The Journal of Mind and Behavior

Vol. 35 No. 1 Winter 2014
Vol. 35 No. 2 Spring 2014

ISSN 0271–0137

The Journal of Mind and Behavior (JMB) is dedicated to the interdisciplinary approach within psychology and related fields. Mind and behavior position, interact, and causally relate to each other in multidirectional ways; JMB urges the exploration of these interrelationships. The editors are particularly interested in scholarly work in the following areas: □ the psychology, philosophy, and sociology of experimentation and the scientific method □ the relationships among methodology, operationism, and theory construction □ the mind–body problem in the social sciences, psychiatry and the medical sciences, and the physical sciences □ philosophical impact of a mind–body epistemology upon psychology and its theories of consciousness □ critical examinations of the DSM–biopsychiatry–somatotherapy framework of thought and practice □ issues pertaining to the ethical study of cognition, self-awareness, and higher functions of consciousness in nonhuman animals □ phenomenological, teleological, existential, and introspective reports relevant to psychology, psychosocial methodology, and social philosophy □ historical perspectives on the course and nature of psychological science. We typically do not publish empirical research. The Journal also recognizes the work of independent scholars.

JMB is based upon the premise that all meaningful statements about human behavior rest ultimately upon observation — with no one scientific method possessing, a priori, greater credence than another. Emphasis upon experimental control should not preclude the experiment as a measure of behavior outside the scientific laboratory. The editors recognize the need to propagate ideas and speculations *as well as* the need to form empirical situations for testing them. However, we believe in a working reciprocity between theory and method (not a confounding), and in a unity among the sciences. Manuscripts should accentuate this interdisciplinary approach — either explicitly in their content, or implicitly within their point of view.

JMB offers a publication outlet on a quarterly basis. The Journal publishes one volume per year in the following sequence of issues: Winter, Spring, Summer, and Autumn. There are no submission fees or page costs for accepted manuscripts. JMB is a peer-reviewed, refereed journal, and all decisions will be made by the assessing editors, associate editors, and chief editors. Commentaries and responses to individual articles and reviews are welcome. Authors subscribing at the time of manuscript submission are eligible for reduced subscription rates (see below).

All manuscripts should follow the style and preparation presented in the *Publication Manual of the American Psychological Association* (sixth edition, 2010). Particular attention should be paid to the citing of references, both in the text and on the reference page. [Note exceptions to APA style: JMB uses *no* ampersands or city/state abbreviations in referencing; the Journal uses three levels of headings: level 1, level 3, and level 4, see pp. 113, 114, 115 from the fifth (2001) edition APA Manual.] Authors requesting blind review must specify and prepare their manuscripts accordingly. Manuscripts may be sent to the Editor either by e-mail to jmb@maine.edu or by post (one copy) to:

Raymond Chester Russ, Ph.D., Editor
The Journal of Mind and Behavior
Department of Psychology
The University of Maine
5742 Little Hall
Orono, Maine 04469–5742
Tel. (207) 581-2057

Yearly subscription rates are $32.00 for students or hardship; $35.00 for past/present JMB authors or for those submitting manuscripts; $46.00 for individuals; $185.00 for institutions. Air mail rates upon request. All back issues are available and abstracts are located at www.umaine.edu/jmb. For detailed information contact our Circulation Department at The Institute of Mind and Behavior, P.O. Box 522, Village Station, New York City, New York 10014; Tel: (212) 595-4853.

continued inside back cover

The Journal of Mind and Behavior

The Journal of Mind and Behavior

Vol. 35 Nos. 1 and 2

Winter and Spring 2014

Library of Congress Cataloging in Publication Data

The Journal of mind and behavior. – Vol. 1, no. 1 (spring 1980)–
– [New York, N.Y.: Journal of Mind and Behavior, Inc.]
c1980–

1. Psychology–Periodicals. 2. Socialpsychology–Periodicals. 3. Philosophy–Periodicals. I. Institute of Mind and Behavior
BF1.J6575 150'.5 82-642121
ISSN 0271-0137 AACR 2 MARC-S

The Journal of Mind and Behavior
Winter and Spring 2014 Vol. 35 Nos. 1 and 2

CONTENTS

The Journal of Mind and Behavior
Winter and Spring 2014, Volume 35, Numbers 1 and 2
Pages 1–20
ISSN 0271–0137

Feeling

Jason Brown

New York City, New York

This article attempts to establish on a psychological basis some foundational principles of a philosophy of mind grounded in process (microgenetic) and evolutionary theory, with a focus on the micro-temporal or diachronic aspects of mental contents and the derivation and intra-psychic structure of the mind/brain states in which they are ingredients. The subjectivity of the approach is in contrast to the externalist stance of cognitivist theory, a distinction with a venerable history. For example, Bosanquet asked, "is mental growth a process of compounding units . . . or a process of discrimination?" and cited James as preferring to begin with "the more concrete mental aspects . . . (and go) to elements we come to know by way of abstraction." James went on to write that the "process of 'building-up' the mind out of its 'units of composition' has the merit of expository elegance, and gives a neatly subdivided table of contents; but it often purchases these advantages at the cost of reality and truth." James insisted on a focus on entire conscious states rather than "the *post-mortem* study of their comminuted 'elements' (which is) the study of artificial abstractions, not of natural things."

Keywords: emotion, energy, feeling

. . . the individual can be cheerful and happy only if he has the courage to feel himself
Goethe

For microgenesis, the process leading to a conscious endpoint is, together with the final content, part of an epochal state. the outcome of which — an act, object, word — is not a resultant of the preceding series but incorporates its earlier segments — value, meaning, belief — as part of what it is. An object includes its formative phases. The subjective has inner and outer segments. The world is the surface of the mental state. Final actualities specify pre-object phases which detach and articulate mind-external. One effect of process-thinking is a revival of the underlying continuities in the diverse aspects of cognition fractured

Correspondence concerning this article should be addressed to the author at 66 