space of introspection prior to the exteriorization of object space. Awareness, self, choice, and imagery are due to the postponement (retardation) of mental process at earlier phases. The ontogeny of human brain and mentality reflects neotenous change in microgenetic process.

CONCLUSION

This chapter has examined the relevance of two morphogenetic phenomena, parcellation and heterochrony, to processes underlying mature

cognition. It is argued that the specification of elements in growth establishes a physiological dynamic that continues as the specification of contents in mind. Growth is not an open-ended sequence but a reiterated dynamic. The parcellation that is reiterated in growth continues in whole-to-part or context-to-item shifts in adult mentation.[46] A process that is stretched out (but is, in reality, *repeated*) over ontogeny lays down a "track" for the momentary unfolding (microgenesis) of the mental state.

In development, structure is deposited as a milestone of change. The slow development of structure creates an historical perspective that is difficult to resolve with the idea of growth as a resurgent process. Yet growth can be interpreted as a cyclic or reiterated dynamic. Cell migration begins in the periventricular region with the deeper layers of neocortex established prior to the superficial ones. Rhythmic patterns of depth to surface layering occur through successive waves of proliferation, migration, and pruning. After the morphology has been established, these patterns persist as "force-lines" which determine the pattern of physiological activity. The pattern of growth and physiological activity continues into microgenetic process, which deposits the depth to surface "layers" of mental structure.

Specification and heterochrony are related as the pattern of a process to its rate. Retardation leads to incomplete specification at a given segment, while progressive arborization (context–item transforms) requires that every branch (fractal?) issuing from a context remain in phase with other parallel branches; i.e., that the rate of specification is uniform across concurrent segments.

As the saying goes, timing is everything. The cognitive process (microgeny) can be considered a speeded-up ontogeny. But what aspect of ontogeny is speeded up? Microgeny does not retrace ontogeny; that would require progressive delay in the duration of recapitulation or continuous acceleration to accommodate new layers of acquisition. It is not stages that recur but the process leading to the stages. The recapitulation is for the process, not its actualized elements.

The timing of process is critical. The notion of time is bound up with the nature of process. Acceleration, retardation, and recapitulation are temporal concepts. Indeed, the shift from a context to an item is like the analysis of a category. A category is like a duration. Duration and categorization are abstract entities with vague boundaries and indistinct centers. The passage from a category to an instance in that category is comparable to the isolation of instants in a duration of time.[47] Exemplars in categories are virtual, like instants in a duration: both categorization and duration are fundamental properties of mind. The nature

of time and duration are crucial problems for any theory of physical growth (morphogenesis) and mental process (microgenesis).

The idea that ontogeny retraces embryonic stages in evolution implies that deep phyletic time is collapsed in developmental or lifespan process. In fact, phylogeny is not an extended sequence but a reconstruction of a series of ontogenies. Ontogenetic process is the primary reality with phylogeny an extrapolation from the ontogenies of a given line. What, then, is the time of an ontogeny? Is it the lifespan of an organism, a day, a year, a century? Alfred North Whitehead thought the duration of a thing (atom, object) was the minimal period for the thing to be what it is. I would argue that the duration of an ontogeny is the minimal duration of the process that sustains it and that this process is replicated throughout the life of the organism. In other words, a process of some duration is repeated—a minimal unit of growth (later, of mind)—with ontogeny the sum of the repetitions of these units. With each iteration, the organism changes. The sequence of change appears to be the effect of a lifespan process, when actually the sequence is the pattern of change the process lays down.

NOTES

1. The genes encode amino acid sequences through RNA. Moreover, the increase in DNA over the mammalian series is nonlinear in relation to the increase in complexity of the brain.
2. G. Stent, "Strength and Weakness of the Genetic Approach to the Development of the Nervous System," *Annual Review of Neuroscience* 4(1981):163–194.
3. Brown, *Self and Process*, 41–45; G. Streidter and R. Northcutt, "Biological Hierarchies and the Concept of Homology," *Brain, Behavior and Evolution* 38(1991):177–189.
4. K. Pribram, *Brain and Perception* (Hillsdale, NJ: Erlbaum, 1991), 25.
5. C. von de Malsburg and W. Singer, "Principles of Cortical Network Organization," in *Neurobiology of Neocortex*, ed. P. Rakic, and W. Singer, (New York: Wiley, 1988). R. Thatcher, "Cyclic Cortical Reorganization: Origins of Human Cognitive Development," in *Human Behavior and the Developing Brain*, ed. G. Dowson and K. Fischer (New York: Guilford Press), 232–268.
6. D. Tucker, "Developing Emotions and Cortical Networks," in M. Gunnar, and C. Nelson (Eds.), *Minnesota Symposium on Child Psychology, Volume 24. Development and Behavioral Neuroscience* (75–128). (Hillsdale, NJ: Erlbaum, 1992); J. W. Brown, Overview. In *Neurobiology of Higher Functions*, ed. A. Scheibel and A. Wechsler (New York: Guilford, 1990).
7. B. Goodwin, "Changing from an Evolutionary to a Generative Paradigm in Biology," in J. Pollard (Ed.), *Evolutionary Theory: Paths into the Future* (New York: Wiley, 1984), 99–120.
8. M. Katz, "Ontophyletics: Studying Evolution Beyond the Genome," *Perspectives in Biology and Medicine* 26(1983):323–333.
9. G. Edelman, *Neural Darwinism* (New York: Basic Books, 1987).

10. P. Rakic, "Competitive Interactions during Neuronal and Synaptic Development," in *From Reading to Neurons*, ed. A. Galaburda (Cambridge, MA: MIT Press, 1989); P. Rakic, "Developmental Origin of Cortical Diversity," Schmitt lecture, Rockefeller University (1992). The basis for parcellation is unclear. There is evidence that the activity of the postsynaptic cell regulates the stability of the synapse in a retrograde manner. See J.-P. Changeux, *Neuronal Man*, trans. L. Garey (Oxford: Oxford University Press, 1985).
11. S. Ebbeson, "Evolution and Ontogeny of Neural Circuits," *Behavioral and Brain Sciences* 7(1984):321–366.
12. G. Innocenti, "Commentary," *Behavioral and Brain Sciences* 7(1984):340–341.
13. S. Witelson, "Structural Correlates of Cognition in the Human Brain," in *Neurobiology of Higher Functions*, ed. A. Scheibel, and A. Wechsler (New York: Guilford, 1990).
14. O. Creutzfeldt, "Generality of the Functional Structure of the Neocortex," *Naturwissenschaften* 64(1977):507–517.
15. Ebbeson, "Evolution and Ontogeny."
16. G. Coghill, *Anatomy and the Problem of Behavior* (New York: Hafner, 1964).
17. J. Semmes, "Hemispheric Specialization: A Possible Clue to Mechanism," *Neuropsychologia* 6(1968):11–26; J. W. Brown, "Lateralization: A Brain Model," *Brain and Language* 5(1978):258–261.
18. P. Wall, "Recruitment of Ineffective Synapses after Injury," in *Functional Recovery in Neurological Disease*, ed. S. Waxman (New York: Raven Press, 1988).
19. A. Karmiloff-Smith, "Beyond Modularity: Innate Constraints and Developmental Change," in *The Epigenesis of Mind: Essays on Biology and Cognition*, ed. S. Carey and R. Gelman (Hillsdale, NJ: Erlbaum, 1991).
20. Pribram, *Brain and Perception.*
21. C. Best, "The Emergence of Native-Language Phonological Influences in Infants: A Perceptual Assimilation Model," *Haskins Laboratory Status Report on Speech Perception.* SR 107–108:1–30.
22. F. Keil, "The Emergence of Theoretical Beliefs as Constraints on Concepts," in *The Epigenesis of Mind: Essays on Biology and Cognition*, ed. S. Carey and R. Gelman (Hillsdale, NJ: Erlbaum, 1991).
23. E. Perecman and J. W. Brown, "Phonemic Jargon," in *Jargonaphasia*, ed. J. W. Brown (New York: Academic Press, 198, 177–258.
24. C. Trevarthen, "Emotions in Infancy," in *Approaches to Emotion*, ed. K. Scherer and P. Ekman (Hillsdale, NJ: Erlbaum, 1984).
25. A. Diamond, "Developmental Time Course in Human Infants and Infant Monkeys, and the Neural Bases of Inhibitory Control in Reaching," in *The Development and Neural Bases of Higher Cognitive Functions*, ed. A. Diamond. *Annals of the New York Academy of Sc*ience 608(1990):637–69.
26. P. McNeilage, M. Studdert-Kennedy, and B. Lindblom, "Primate Handedness Reconsidered," *Behavioral Brain Sciences* 10(1987):247–303.
27. G. Goldberg, "Supplementary Motor Area Structure and Function:Review and Hypotheses," *Behavioral Brain Sciences* 8(1985):567–616.
28. L. Weiskrantz, *Blindsight: A Case Study and Implications* (Oxford: Oxford University Press, 1986); L. Bard, "De la persistance des sensations lumineuses dans le champ aveugle des hémianopsiques," *Semaine Médicale* 25(1905):253–55. M. Bender, and H. Krieger, "Visual Function in Perimetrically Blind Fields," *Archives of Neurology and Psychiatry* 65(1951):72–99.
29. A. Reber, "The Cognitive Unconscious: An Evolutionary Perspective, *Consciousness and Cognition* 1(1992).

30. R. Van Sluyters, J. Atkinson, M. Banks, R. Held, K. Hoffmann, and C. Shatz, "The Development of Vision and Visual Perception," in *Visual Perception: The Neurophysiological Foundations,* ed. L. Spillman and J. Werner (New York: Academic Press, 1988).
31. Brown, "Overview" (1990).
32. B. Goodwin, "Development and Evolution, *Journal of Theoretical Biology* 97(1982):43–55.
33. P. MacLean, "Neofrontocerebellar Evolution in Regard to Computation and Prediction:Some Fractal Aspects of Microgenesis," in *Cognitive Microgenesis: A Neuropsychological Perspective*, ed. R. Hanlon (New York: Springer-Verlag, 1991); D. Robertson, "Feedback Theory and Darwinian Evolution," *Journal of Theoretical Biology* 152/4(1991):469–484; L. Vandervert, "Systems Thinking and a Proposal for a Neurological Positivism, *Systems Research* 5(1988):313–321; 7(1990):1–17.
34. S. Gould, "Change in Developmental Timing as a Mechanism of Macroevolution," in J. Bonner (Ed.), *Evolution and Development* (pp. 333–346). Berlin: Springer-Verlag, 1982.
35. S. Gould, *Ontogeny and Phylogeny* (Cambridge, MA: Harvard University Press, 1977).
36. J. Bonner and H. Horn, "Selection for Size, Shape and Developmental Timing," in *Evolution and Development*, Dahlem Conference, ed. J. Bonner (Berlin: Springer-Verlag, 1982), 259–276.
37. Serres (1860), cited in Gould, *Ontogeny and Phylogeny.*
38. J. H. Jackson, *Selected Writings*, 2 vols., ed. James Taylor (London: Hodder and Houghton, 1931).
39. R. Jakobsen (1968) *Child Language, Aphasia, and Phonological Universals* (The Hague: Mouton, 1968).
40. R. Thatcher, "Are Rhythms of Human Cerebral Development "Traveling Waves"? *Behavioral Brain Sciences* 14(1992a):575.
41. R. Hoffman, "Computer Simulations of Neural Information Processing and the Schizophrenia-Mania Dichotomy, *Archives of General Psychiatry* 44(1987):178–188.
42. B. Sahlén, *From Depth to Surface: A Case Study Approach to Severe Developmental Language Disorders.* Studies in Logopedics and Phoniatrics No. 1, Lund University (1991).
43. E. Bishop, "The Underlying Nature of Specific Language Impairment," *Journal of Child Psychology and Psychiatry* 33(1992):3–66.
44. Bishop, "Language Impairment."
45. F. Sanides, "Comparative Neurology of the Temporal Lobe in Primates Including Man with Reference to Speech," *Brain and Language* 2(1975):396–419.
46. The reverse process, an item to context or part to whole transition, is a more complex issue relating to emergence and the anisotropy of time (see Chapter 12).
47. *Self and Process*, 39.

CHAPTER 12

Emergence

ARGUMENT: *Synchronic emergence across categories, e.g. the mental from the physical, is insoluble because the concurrence of events eliminates the time for intervening process. Diachronic emergence within a category, e.g., a chemical reaction, evolution, is a process of part-to-whole transformation. The emergence (synthesis) of parts to wholes is not an assemblage of parts to aggregates, but the perishing of partlike wholes as they are replaced by larger ones that also perish. Every whole embodies a potential implicit in the capacity of its parts to be united. The synthesis of parts to wholes is their replacement by this potential. In this sense, the whole is prior to the parts. The unity of an object is a temporal phenomenon. Temporal order appears in the surface progression of nature. More deeply, it is the passage through the precedence in becoming.*

> Nature always follows an analytic course—
> development out of a living mysterious whole.
>
> GOETHE

For many, emergence is mysticism in the cracks of scientific progress. Because it has been employed in many different ways and for relations of mystery and uncertainty, the concept of emergence like that of entelechy or elan vital is commonly dismissed as obscurantist and prescientific. The emergent step is claimed to be a resultant even if its causal history cannot be specified. Or, it is maintained, emergent organizations entail novel arrangements of atoms that in principle can be derived from a knowledge of their properties.[1]

The difficulty arises in part because emergence is invoked for fundamental problems that confound scientific understanding, such as the

transition from inorganic matter to life, or the relation between the physical and the mental, as in the appearance of consciousness from brain activity. In systems theory, emergence characterizes the transition from one level of organization to the next, for example from the intracellular to the cellular, from cell to tissue or from neuron to network.

That emergence is a limit phenomenon waiting for causal explication is not the sole reason for skepticism. The deeper reason, having to do with the notion of time, is that causal science is reluctant to accept the possibility of genuine or nonprobabalistic novelty as an outcome of change. Causation is a theory of change that is centered in relations *between* entities. A change from one entity to another cannot be emergent if change is located at the interface of the pair. Emergence entails that entities are the embodiments of their relations. Entities do not undergo emergent change in a pursuit after novelty but are the outcome of a flux of relations through which the entities evolve.

Emergence, therefore, is linked to the concept that relations are primary, i.e., that entities are not changing to other entities but are the products of change, that change deposits momentary entities. Emergence is one way of characterizing change, causation another. If all change is fundamentally the same, i.e., if there is a common basis for change in the universe, emergence would not be *intermittent*. Either change is never emergent or continuously so. Whitehead wrote, "causal connection is merely one typical instance of the universal ruin of relatedness."[2] For emergence to be an alternative to causation, relations must be conceived as more fundamental than entities.

Part and Whole

Emergence can be examined in the relation between part and whole where the whole is not explained as the causal effect of the action or interaction of the parts. Nor is the whole exhausted when its constituents are expressed. Even if the constituent relations are included, e.g., binding properties, the whole, as von Ehrenfels[3] said, is still greater than the parts. Put differently, the properties of the whole cannot be explained by the individual and relational properties of the parts.

Wholes cannot be identified with their constituent structure but parts are equally problematic. The parts of a whole are themselves wholes that can undergo further analysis. An atom is a part of a molecule, but it is a whole in relation to its subatomic structure. With the advent of superstring theory, the whole to part analysis is essentially bottomless.[4] The part is more than an element of a whole; i.e., the properties of the part are not predicated on, or ingredient in, the properties of the whole. If

the transition from part to whole is emergent, so is the transition from whole to part. One reason for confusion over the part–whole problem lies in the distinction that is made between parts and wholes. The problem is the distinction. It is a mistake to assume that parts in isolation are the same as when embedded in a field. An isolated part is itself a field until its particularity is usurped by the field in which it is embedded. A part of a *natural* whole, i.e., a unitary whole that is not a mere aggregate, loses its particularity when it is bound to other parts of the same whole.[5] The boundaries that separate isolated parts dissolve into zones of relations that bind the parts together. Every part is a whole to its constituents, every whole a constituent of a field. A whole is nested in a field, a part is a subordinate whole. Parts and wholes are whatever actually develops at a given moment. The greater whole is the field from which an entity materializes, no matter how large.

Synchronic and Diachronic Perspectives

When part–whole relations are viewed from a spatial or *synchronic* perspective, i.e., as a simultaneous lattice in a slice through change, there is a loss of the temporal relatedness of parts and wholes. Parts cannot be comprehended as a two-dimensional section of interlocking jigsaw pieces. The part–whole relation is grounded in process, in the becoming of the whole or the part. A synchronic perspective gives parts with boundaries that are artificial and wholes that are mere compilations, not the process through which parts *become* wholes.

Conversely, a temporal or *diachronic* perspective captures the transition from part to whole, or the reverse, and gives the relatedness, which is the process of binding and the changing configural properties of the parts and wholes, but the parts of a whole and the whole with its parts are not grasped at the same time.

Both spatial and temporal perspectives incorporate different types of part–whole distinctions. For example, *within* the same category, one can distinguish "levels" such as atoms and molecules or the evolution of a behavior. One can also distinguish "levels" *across* two seemingly different categories, such as inorganic to organic, or physical to mental. Micro and macroperspectives on a common substrate, e.g., mean kinetic energy and temperature, are of the latter (across category) type. Of course, "within" and "across" suggest that one can distinguish a breach of one category from that of another. This is not always possible. Indeed, the shift from the atomic to the molecular may not be inherently different from the shift from the physical to the mental.

Within-category emergence can have a micro to macrolevel quality and be synchronic, as an atom and its internal structure, or it can be diachronic, as in a chemical reaction. Across-category emergence is often treated as synchronic where the cooccurrence of states is collapsed to an identity of the categories. This strategy is common in philosophical approaches to the mind–brain problem. For example, a mental quality, say, color, is said to be identical to a physical state such as wave length, even though the mental properties of the color are qualitatively different from the physical ones, e.g., in the dissociation of color perception from its sensory induction or discontinuities in pathological cases between color and shape, and the uncoupling of color categories in perception from the continuous physical spectrum.

If something more needs to be said about the mental to accommodate its identity with the physical, i.e., if the physical fails to account for the range and quality of the mental, this "something more" is what is lost when cooccurrent properties are identified. What is lost is the process through which one set of events becomes another set of events, whether or not the process is emergent. The claim of identity, e.g., between the physical and the mental, ignores the uniqueness of cross-category instances and eliminates the (possibly emergent) process linking the two categories.

In physical systems, to take a common example, the relation of liquidity, which is a property of a whole, to the properties of single molecules in a liquid is a type of micro to macrolevel synchronic "emergence." The properties of the whole are held to be explicable in terms of the physical relations that obtain between elements or, in the case of the synchronic emergence of the physical to the mental, e.g., wavelength and color, the mind-dependent properties of wholes. The attempt to identify the properties on either side of a possibly emergent transition abolishes the transition and with it the emergence. Identity across two categories assumes a lack of relatedness between two sets of correlated events since time is required to accommodate the relation, whether it is conceived as an "interaction," a causal interface, or the realization of binding properties among parts. This temporal aspect, e.g. the time for a physical event to *cause* a mental event, sunders the synchronic perspective of identity theory and relations at-an-instant.

Emergence, therefore, confronts a more artful and certainly more parsimonious refutation by the postulation of an *identity* of parts and wholes or levels of description than by an effort to show that emergent change is incoherent or a cryptic form of causation. The attack on serial or diachronic emergence of the part to whole type requires a description of the putatively causal series of intervening steps leading from parts to wholes. Could this be done? At what point in their binding does an

assemblage of individual molecules of a liquid show the properties of liquidity?

An entity or event in isolation, i.e., a part such as a molecule of water, is not the same entity or event when it is bound to other entities. The entity, which is a whole to its parts, is lost the moment it is fused with, i.e., becomes a part of, something else. Wholes are not simultaneous with their parts, since wholes and parts continuously appear and vanish. The novelty or emergent change created in the relatedness of parts to wholes is displaced by an observer from the becoming of the parts to their aggregation. The change that characterizes the coming together of parts to wholes is located by an observer in the shift from an occurrent to a replacing state, which is the gap across becomings. In truth, however, this "gap" is an absence of change that is the glue of passage from one state to the next.

The incessant transition (becoming) to a part or a whole is the process of change. Emergence is linked to a theory of change, which determines the degree to which an emergent transition is plausible, given the constraints of the theory. If all change is causal, emergence is a deviation from the "laws" of nature. If all change is probabalistic, emergence is the probability of a deviation from an expectancy. If other "laws" such as those of microgenetic or fractal systems apply to successive microstates, emergence might well be ubiquitous. In a microgenetic system, every entity or event (the domain that is changing) is a becoming, i.e., an iterated whole to part transition. Change is in the replacing of an entity that no longer exists. In such a system "causation" would be the minimal deviation from an expectancy given to emergent change by the constraints on recurrent form.

Emergence in Cognition

Relations within and across phases in the mental state are those of nested whole to part transitions. These relations can be diachronic, e.g., gestalt to feature, category to item, semantics to phonology, or they can be constitutive or synchronic, e.g., lexical items in a semantic category.

In the first case, there is a progression from one qualitative state (e.g., semantics) to another (e.g., phonology). That this progression may be emergent is suggested by the fact that relations among elements in one state, e.g., words, do not appear to predict the transition to the next nor the relationships among the elements of the ensuing state, e.g., speech sounds. Even if we can describe certain of these relationships as context–item shifts, we still cannot specify the change at each

segment in the becoming sequence that corresponds with the emergent step.

In the second example, where the elements, say, names (lexical concepts) of animals, constitute the category (animals), the usual approach is a reduction of the category to its member items (e.g., dogs, cats, etc.). But the category is not the sum of what it contains, i.e., the sum of the members and/or their properties (attributes). Even if a category were to be exhaustively searched and depleted of its members, the category would still remain. The category is not discharged by counting every member, as the quality of liquidity is not grasped by enumerating every molecule in a fluid. The category is abstracted or averaged from its contents.[6] The abstraction of the category accounts for the superordinate relation to the members. The properties of the category can be defined as belonging to every member but not the reverse. The properties of the category are less rigid, or more inclusive, than those of the members. A member can belong to more than one category depending on its use or behavior. The properties of the members are not exclusive to one or several categories. A shoe can be a piece of clothing, a cup, or a weapon. The members and their properties are virtual in the category. They are realized individually in the passage from category to item.

When an item (say, the word or image *dog*) is in focus, the category recedes to the background. At that moment we are thinking dogs, not animals. When the category (animals, pets) comes to the fore, the item is a possibility. We are then thinking animals, and dog is the potential of the category to generate that item. Since the members issue from the category (the member "entities" and their properties are derived in a traversal through the potential of the category) the properties that are realized, whether of the category or the members, their specificity, and their "distance" from related members, depend on the extent of the derivation. Clinical study reveals the graded nature of item selection. Any member derived from a category is incompletely specified since an item always contains a potential for further analysis.

Member items are themselves abstract wholes that are not discharged in the realization of any or all possible physical instances (saying *dog*, seeing a *cat*). The word *dog*, as a lexical image in the mind, has millions of unique instantiations; it does not correspond with a specific object. In the passage from a semantic category to a lexical item to its phonetic realization (or from an object concept to an object to a feature of the object), there is a shift to a qualitatively different state in which it is often unclear what is a whole and what is a part of the whole. As noted, the properties of parts do not predict those of wholes, but the reverse is also true. If emergence applies equally to the shift from whole to part, i.e., if it is reciprocal, the analysis of a whole to its parts should

provide the flip of an account of the synthesis of the parts to the whole. Is this the case?

In perception, any arbitrary element is a local contrast in the "soup" of visual space. A shape is a field of color. Contrasts of color are the boundaries of objects. In pathological cases, the color of an object, thus the object, can melt into space. When this occurs the object dissolves into space or space is perceived as continuous with the object that is dissolving. In hallucination, one can "feel" the space between images. In dream and hallucination, the "empty" space of waking perception is an object of lesser density. The perceptual field is one object articulated into parts. It is a matter of contrasts within the field as to which featural elements differentiate out of which background.

For gestalt theory, it was axiomatic that the whole was greater than the sum of its parts (constituents).[7] There is evidence that the recognition of gestalts precedes the perception of details. The transition from whole to part is fundamental in microgenetic theory. In mind, perhaps in the physical world as well, there are continuous whole–part shifts. Mind is part of nature with the same lawful properties as natural phenomena. The microgeny of the mental state in a direction toward analysis is, presumably, an instance of becoming in the material world.

If the whole is prior to the parts, a part can only become a component of a whole by regaining the potential out of which it, the part, developed. Any part can serve as a point of departure for a return to, i.e., revival of, the whole. One view of the part to whole transformation, then, is that wholes are not created by the parts coming together but reappear in the perishing of parts. The wholes are antecedent to their parts. The parts perish and are replaced by novel wholes that are attenuated prior to a phase through which the parts would, in the next present, be specified. The "mechanism" of a part to whole shift is a truncated or more preliminary actualization of the whole to part transition. The whole is recaptured prior to a phase of actuality through which its previous parts emerged. This is not a reversal of the whole to part transition but a replacement of parts by the wholes that anticipate the next state of actualization. I believe this to be the case for emergence in the mental state (see below).

The potential that precedes an actuality, even if that actuality could serve as the potential for a further specification, is a whole that is prior to its actualized elements. The whole appears in the extinction of the elements (parts) of a prior actualization. The whole that follows the parts of the present actuality becomes part of an ensuing phase of analysis. A whole or field is constantly resolving to its parts, while parts are constantly perishing as the next field approaches.

Take again the mental category: *animals*. This category is not pieced together by an accretion of knowledge concerning its many singular instances or exemplars. We do not add the properties of dogs and cats and lions and tigers together to build up a category. Common features are extracted in the resolution of the particular, features that enlarge the category, but the extraction is *prior* to the individuation of features that are specific to the instance. In other words, the specific is not assembled to the generic but is delivered from the generic by a process of specification.

Categories expand from within. An event of infancy arises in a still forming category. When my 12-month-old son first saw a ball and it was named for him, he did not say the word *ball* until a few days later when a pea rolled from his plate at dinner. The category of, say, round objects is not created by adding two round objects together. With only one object there is no strategy to guide the additive process; every instance is unique. Rather, the perceptual category of roundness and its semantic equivalent (ball) configure the first instantiation, while successive objects average the category as they develop through it. The features of a specific object are achieved by delimitation in the actualization from category to world.

Philosophical Studies of Emergence

Synchronic Aspects

Most recent philosophical discussion of emergence has centered on the synchronic aspects, referred to as the many–one problem.[8] Baxter[9] attempts to resolve the many–one identity on the basis of the identity of different counts. The cross-mapping of the counts, i.e., a count of the multiplicity of the parts and a count of the oneness of the whole, is the identity of the two counts, but the concept of identity is left unspecified.[10] The "same" thing is not identical with itself at different times. If an entity (whether part or whole) is not the same entity a moment later, how can successive counts be correlated? The distinction between part and whole is not just perspectival. Even if one could observe a whole and its parts simultaneously over a segment of change, the whole would still not be predicated on its parts. That is because the parts can be specified in terms of their particularity or in terms of their relatedness, i.e., as objects or as processes, but not in terms of both.

The emergence of mental states or properties from brain states is usually considered an example of synchronic or across-category emer-

gence. A mental state arises from (corresponds to) a concurrent brain state. A complete account of the brain state—in theory—will still not yield the mental state, nor the reverse. Synchronic emergence of this type is probably insoluble. The insolubility is a function of the impossibility of "reducing" one state to another, other than by defining the properties (see below) both states have in common. As mentioned, this (synchronic) way of thinking evaporates the *process*, causal or emergent, by which a brain state *becomes* a mental state. The process is eliminated by the simultaneity of events or properties across the two states. There is no "temporal space" for one state to *become* another since the identity, the correlation or the epiphenomenal quality of the supervening (emergent) state is too tightly bound to the lower level to allow a *time* for a transition across states. The emergence "just is" or is not, with no explanation of how it occurs or is eliminated.

Mental states may be identical to brain states but for some philosophers, psychological properties could still supervene or emerge from, i.e., not reduce to, the physical properties of the brain. The emergent step is unidirectional, though in principle brain states could as well be emergents of mental states. It has been argued that emergence is compatible with identity theory. This requires that subjective properties survive the reduction, i.e., that the distinction of (some) mental and (some) physical properties is maintained.

According to the concept of supervenience,[11] a mental event depends on a correlated brain event but differs in some way from the brain event and is not reducible to it. Philosophical studies have focused on the possibility of multiple cognitive realizations of a given brain state or whether there could be a difference in a supervening state without differences in the subvening state. These studies have explored the conditions of difference or identity but not their bases.

Supervenience is a concept built on the assumption of identity. It does not obligate that the mental is identical to the physical except for properties that are irreducible. If the difference between the mental and the physical is inexplicable, why is the extent of this difference or identity not arbitrary, since the identity is no less speculative than the supervenience. While the concept of supervenience permits mental properties to be analysed without concern for their eventual reduction, the lack of explanation for the supervenient relation or emergent *step*, which must be accepted as an act of faith, as Alexander[12] put it, of "natural piety," vitiates its credibility.

A theory in which the emergence is *from* the physical *to* the mental (rather than the reverse), and where the mental is identical, but irreducible, to a physical state, contains an unavoidable residue of dualism. An identity between two sets of irreducible properties is incoherent. If

the mental is *caused* by the physical, the physical would have to precede the mental. But the commonality of mind–brain states in identity theory does not permit the temporal step that is necessary for causal or emergent events to occur.

Identity theory attempts to account for subjectivity by the postulation of "connecting principles" or "bridge laws" that relate mental states or properties[13] to their physical correlates. The supervening (emergent) mental states or properties are reduced to the more fundamental physical ones. Bridge laws are algorithms for translating the mental to (or from) the physical. A translation that entails events extrinsic to the bridged states, e.g., a transformational rule or operation that is applied to a state, is open to the above criticism. Bridge laws are hypothetical; they are not laws in the sense of physical science, but sets of relations no less mysterious than the mind–brain relation they seek to replace. In their absence, the possibility of emergence is justified by the resistance mental properties show to intertheoretic reduction (to brain). Put differently, emergence remains a possibility as long as there is a substantial disparity between what is resolved and what is "left over" after a causal explication of the relation between the physical and the mental is exhausted.

Diachronic Aspects

Emergence is frequently postulated by systems theorists to explain transitions over successive levels of organization. Writing from this point of view, Bunge and Ardila argue that every system has some global or systemic properties that are emergent relative to those of its components, i.e., properties that are lacking in the components. Emergence is widespread and must be approached on a historical basis. Molecular properties are explained in terms of their precursors and modes of binding. Biological assemblies are explained by reference to antecedent stages. Bunge adapts an emergentist (or systemic) reductionism in which mental properties develop as qualitative shifts at successive levels of organization.[14]

In a systems approach, emergence can explain the appearance of increasingly more complex organizations out of simpler ones, where the more complex involves a quantal shift not attainable through an expansion of the prior state. The shift, for example, from the intracellular to the cellular, from cell to tissue and organ, or from the individual to the social, whether in human behavior or the flight patterns of birds, subsumes the properties of a prior state in those of an "emergent" one, though the latter is qualitatively unique. The systems orientation is to macrodiachronic change, where ignorance of mechanism at the mi-

crolevel supports inconclusive argumentation on both sides, when what is needed is an examination of the minimal events surrounding the shift from part to whole from a temporal or diachronic perspective.

Some of these issues have been discussed by Van Cleve,[15] who is concerned with co-occurent sets but alludes to diachronics or serial change as the basis of emergent properties, which differ from causal properties that occur at the same time. In the relation of part to whole, the supervening properties *depend* on the properties of the parts without being *determined by* them. The relation of dependence is lawful, that of determination is logical. But what are the laws through which emergent properties arise?

Synthesis of Parts to Wholes

The relation of whole to part is not in the "block structure" of the whole but in the change that occurs over the separate times that wholes and parts are observed. When parts become wholes, their partlike properties change in the becoming. The loss of these partlike properties is the gaining of wholelike properties. This is an inherent feature of change into larger or smaller assemblies.

A part does not become a whole as an aggregate with other similar parts. Wholes are not created by accumulation. A rockpile is not a whole composed of rocks. The rockpile *in perception* has an *organic* wholeness. The unity of the rockpile reflects its holistic *mental* properties, not the combined physical properties of the individual rocks. A rockpile in perception is a sector of the visual field. The sector can deposit as a pile of rocks, or as a single rock surrounded by the whole portion of mental space that formerly was filled by the pile. This wholeness is the implicit ground of any potential part.

The part to whole problem relates to the idea of *synthesis* and *integration*. In neuroscience, these terms denote the binding (connectivity) of physical or psychic elements to larger assemblies; e.g., the compilation of reflexes to "higher" levels of organization. But how do reflexes come together to form more complex groupings? In the evolutionary progression from the spinal gray to the reticular formation, cell proliferation creates neuronal populations that can mediate more complex performances. But is the complex a composite or multiple of the less complex? If there is proliferation from a common stem, the unity in the system is present from its inception. Sherrington[16] wrote, "If the nature of an animal be accepted as being that of a whole presupposed by all its parts, then each and every part of the animal is integrative." Parts are wholes nested in larger wholes. The holistic properties of cell

populations or networks are replicated at the level of the nerve cell. The occurrence of holistic principles at all levels of organization is expected if the lowest level is a starting point in the fractionation of successive stages in growth.

Sherrington went on to write that, "behavior seems to become less and less reflex as the animal individual becomes more and more complexly individuated." The shift from reflex to more complex operations is through *individuation*. I take individuation to be a process of analysis through which part-functions are elaborated. This is a whole to part or context–item specification, not an augmentative or constructive process, which implies an aggregation of elements.

Synthesis as a form of emergence can be approached in the following way. If the parts A and B come together to produce AB, and if AB has properties other than A + B, this would seem to be an example of synthesis by addition. In this example, it seems obvious that AB is the sum of A + B, that the whole (AB) is a product of the union of the parts A and B. Yet, I have said the parts develop from the whole. How can A and B produce AB if the whole (AB) is prior to its constituents? The problem of priority, or precedence, therefore, and the account of precedence in a theory of change are crucial for an interpretation of synthesis.

To say that wholes are prior to their parts is to say either that the parts, A and B, are the realizations of an AB in the past of A and B or, more to the point, that A and B perish when bonded and are replaced by the potential implicit in their capacity to be united. A and B are wholes that disappear in the replacement by another whole, AB, that another moment can *become* A and B. As soon as A and B are joined, they exist no more as particulars. The moment of A and B as distinct entities has terminated. This is the perishing or extinction of A and B as actual entities. The next moment there is only AB. What is AB, but the potential to fractionate to A and B, or the fate to perish as a particular for the sake of another potential entity.

A and B do not change into AB. The parts do not combine to a whole, but are replaced by the next actuality. This actuality *appears* to contain the now nonexistent A and B, which have perished as actual entities, just as AB perishes to be replaced by another actuality. The change is not in the transition from A and B to AB, which are successive termini, but in the resurgence of AB over the extinct A and B of a moment ago (see Figure 12.1)

The fact that smaller wholes (A or B) may actualize before the larger ones (AB), to which they seem to combine, gives the impression of a part to whole synthesis, but every occurrent whole is perishing, no matter how large an entity. The whole has already individuated to its limit.

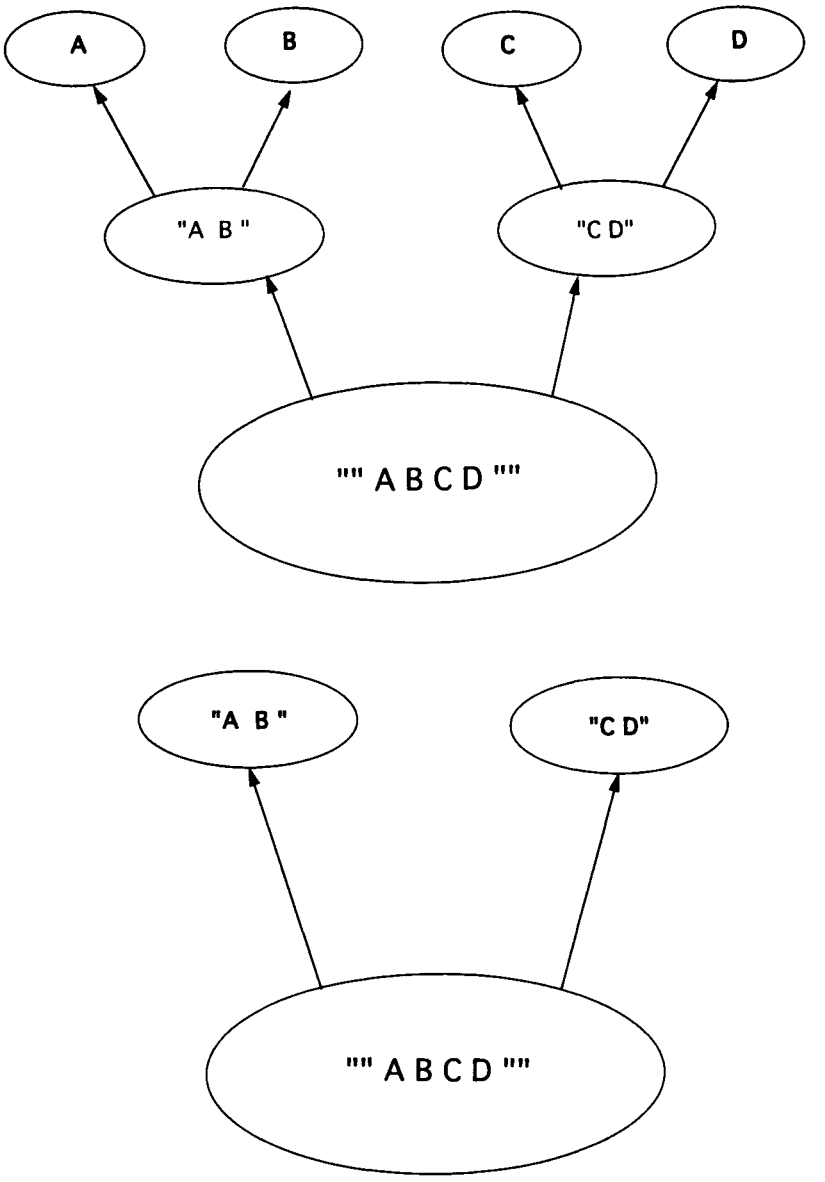

FIGURE 12.1. The whole ""ABCD"" gives rise to "AB" and "CD," which in turn individuate to A, B, C, and D. This is a depth to surface actualization in a single mental state (above). In the ensuing state (below), the specification proceeds only to "AB" and/or "CD." This appears to be a combining of A + B to "AB," and C + D to "CD," when it is actually a replacement of A, B, C and D by the antecedent wholes, "AB" and "CD." In these diagrams, A, B, C, and D are not copies of the content in "AB" and "CD," nor are the latter copies of ""ABCD"." Each phase is a potential to develop to the next.

This is its actuality. The next moment, other smaller or larger wholes appear. These too are termini, extinct the moment they actualize as "parts" of wider fields of potential.

Take an example in the physical world. When rain falls on an ocean or when a river flows to the sea, parts are added to a whole. When mist rises from the waters, wholes are given up to their parts. But should one not rather say, the momentary entities of rain and river and mist and ocean are continuously being replaced by novel entities that are reconfigured. The reconfiguring is change in the world as its entities actualize. Every entity is an unequal "part" that is realized within a "whole" field of potential.

The process of synthesis, then, can be defined (minimally) as the replacement of two partlike wholes by one wholelike part when the two parts perish and give way to a whole that is also perishing. Every actual whole is an about-to-be-extinct part of a wider field. The degree to which the whole actualizes out of this field determines whether it is a whole or a part. The process is always in the direction of greater partition. A part is a whole embedded in a wider whole, a whole is that part of the wider field that momentarily actualizes.

What is Potential?

To say that the transition from parts to wholes is the resurgence of potential within the parts is to say that the apparent coming together of parts is, in reality, the recurrence of a potential that does not actualize to the limit of the prior series. But what exactly is a potential? The potential of a whole has been interpreted in different ways, as a mysterious something in the background of the world, as a capacity or competence, or as a collection of preformed or incipient parts.

As a capacity or propensity to give rise to further elements, a state of potential can be described in terms of probabilities. Popper reviews the argument "that the physical parts constituting a new structure (like an organism) must possess beforehand the possibility or potentiality or capacity for producing the new structure in question." To avoid the problem of preformation, he replaces the concept of potentiality with probability or probabilistic propensities. A propensity would seem to be a potential with some direction or bias. The bias is important since a probability cannot be applied to a directionless potential. A propensity has a quantitative aspect that incorporates a probability. On this account, potential is the probability for a given propensity to be realized.

The likelihood that a whole will undergo further analysis and not deposit as a wholelike part, and the occurrence of those parts that are

possible given the whole as a starting point, could be captured by a set of probabilities, but to do this would require that ***given the whole, one knew or could generate all of its possible outcomes in advance.*** The potential of the whole is seized only when its possible outcomes are known or imagined. This is the basis for the postulation of propensities which are closer in some sense to outcomes. Still, since probability describes the distribution of a change and not the process through which the change occurs, it does not provide a satisfactory account of the meaning of potential.

Preformation entails that a future part can be uncovered in the static content (precursor, code, etc.) of a present whole. However, since every part can be a whole for a further stage of analysis, preformation would require not only that every part in a whole to part series is represented in the initial state, but that the initial state prefigures every unrealized or, in some versions, counterfactual, derivation.

The potential–actual, whole–part or many–one relation recalls the type–token relation in functionalist philosophy. If a single "type" is common to a manifold of "tokens" of that type, each token must have its correlates in the type-initial state. If the token maps to a constituent of the type, is the type recaptured by a portion of its content? For a type to be reconstructed in part from its tokens requires a multiplicity of copies in the type-initial state. A potential consists of actual and unrealized tokens. A whole–part or type–token correspondence, even if partial, begs the question of underived parts and the factors that account for parts that do or do not become actual. For these reasons, a correspondence theory is vulnerable to the objection that types (wholes) are not just incomplete specifications but are qualitatively distinct from the tokens (parts) they anticipate.

A state of potential connotes a power or a multiplicity that delimits to a future actual state, whether *over* time, e.g., the potential of a seed to become a flower, or momentarily in becoming, e.g., the potential of an idea to call up an action. The present state instantiates the potential for a past state or it expresses the potential for a future one. The judgment of potential, therefore, whether realized in a product or residual in a whole, depends on whether a state is deemed initial or final. However, a judgment as to what is final does not rest on the perceived actuality of the state. A seed is an object with a potential for growth. Lacking this potential, the seed is a terminal object. The assignment of potential to an object is linked to its capacity for change, i.e., recurrence. In a seed, recurrence is identified with growth. Recurrence can also lead to decay, which signals a loss of potential. What is meant by potential in this case? A seed has a potential for growth or decay depending on the change it undergoes.

In microgenetic theory, potential is the possibility of change at every moment in a series of whole to part transitions. Wholes and parts are arbitrary designations. So also is the distinction of potential and actual. The progressive delimitation of possibility—the cutting-off of potential as successive states are realized—is an artifact of our perception of what is terminal. In mind, potential implies the capacity for a manifold of actual (objective) events, e.g., a concept that discharges into objects or words. We tend to think that potential is discharged in an actual object but an actual object is what we perceive. A different perceptual system might carry the whole–part analysis more deeply into the object through a series that is essentially bottomless, e.g., to its atomic structure.

Every "slice" of becoming retains a potential to become the next "slice" in a continuum, so that potential is not in the content of a given slice but in a "law" of transformation. This "law," which combines the pattern and rate of context to item shifts, defines the process in a becoming sequence. The configural properties at a given moment determine the range of possible outcomes. The momentum of the transformation and the configuration that is being transformed constitute the potential at a given state.

On this view, the "capacity" residual in a particular content is implicit in the transition the content is undergoing. If this transition is abbreviated and the content becomes actual, the potential in that content, i.e., the transformational dynamic, is lost for another fractionation. The ensuing series replaces the content and determines in the next deposition how the dynamic terminates, e.g., into growth or decay. Potential is the term we use to express this tendency which, apparent or not, inhabits every content, while *content*, in an object or a mental representation, is the term we use for the momentary terminus of a given transformation. This implies that potential is ubiquitous and that observed states of potential are only moments of apparent wholeness. What are parts and what are wholes depends on our perceptual apparatus, with every state of wholeness the configural moment and pattern of transformation from one segment of change to the next.

ORDER

Order applies to the structure of a given event in the harmony of its parts; and to the temporal pattern of an event sequence. Is there a common framework in which these two (spatial, temporal) meanings of order can be grasped? With respect to the temporal pattern, there is order within and between appearances of the same or different objects. There is a hidden order within a becoming and an apparent order from one

actual object to the next. The apparent order, or succession of actual objects, is the surface progression of nature. The pattern of reconstruction of an object concealed in the imagination is assigned to the linkage *between* actual objects.

In the deeper sense of passage, temporal order is the precedence of phases *within* a set of qualitative layers of becoming only the surface of which is perceptible. The transition from initial to final stage is one indivisible object. Order implies events in succession while becoming extends over an indivisible duration. Order is a product of a becoming, not a series into which a becoming deposits. The actualization of whole to part is a bridge to the idea of order as harmony in a spatial pattern.

What, then, does it mean for there to be order in a system? We can say that order is maintained through the expenditure of energy but what is the order that is maintained in this way? Order and disorder are difficult to specify.[17] An object changes so its elements coincide. Entropy involves change in the equilibrium of the component particles of a system. Increasing entropy is possible only in irreversible systems. Life is such a system. Novelty is also a feature of irreversible systems. If direction in time is mind-dependent, even if entropy is unrelated to time, the subjectivity of disorder is the same subjectivity from which irreversibility is derived.

UNITY

> . . . where begins
> The union, the partition where, that makes
> Kind and degree, among all visible Beings.
> WILLIAM WORDSWORTH, Excursion IV, 334-337

The object-to-be generates phases in its growth that exhibit novelty as a property of their actualization. Unity obtains in the coherence and thematic of the process, novelty in the shift from one unity to another. The thematic is the direction, the coherence is the connectedness, and the unity is a web of spatio-temporal dependencies. Each phase delineates the manifold to a further definition. This is how objects grow into the world.

Unity is a way of characterizing relatedness. Since the relatedness *within* an object is not directly accessible to perception, the feeling or intuition or perception of unity cannot reside solely in the relatedness. This is not something that can be stated with assurance. What is the criterion of unity? A crystal has a unity that is not shared by an aggregate of disparate elements. Is this unity a function of the magnitude of the perspective or a distinction grounded in intrinsic relations of matter.

Unity must require the harmonious and coherent participation of elements. If elements are not isolated, i.e., if they are never truly elemental, their unity must owe to the relations they derive from the whole. The uniformity or redundancy of crystal structure is a crude form of unity where the same elements repeat a common pattern. A different type of unity results when a pattern is shared by diverse elements. In a chess game, the unity is in the (synchronic) pattern and logic of the pieces but even more in a (diachronic) array poised between an immediate past and an oncoming future. The pattern resolves a prior state of potential and relates the resolution to a future actualization.

The basis of unity is that elements express or exhibit properties of the whole where a transition from whole to part is central. Unity is in the potential for an ensuing actualization. The unity of an assemblage in space derives from its coherent passage in time. Unity is not the distribution of elements in a simultaneous slice of nature but a transition from whole (past) to part (present). This is how unity is linked to order in the sense of harmony. In a unified system, elements replicate patterns of progressively larger domains until a single pattern resonates throughout the system, every element partaking of the whole. This partaking is the basis for an aesthetic of harmony.

NOTES

1. K. Popper and J. Eccles, *The Self and Its Brain* (New York: Springer, 1981). Popper argues that objective probabilities underlie the possibility of emergent entities. Probability, however, only makes room for, i.e., entitles or logically enables, emergence; it does not explain it.
2. A. N. Whitehead, *An Enquiry Concerning the Principles of Natural Knowledge* (Cambridge, UK: Cambridge University Press, 1919), 12.
3. C. von Ehrenfels, Uber Gestaltqualitaten. *Vierteljabrschrift für wissenschaftliche Philosophie* 14:(1890). There are prior references, R. Frondizi: *The Nature of the Self*, (New Haven, CT: Yale University Press, 1953). Lao-tse in 600 B.C. may have been the first to refer to whole–part relations in his *Tao-te-ching*: "the sum of the parts is not the whole."
4. For example, the relation in size of the visible universe to the earth, the earth to an atomic nucleus, and a nucleus to a superstring is roughly of the same magnitude. See F. Dyson, *Infinite in All Directions*, Gifford Lectures (New York: Harper & Row, 1985); D. Bohm, *Causality and Chance in Modern Physics* (London: Routledge and Kegan Paul, 1957).
5. As Scaltsas put it, elements are consumed by their unification, the relation between elements "destroys the relata, leaving nothing standing but the whole," T. Scaltsas, "Is a Whole Identical to Its Parts?" *Mind* 99(1990):583–598. From a different perspective, Heller writes, "We do not start with parts and somehow build extended wholes out of them; we start with filled regions of spacetime and then look to the relationships between those regions to discover the part-whole relationships between objects." M.

Heller, "Things Change," *Philosophy and Phenomenological Research* 52(1992): 695–704.

6. Popper and Eccles (*The Self and Its Brain*) note that averaging, e.g., the average velocity of molecules to give temperature, adds something not present in the elements.
7. On whole–part relations in Gestalt psychology, see B. Smith, ed., *Foundations of Gestalt Theory* (Munchen/Wien: Philosophia Verlag, 1988); D. Sweet, "The *Gestalt* Controversy: The Development of Objects of Higher Order in Meinong's Ontology," *Philosophy and Phenomenological Research* 53(1993) 553–575. For some theorists, e.g., Köhler, emergentism was not obligatory.
8. For the history of the concept, see C. L. Morgan, *Emergent Evolution* (New York: Holt, 1931); B. McLaughlin, "The Rise and Fall of British Emergentism," in *Emergence or Reduction: Essays on the Prospects of Nonreductive Physicalism*, ed. A. Beckermann, H. Flohr, and J. Kim (Berlin: Walter de Gruyter, 1992).
9. D. Baxter, "Many–One Identity," *Philosophical Papers* 17(1988):193–216.
10. For Lewis, composition is identity; D. Lewis, *Parts of Classes* (Oxford: Basil Blackwell, 1991).
11. T. Horgan, "From Supervenience to Superdupervenience: Meeting the Demands of a Material World," *Mind* 102(1993):555–86. Also, papers in A. Beckermann, H. Flohr, and J. Kim, *Emergence or Reduction*.
12. S. Alexander, *Space, Time, and Deity* (London: Macmillan, 1920).
13. The identification of properties rather than the events on which the properties depend would seem to be a ploy to satisfy the *logical* demands of identity theory and avoid the deeper problem of subjectivity or the actual *nature* of the identity. For an insightful discussion of these issues, see, J. Kim, "On the Psycho-Physical Identity Theory," *American Philosophical Quarterly* 3(1966):227–35.
14. M. Bunge and R. Ardila, *Philosophy of Psychology* ((New York: Springer-Verlag, 1987).
15. J. Van Cleve, "Mind-Dust or Magic," *Philosophical Perspectives, 4, Action Theory and Philosophy of Mind* (Atascadero, CA: Ridgeview, 1990), 215–226.
16. C. Sherrington, *The Integrative Action of the Nervous System* (1906; reprint New Haven, CT: Yale University Press, 1947).
17. Penrose comments on the subjective element in the judgment of "manifest" (macroscopic) disorder. R. Penrose, *The Emperor's New Mind* (Oxford: Oxford University Press, 1989).

CHAPTER 13

Process and Creation

ARGUMENT: *Creativity is novelty in the domain of concepts. The creative idea is a cognitive whole that partitions into its varied expressions as concepts empty into acts, images, and objects. The errors of the brain-damaged and psychotic cases exhibit the same part–whole relations that characterize metaphor and new concept formation. The whole to part transition in cognition is a species of creative advance in physical passage.*

> The Gifted have told us for years that they want to be loved
> For what they are, that they, in whatever fullness is theirs,
> Are perishable in twilight, just like us.
>
> MARK STRAND

From the different perspectives that have been brought to bear on the study of the creative personality over the past century on the effects of cultural and political conditions, family structure, early development, and so on,[1] it is clear that a complex of events must converge at successive stages in the life of an individual to nurture creative ability. In addition to this complex of external conditions, a set of innate dispositions is probably no less essential. The balance of these factors determines not only whether an individual is creative but the scope, the drive, the intensity, the confidence to follow through, and the discipline that a productive creativity demands.

In discussions of creativity, the life and times of the creative person have been given so much attention that they tend to displace an account of the creative ***process*** from the mind of the individual where it belongs to circumstances of biographical detail. Collectively, these events impact on the development of the creative *personality* but separately they are

incidental to the process of creative *thinking*. An explanation centered in the facts of an individual life leaves the creative process itself unexplained.[2]

In recent years, attempts have been made to demystify the act of creative thinking, especially inspiration, as a phenomenon of almost magical signification. Experimental studies have dispelled the mystique of creativity as an irrational mode of cognition, linking it instead to normal problem solving.[3] The incremental nature of creative thinking and its continuity of expression over time suggest that it is related to the normal thought process.

Creativity is certainly present in all people to varying degrees. The sudden insight to the solution of a game or a puzzle, the Aha experience on apprehending the answer to a challenging problem, are not equivalent to the discovery of relativity theory, but such experiences presumably reflect a continuum of performance in a common mental process. Have we not all had a fragment of the experience of a Mozart on leaving a concert and "hearing" the music all at once in our mind? For Mozart, a musical concept was generated endogenously, while in listening to Mozart a memory image of the music is generated by a template. As incomprehensible as the genius of a Mozart may be, it is probably an elaboration of certain aspects or domains of normal thought. But what exactly is normal thinking? Is the idea of a common mechanism for the normal and the exceptional, both the gifted and the abnormal, consistent with *degrees* of originality? More deeply, what is the nature of creative thinking and how does creativity in mind relate to novelty or creativity in physical brain process?

NOVELTY

The distinction between novelty and creativity is often cast as a difference between the physical and the mental, novelty being the appearance of a new entity, creativity a mental act in which the novel is created. We tend to attribute the possibility of novelty to purely material events, though in everyday discourse novelty is applied to human or animal behavior and can refer to both physical and mental events.

On the other hand, creativity requires a mind to bring something novel into existence the novelty of which is judged by others to reflect the talent, originality, giftedness, or genius in the creative act. Creativity is not ordinarily assigned to the material world, i.e., the world "machine," except for accounts of the "creation" of the universe, in which case the attribution of creativity follows on the assumption of a mind, i.e., a creator, that is responsible for the creation. For example, in the

metaphysics of Whitehead, change in the world is conceived as a creative advance into novelty. Whitehead wrote, "the creativity of the world is the throbbing emotion of the past hurling itself into a new transcendent fact"[4]

In the relation to material events, novelty could be a property of the brain state which is a physical process, while in the relation to mental events, creativity could be a property of the mental state. The distinction entails an implicit dualism, i.e., novelty in the brain state, creativity in the mental state, unless creativity is understood as a "higher" order or evolved form of novelty, in which case the ***brain*** process that generates a creative idea is no less creative than the mental process to which it corresponds. In this way of thinking, novelty could occur in the absence of creativity, with creativity an expansion of novelty at a certain level of organization, i.e., the creative is a more complex expression of the novel.

A judgment of novelty requires a comparison of a prior and an occurrent state. There must be a departure from the expectations of the prior state for the judgment to be made. Apart from the mind-dependence of this judgment, we can ask what is the nature of the departure. If the step leading to novelty reflects a set of contingent events that impacts on change, the shift from a causal effect to a novel outcome could in principle be specified, say through a post hoc analysis of the effects of the (known) contingencies.

In contrast, the creative seems inexplicable from its prior states. There is a causal "gap" between the antecedents of the creative step and the moment of creation. Partly, this feeling of a gap reflects the definition of creativity as a coming into existence. The feeling is heightened by the basis of creative work in the mind, in contrast to the "causal certainty" of physical (brain) mechanism. The creative outcome is further obscured by its presumptive origins in the subconscious with a realization in consciousness, a step more like a quantal jump than a continuous sequence.

We tend to think of physical change as causal, with novelty improbable or rare, like a miracle in a deterministic world. A deviation from the expected is impossible in a world of universal causation. Whether or not there is novelty depends on a theory of change. A novel change should be unpredictable but the inability to predict a change, even if all information is available, does not obligate that the change is novel. Random or unpredictable change is not necessarily an occasion of novelty; a random process could generate a (nonnovel) recurrence. An approach on the basis of probabilities has the same drawback. Neither approach gets at the inner ***nature*** of change. The indeterminism of microphysics is compatible with novelty even if it does not elucidate

the ***process*** through which novel states appear. John Dewey wrote that only a philosophy of "genuine indeterminism, and of change which is real and intrinsic gives significance to individuality. It alone justifies struggle in creative activity and gives opportunity for the emergence of the genuinely new."[5]

CREATIVITY

Creativity is not an accumulation of elements into more complex aggregates. Suppose I imagine or draw the head of a frog on the body of an elephant, and suppose also that no-one has drawn or even imaged this combination before, this would not qualify the image or drawing as necessarily creative. The parts are unchanged by their recombination. They are the same parts differently put together. One might as well have a rockpile that is rearranged by an earthquake.

One can say that creativity is novelty applied to concepts; or that in creative thinking there is a conceptual basis for the novel. Essentially, creative activity is the forming of new concepts in the mind of an individual. In concept formation the creative is achieved as a cognitive whole not decomposible to a set of antecedent or constituent elements. Indeed, there are no basic elements, only the emergence of the whole from its ingredients, i.e. the replacement of prior elements by the novel whole, and the potential of the whole to fractionate into novel parts.

While the understanding of the whole may, it often does, develop on a profound knowledge of the parts, the parts that were preparatory in the elaboration of the whole, e.g., elements of prior theory, are not the same parts after the whole has been established. The parts that go into the creative transformation, e.g., data, tradition, technique, differ from the parts that emerge from the new concept. After the new concept is grasped, the parts are conceived differently and thus are different parts. The parts are reconfigured or transformed by the new mode of understanding.

In a discussion of scientific creativity, David Bohm has written that, "in most cases it is not (an) experiment . . . that falsifies earlier theorys and conceptions; rather it is some new understanding which arises in response to reflection on the *total* situation,"[6] To reflect on a total situation or to achieve a new understanding is to grasp an original perspective, and this perspective is a conceptual whole with the potential to develop into the parts that it anticipates.

The intuition, grasp, or total understanding surfaces to awareness as its implicit "content" becomes more factlike, thus more explicit. The moment when the conception becomes explicit is the inspiration or total

understanding. Still, even at this stage, the explicitness is not that of actual fact. The creative *product* is generated through a resurfacing of fragments of the original concept.

Inspiration, then, is the presentation in the mind of a conception too replete for expression that enfolds the entirety of the work. This conception is replaced by part-concepts that, through repeated presentations, empty into the art, the theory, or the science. In the translation of a concept "in the head" to a canvas or a page, anticipatory concepts develop to actual facts through the iterated analysis of conceptual wholes.

A more precise formulation, then, is that creativity is the articulation of transformed parts out of novel wholes, where the novelty in the parts owes to the constraint of the whole that configures them, and the novelty of the whole is the originality and degree of empowerment given to the parts that emerge.[7] Emergence is the key; it is a general principle of physical and mental process that every effect is in some degree, however slight, an emergent whole,[8] and that every whole is an emergent part derived from a larger, antecedent whole. The creative *process* can be studied from these several aspects: concept formation as the basis of creative thinking; this process as an instance of the whole–part relation; and the whole–part relation as the basis of creativity in the mind and creative advance in the world.

Pathology of Concepts

Disorders of thinking are important to study because pathology exposes in the symptoms of abnormal thinnking the very mechanisms that are involved covertly in the normal thought process. The neurologist most closely associated with the topic of thought and its disorders was Kurt Goldstein,[9] though the conditions he studied the most thoroughly, the disturbances of language and perception, he considered to be disorders of the *instrumentalities* of thought, not disorders of thought itself. Goldstein described an impairment of the abstract attitude in certain aphasic, amnestic, frontal lobe and other cases which he believed to reflect the disruption of a thought process prior to its implementation in speech, action, or perception. For Goldstein, the abstract attitude, i.e., conceptual or categorical thinking, was basic for the establishment of a voluntary mental set, for shifting from one set to another, for grasping a whole, breaking it into its parts, and recombining them, and for holding multiple aspects of a situation in mind simultaneously.

The abstract–concrete dichotomy deserves closer attention for it is linked to the whole–part problem and is a central aspect of thought

and its disorders. In concrete behavior, the category cannot be accessed from the instance, e.g., the color *red* is not abstracted independent of a red apple. If the patient can accomplish this task, he may not arrive at two or more categories from an instance, e.g., that an apple is a member of the category of shape (round), of color, and of food. The impairment of abstraction can be relatively selective, as in defects of color naming or sorting, e.g., naming colored objects or grouping them in color categories, or it can be generalized and affect a great many perceptual tasks, as in frontal lobe patients.

In the latter cases, the deficiency is the basis for impairments on tests such as the Wisconsin Card Sorting Test, in which the patient is required to sort objects along several different dimensions. Given an instance of a category, e.g., shown a round red object and asked to group it with similar objects, the patient cannot derive the target category from the member items. Nor can the patient sort according to several dimensions (color, shape, etc.). Such patients may verbalize the correct strategy for the task but not implement it in action. Luria[10] referred to this as a disorder of verbal regulation, i.e., a dissociation between thought and action.

The opposite behavior is also commonly observed. For example, on a naming or reading task an aphasic or dyslexic patient may sample a word or object category, e.g., saying grapefruit instead of apple, or reading zebra instead of horse, relying on the lexical–semantic and, less often, perceptual features of the object, or giving responses such as eat or tree for apple, where the situation or the nexus to experience is prominent. One might speculate that the correct category is realized, e.g., fruit, but the instance, e.g., apple, is not derived out of the category. In these cases, the difficulty is in going from category to exemplar, not the reverse. Indeed, there are cases where the patient is able to sort written words and objects according to their category without being able to recognize, not just name, the categorized item.[11]

The inability to go from the category to the item, i.e., having the category (fruit) but not the item (apple), suggests a priority of wholes in the progression from whole to part. However, those cases with an inability to derive the category from the item do not imply the opposite process, i.e., going from item to category, or that the patient recognizes the items but not the concepts or categories that stand behind them. Every item (word, object, act) develops out of a category or concept. Sorting tests involve object and category identification. A patient asked to group various objects into their categories must access multiple phases in the same object, i.e., recognize both the category and the item. There are cases (agnosic) who "lose" the concept of the object and are unable to identify it. The object name cannot be "found" unless

the concept or category to which the object belongs has already been traversed.

Brain-damaged patients without aphasia have difficulty with polysematous words. For example, given a word such as ***bank***, and asked to point to words such as money, river, etc., they will often select only one meaning. Other cases, asked the color of an orange, may say "yellow" or "red." A similar phenomenon occurs with verbal nouns, e.g., what do you shovel snow with? Such patients are unable to deal with more than one meaning or interpretation at a time. There is an inability to revive the alternative concept, perhaps due to blocking or persistence of the initial interpretation out of which the polysematous item develops. The fact that patients with severe aphasia (or schizophrenia) show enhanced semantic priming[12] even for words they do not recognize or misidentify, or that one meaning of a polysematous word may prime the other, or that patients with cortical visual defects can extract the meaning of words they do not consciously perceive,[13] argues strongly for early or preliminary access to the semantic or object category—even to many related semantic categories, i.e., shades of meaning—with subsequent derivation of the specific exemplar or member item. This, I would argue, is another instance of the transition from whole to part.

Most patients recognize the instance as a feature of the category, but there are cases, such as those described by Weinstein and Kahn[14] where disparate objects are identified on the basis of shared or overlapping features, e.g., a patient who names a doctor as a butcher based on the white jacket or perhaps the comparable level of skill. This is not just facetiousness; the patient believes in the identity of the two objects.

A similar behavior in schizophrenics (the von Domarus effect)[15] is the identification of disparate topics on the basis of shared attributes. For example, an apple is thought to be poisonous due to the shared property of, say, being able to be ingested or on the basis of literary mediations. The motivation for the identification, e.g. paranoid ideation, is a sign of the depth at which the error arises. The deeper the origin of the error, as in dream or schizophrenic thought, the more it samples personal memory and subconscious cognition.[16] This is the irrationality of psychotic thought.

In hallucination or dream, shared perceptual or semantic features provide the basis for the substitution of objects, e.g., a knife may be a symbol for a penis based on the shared attribute of shape, penetration, etc. A similar mechanism is at play in so-called schizophrenic paralogic, e.g., Mary is a virgin/I am a virgin/I am the virgin Mary. A similarity by way of a common feature or predicate suffices for an identity of otherwise different objects or subjects.

These examples, I believe, are all variations on the same theme, which can take inter alia the following forms:

1. Failure to go from one or more instances (items, attributes, predicates, features) to their context or background category, and the reverse, failure to recognize the item given the category.
2. Inability to stabilize or apprehend the *partial* relatedness of an item given its category, so that different categories (objects, subjects) are identified on the basis of common attributes
3. Failure to apprehend an item as belonging to multiple categories or to apprehend an object or category as having multiple features or meanings.
4. Inability or lability of shift from one context–item pair to another; inability to suppress one context or meaning in favor of another.
5. Ability to access the category but not identify the item
6. Ability to identify the feature but not the object, or the object but not the category.

The different manifestations of context–item or category–member relations are complex, but a thorough analysis of this topic should clarify the relation of pathological to normal thinking. What is item or member and what is context or category accounts for the difference in symptomatology across patients and from the neurological to the psychiatric series. In psychiatric disorders, shared features tend to be the basis for an identification or substitution of categories, whereas in neurological cases, a single category tends to be sampled for items having features in common with the target object. The continuum from the psychiatric to the neurological reflects the depth of cognition sampled—thus the affective intensity and/or delusional quality of the error—and the content or modality, i.e., the degree of generality or specificity to language, perception, or action, which is also a function of the depth of the error.

Every item is an element for a larger domain of study or a ground for a further analysis. The transformation of context to item or ground to figure is bottomless. Moreover, the context–item or concept–feature relation, as Goldstein noted, is a relation between wholes and parts. The part–whole relation is also central to schizophrenic paralogic. Arieti noted that, "the more difficult it is to abstract a part from wholes, the stronger is the tendency to identify the wholes which have that part in common". Moreover, the identification of topics in the von Domarus effect is not due to "shifting cathexes", as in psychoanalytic theory, but is based on the "cognitive equivalence of members of a primary class".[17]

I believe the whole–part transition is in a direction from whole to part.[18] The reverse direction, the construction or emergence of wholes

from parts, is not the opposite of that from whole *to* part. The part to whole transition is the replacement of the mental state that leads to a given part by an antecedent whole that is reconfigured in the next mental state. A continuous whole to part shift is the process-equivalent of the laying down of the mental state. The opposite direction, the shift from part to whole, requires this process to be reversible. Since the process is linked to the asymmetry of time awareness, a shift from part to whole might entail a reversibility of subjective time.

Pathology unveils normal process. That is the importance of the pathological material. The varied relations of parts and wholes in the pathology of lexical and object concepts provides a basis for thinking about normal cognition[19] and creative ideation. The possibility that the part–whole relation is fundamental was first thoroughly explored by the Gestalt psychologists. Thus, Wertheimer argued that thinking concerns "the relations between parts and wholes . . . involving operations as to the place, role, function of a part in its whole." Among these relations are the division of wholes into parts (subwholes), seeing the parts together without losing sight of the whole, and the achievement of closure in a "good gestalt".[20]

Metaphor and Whole–Part Relations

Simile and metaphor are ubiquitous phenomena the importance of which to psychology can be appreciated by just one example, the metaphor of the brain as or *like* a machine (computer, hologram) that has been investigated as literal fact.[21] Metaphor develops out of *perceptual* part–whole relations fundamental to the human conceptual system.[22] These relations are primary. One could discuss creativity in terms of lexical concepts and still not tap the pre-lexical sources of the creative imagination.[23] Metaphor is one way that language extends the figure–ground or feature–gestalt relations of spatial cognition.

In metaphor, a topic is assigned to a category in which the metaphoric vehicle is an instance. Take the example: *My doctor is a butcher.*[24] Attributes of the metaphoric vehicle *butcher* serve as "connecting links" to the topic *doctor*, i.e., the two categories are related by virtue of shared attributes. In metaphor, the relation can be the basis of a creative use of language. When an aphasic calls a doctor a butcher, the relation is not metaphoric since the categories are identified. Whatever is metaphoric in the expression is inferred by the listener. The speaker is not attempting to convey an unusual meaning and is unaware of the error.

At one level, a metaphor is a comparison. In other comparisons, e.g., "a grapefruit is larger (sweeter, etc.) than an apple," or in simile, "the sun is like an oven," two similar or dissimilar items overlap or are related by an explicit attribute (size, heat, etc.). The attribute is not just a nexus uniting the categories to which the terms refer, or uniting terms in the same category, but can also serve as an ad hoc category, as when similar or dissimilar items are related or listed according to size, sweetness, etc.

When an aphasic names an apple a "grapefruit," he is sampling the category (context) of these items, i.e., fruit. The misnaming is an implicit comparison, since it relies on common (in-class) features for the identification. The background category of fruit is a whole that has the potential to actualize to parts. One could say the parts, e.g., grapefruit, apple, share features in the category that are the basis for the misnaming. Conversely, there is incomplete elicitation of the lexical items from the antecedent whole of the category.

When a schizophrenic calls himself "football," as did one of my patients, items from disparate categories are misidentified on the basis of common features, e.g., being "kicked around." In aphasia and schizophrenia, regardless of whether errors arise on features within or across categories, but particularly when experiential or functional attributes are involved, patients are not fully aware an error has occurred; i.e., the patient does not apprehend the relation implicit or collapsed in the error. In normal thought or language, in comparison, simile or metaphor, the individual is conscious of these relations as a device for the communication of new meaning. The inability to bring the whole–part relation to awareness is due, partly, to the inability to retain two items simultaneous with their relatedness. Certainly, the exposure and vulnerability of these relations reflects the difficulty in comprehending wholes and parts as distinct but related contents.

The apprehension of new meanings and the awareness of metaphoric (or other part–whole) relations are important differences between the pathological and the creative. Psychotic speech and certain types of aphasic jargon may resemble some forms of poetry. For example, a patient of mine with semantic jargon described his difficulty with vision as "My wires don't hire right." Another aphasic wrote that she: "found Brooklyn about her troubles, a small nature in the pink yellow garbage from motion." The schizophasic of Arieti with word-salad said, "The cow burnt the house horrend(end)ously alway." Such errors can be analyzed in terms of contextual effects on word substitution.[25]

The similarities between semantic deviance in aphasia and schizophrenia with some forms of poetry can be striking; e.g., Dylan Thomas wrote "If my head hurts a hair's foot/ Pack back the downed bone." A

more subtle anomaly of word meaning is found in verse by the schizophrenic poet Ezra Pound: "Shines in the mind of heaven/ God who made it/ More than the sun in our eye." In this passage, the object referred to is not given in the neighboring text.[26] Derailment in word meaning and the sampling of the contextual background of target lexical items can be turned to artistic effect. We assume that such writing is deliberate in the poet and involuntary in the psychotic or brain-damaged patient. However, the poet may be no less passive than the schizophrenic to the depth of origin of his or her verbal imagery but differs from the schizophrenic in being able to edit the material later on. Conversely, the aphasic or psychotic has little or no awareness of the defectiveness of the utterance and cannot play with creative output. The ability to edit is essential. Dali alludes to this in his witty comment that he was critical of his own paranoia.

Metaphor and Concept Formation

Ordinary conversation is often a description or recounting of one's states, opinions, or experiences where the speaker is the topic; e.g., "I (remember, anticipate, enjoyed, etc.) drinking Pastis with Pierre." When the speaker detaches from the topic, and facts or events are not simply recounted, novel entities may occur.[27] For example, a relation of simile could introduce a new meaning to extrapersonal terms; e.g., "Pierre is (like) a saint." Here, Pierre shares features or attributes of saintliness, or an attribute that is a part of the whole concept of Pierre is an attribute of the whole concept of a saint. The common part brings two wholes together. In synechdoche, e.g., "The saint has departed," an entity (Pierre) is replaced by one of its features or parts, saintliness. The correlate of this in pathological states might be the identification of an object by a perceptual or functional feature, naming an apple "red" or "eat." Similarly, the replacement of one word for another in metonymy is a common error in aphasia.

An additional step takes the expression further, e.g., "Pierre has piety without religion." This step complicates the part–whole relation in an interesting way. Now the feature *piety* that links Pierre to a saint is cleaved from another feature of saintliness, that of religiosity. The result is an expression close to irony. When the cleavage is made explicit, e.g., Pierre is no saint," the feature, say piety or celibacy, becomes the implicit topic in a remark that turns critical. Such examples show how the play of concept–item or category–attribute relations can become very complex with ad hoc concepts or categories continuously being formed.[28] Of interest are studies of aphasics showing ad hoc categories on sorting

tasks when abstract categories are unavailable; for example, sorting tiger with crocodile instead of cat based on their common ferocity.

In the example, "Pierre has piety without religion," a microgenetic account might hold that the categories of Pierre and religion parse to a featural element, *piety*, which is affirmed in the proximate section of the sentence and negated in the distal section. The fractionation of categories or wholes to their exemplars, e.g., Pierre, religion, supplies the main terms, while the shared features take on opposing contrasts. What is a whole and what is a part is arbitrary. Features can become concepts and vice versa. The *feature* piety can become the *concept* "piety" or the *category* "pious things" to which other wholes such as Pierre or religion, can then relate as positive or negative features.

In bisociation,[29] a new concept arises, perhaps as an act of inspiration, in the fusion of previously unassociated concepts. While conceptual growth is probably more often the result of a gradual transformation than a sudden insight, the account of bisociation resembles that of the blending of items or categories in simile and metaphor. Whether the result is a prelexical concept, as in creative thought and imagery, or a more restricted phenomenon such as a new metaphor, depends on which categories, concepts, or words serve as topics or vehicles, and which feature or set of features is the axis of the transposition. The profoundly creative involves concepts of breadth and/or universality in relation to the elements those concepts enclose. Schopenhauer wrote that the fundamental characteristic of genius is "always to see the universal in the particular."[30] This aspect of genius, along with the depth of creative insight (see below), is embodied in the notion that genius explores underlying concepts (intuits) while talent works with surface elements (analyses).[31]

Categories and Concepts

Part–whole relations figure in the acquisition of basic level categories, e.g., dog, car. Such categories depend on gestalt mechanisms of perceptual similarity, especially shape, while superordinate categories such as animals or vehicles tend to share functional features. Perceptual wholes and features are more salient at the basic level—the shape of a car, the wheels, engine, etc. Moreover, in such objects, "the wholes seem to be psychologically more basic than the parts."[32] Basic objects are the first to be learned in childhood. The whole to part transition is characteristic of early cognitive development and a part–whole relation is critical in the learning of basic objects. Part–whole relations are easier than class

inclusions, and there is a shift in childhood learning from a reliance on categorical meaning to a reliance on features.[33]

For Tversky and Hemenway, the decomposition of wholes into parts is the basis on which structure is used to "link the world of appearance to the realm of action," and to comprehend, infer, and predict function.[34] This relation characterizes intuition and naive induction and is the basis of novel concepts in scientific and other forms of creative thought. These authors give as an example the shift from holistic concepts of brain function to the concept of functional localization. For Lakoff, basic level objects and their parts engage action and the body schema to generate metaphor and complex categorizations.[35]

A category tends to be the more primordial entity, a concept is a more specific or individuated category. A category is a group of like things that resemble each other along some dimension, e.g., the shared features of *dogs* or *chairs*. A concept incorporates the dimension along which such resemblances are established. We say, the *concept* of a chair, i.e., the shape, the features and functions that determine what a chair is, and the *category* of chairs, i.e. the grouping of chairlike objects, or objects that satisfy the concept *chair*. The category of dogs and chairs includes by implication the perceptual features of those objects.

Animals have primitive object concepts for shapes and features. A dog can recognize the categories of dogs and chairs, prey, shelter, without a concept of the meaning of these objects other than the responses they call forth. An object concept is a whole–part relation that, when generalized over similar objects, is the category of that object type. The configural aspect of these objects and their part–whole relations determine the object concept of dog or chair. A lexical concept, e.g., the word dog, is also a whole–part relation. The word has the potential for different meanings and denotations. The shift from the potential of the word to a specific instantiation is a species of the whole to part transition.

Depth and Surface

Though itself a metaphor, depth of process is important.[36] The depth of creative thought introduces a *microtemporal*, cladistic[37] or genetic dimension to category formation. The relatedness between concepts reflects their immediate prehistory, not actualities at the surface which are mere outcomes. Concepts arise in the subconscious of long-term memory organized around experiential and affective cores and traverse the dreamwork, images, symbolic and metaphoric relations, and the like, on the way to propositions and the rational or logical structures they instantiate as "facts" in the mind or the world.

Jung wrote of a "visionary mode" with its own subconscious autonomous form. In this mode, the individual is passive to the emerging creative product. The passivity is a clue to the depth of origin of the creative idea. As in hallucination and dream, the self is passive to its own emerging content. Fantasy and reverie are associated with a receptive attitude. Goethe said, "Thinking doesn't help thought"; and Beethoven, "You ask me where I get my ideas. That I cannot tell you with certainty; they come unsummoned."

I would align myself with Kris,[38] that creativity is a flight from deliberation in the service of novel concepts. The withdrawal from objects to their anticipatory constructs in spatial and imaginal thought allows a more generic concept, i.e., one with the potential to develop into different modalities, to be realized in a specific cognitive domain. This ability to *dip* into the "pool of the creative unconscious" is an uncommon experience for the average person, those who, as Wordsworth wrote, have a "mind intoxicate with present objects and the busy dance of things that pass away." The creative personality reclaims the conceptual and symbolic sources of those objects. The depth of creative work implies an engagement of fundamental aspects of the personality, whether in the sciences or in the arts. When Piaget pointed out to Einstein that his concept of spacetime resembled time perception in small children, i.e., young but not older children perceive time in terms of spatial relations, Einstein is reported to have wondered whether this might have been the result of his slow maturation.

Creation and Nature

We gain a better understanding of mental process by a study of its pathology and correlated brain mechanisms. In so doing, we move closer to a depiction of the physical, to which the mental is our only contact. If whole–part transitions are the basis of human thought and creativity, and if mind is part of the physical world, the relations that characterize mental process would be the same relations that hold for creative advance in nature.

The world of thought and perception emerges through a graded analysis of wholes into parts. In this process, an object is the outcome of change in a passage to greater definiteness. Every entity in the world is a momentary novelty. The world is never twice the same. Creativity depends on the potential of this activity to regenerate the world and the incompleteness of actualizations through which novel worlds are generated.

Science is the study of the relations of physical succession. Microgenesis is a theory of emergent recurrence. The whole–part or many–one relation, the elicitation of items out of contexts, is a fundamental property of mental process. Is the whole–part relation *the* underlying principle of change in mind and nature? Whitehead thought so, and wrote that creativity is "the ultimate principle by which the many, which are the universe disjunctively, become the one actual occasion, which is the universe conjunctively".[39]

Mind gives duration to the transient events it records, like the tidal waters of a rushing stream, a moment of persistence in the flux of actual events. The link from duration to creativity was the theme of Bergson's great work. He wrote, "the more we study the nature of time, the more we shall comprehend that duration means invention, the creation of forms, the continual elaboration of the absolutely new."[40] Duration is the basis of categorization (p. 19) and basic categories are the nuclei of primitive concepts that give rise to objects. An object is first a concept in memory before it is an object in the world. The recognition of the object is by way of the concept of that object that summons the object up. Memory does not hold on to the world but creates it, and it is memory, or the process that makes memory possible, that sustains the world over its momentary instantiations. In this way, through the continual formation of new concepts, mind gives meaning and stability to the raw succession of physical states. This is the expression of creativity in physical process. After all is said, the creative life is the potential of concepts to expand the novelty of physical succession to the generation of abstract entities that endure.

NOTES

1. For example, from C. Lombroso, *The Man of Genius* (London: Walter Scott, 1891) to H. Gardner, *The Creators of the Modern Era* (New York: Basic Books, 1993).
2. Related to this distinction is that of the genius as an explanation of historical events or a product of historical forces. See W. James, "Great Men and Their Environment," in *The Will to Believe* (New York: Longmans, Green, 1896).
3. See R. Finke, T. Ward, and S. Smith, *Creative Cognition* (Cambridge, MA: Bradford, MIT Press, 1992).
4. A. N. Whitehead, *Adventures of Ideas* (New York: Macmillan, 1933, 227); Also, Henri Bergson, *Creative Evolution* (New York: Holt, 1913), for whom God is creation or a creative principle.
5. J. Dewey, "Time and Its Mysteries," in *Philosophers of Process*, ed. D. Browning (New York: Random House, 1965), 211. The same idea is found in H. Bergson, *Creative Evolution*.
6. D. Bohm, "On the Problem of Truth and Understanding in Science," in *The Critical Approach to Science and Philosophy*, ed. M. Bunge (Glencoe, IL: Free Press, 1964).

7. From this perspective, computer simulations of creative thought, e.g., S. Turner, *The Creative Process* (Hillsdale, NJ: Erlbaum, 1994) will be of interest only to the extent they capture the natural state, i.e., whole–part shifts.
8. C. Hartshorne, *Creative Synthesis and Philosophic Method* (London: SCM Press, 1970).
9. K. Goldstein, *Language and Language Disturbances* (New York: Grune and Stratton, 1948); K. Goldstein, *The Organism* (Boston: Beacon Press, 1939).
10. A. R. Luria, *Higher Cortical Functions in Man* (New York: Basic Books, 1966).
11. A. Schweiger, K. Chobor, and J. W. Brown, "From Diffuse Meaning to Phonology" (Paper presented at meeting of the International Neurolinguistic Society, Krakow, Poland, September, 1995).
12. Increased speed of recognition of a word, e.g., *nurse*, by prior activation of the category through exposure of a related word, e.g., *doctor*.
13. A. Marcel, "Phenomenal Experience and Functionalism," in *Consciousness in Contemporary Science*, ed. A. Marcel and E. Bisiach (Oxford: Clarendon Press, 1988).
14. E. Weinstein, and R. Kahn, "Nonaphasic Misnaming (Paraphasia) in Organic Brain Disease," *Archives of Neurology and Psychiatry* 67(1952):72–80.
15. S. Arieti, *The Intrapsychic Self* (New York: Basic Books, 1967).
16. The depth of origin accounts for other features of the schizophrenic disorder. Hallucinations are truncated object developments, delusions the play of word-meaning relations unencumbered by the drive toward denotation. Derealization is the presentiment of the cognitive origins of objects. The gaining of reality by concepts accompanies the incomplete exteriorization and "detachment" of objects. The loss of reality of objects coincides with increased affect in those concepts that are their precursors. The passive self of the dream invades the waking experience and invites the delusion that the individual is a victim for his own images to persecute.
17. S. Arieti, *Intrapsychic Self* (New York: Basic Books, 1967).
18. See also D. Krech and A. Calvin, "Levels of Perceptual Organization and Cognition," *Journal of Abnormal Social Psychology* 48(1953):394–400
19. Though their interpretation differs from my own, some recent studies are described in L. Robertson, and M. Lamb, "Neuropsychological Contributions to Theories of Part/Whole Organisation," *Cognitive Psychology* 23(1991):299–330.
20. M. Wertheimer, *Productive Thinking* (New York: Harper and Row, 1945), 235.
21. As noted by Karl Pribram. The psychological literature on this topic is extensive, as in science more generally. See; D. Leary, *Metaphors in the History of Psychology* (Cambridge, U.K.: Cambridge University Press, 1990); R. H. Brown, *A Poetic for Sociology: Toward a Logic of Discovery for the Human Sciences* (Chicago: University of Chicago Press, 1989).
22. G. Lakoff, *Women, Fire, and Dangerous Things* (Chicago: University of Chicago Press, 1987).
23. The inability of computational linguistics to account for metaphoric extensions of word meaning, N. Goodman and C. Elgin, *Reconceptions in Philosophy* (Indianapolis, IN: Hackett, 1988), 103–110, may prove lethal to computer simulations of creative thinking.
24. This section is based on S. Glucksberg and B. Keyser, "Understanding Metaphorical Comparisons: Beyond Similarity," *Psychological Review* 97(1990):3–18; and S. Glucksberg, "Beyond Literal Meanings: The Psychology of Allusion," *Psychological Science*, 2(1991):146–152.
25. Aphasic neology is usually semantic *plus* phonological errors; e.g., "twas brillig and the slithy toves." As with errors of word meaning, phonological errors can be interpreted in terms of blends and contextual effects on phoneme production, See H. Buckingham "Phonological Processes in Aphasia." (Paper presented at New York Academy Sciences meeting, October, 1994.)

26. For other examples, see M. Critchley, "The Neurology of Psychotic Speech," *British Journal of Psychiatry* 110(1967):353-64.
27. Conceptually original statements, not just novel sentences, which are the rule in language use. See J. Pind, "Computational Creativity: What Place for Literature?" *Behavioral and Brain Sciences* 17(1994):547–548.
28. L. Barsalou, "The Instability of Graded Structure," in *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, ed. U. Neisser (Cambridge, U.K.: Cambridge University Press, 1987).
29. A. Koestler, *The Act of Creation* (New York: Basic Books, 1964).
30. F. Schopenhauer, *The World as Will and Idea*, vol. 3, ed. R. Haldane and J. Kemp (London, 1907–1909).
31 See N. Hirsch, *Genius and Creative Intelligence* (Cambridge, MA: Sci-Art, 1931).
32. Lakoff, *Women, Fire, and Dangerous Things*.
33. See S. Carey and R. Gelman, *The Epigenesis of Mind: Essays on Biology and Cognition* (Hillsdale, NJ: Erlbaum, 1991); E. Markman, "Two Different Principles of Conceptual Organization," in *Advances in Developmental Psychology*, ed. M. Lamb and A. Brown (Hillsdale, NJ: Erlbaum, 1981), 199-236; F. Keil, "Conceptual Development and Category Structure," in U. Neisser, *Concepts and Conceptual Development*.
34. B. Tversky and K. Hemenway, "Objects, Parts, and Categories," *Journal of Experimental Psychology: General* 113(1984):169–93.
35. G. Lakoff, *Women, Fire, and Dangerous Things*.
36. G. Smith and I. Carlsson, *The Creative Process*, Psychological Issues 57 (Madison, CT: International Universities Press, 1989).
37. Cladistic categorization is based on a shared derivation in contrast to categories based on overall similarity. See E. Meyr, *The Growth of Biological Thought* (Cambridge, MA: Belknap Press, 1982), 226–233; see also Lakoff, *Women, Fire, and Dangerous Things*, 118–121.
38. E. Kris, *Psychoanalytic Explorations in Art* (New York: International Universities Press, 1952).
39. A. N. Whitehead, *Process and Reality* (New York: Macmillan, 1929), 31; E. Pols, *Whitehead's Metaphysics* (Carbondale, IL: Southern Illinois University Press, 1967).
40. H. Bergson, *Creative Evolution*, trans. Arthur Mitchell (New York: Modern Student Library, 1944).

Author Index

Subject Index

ISBN 0-306-45231-6
90000
9 780306 452314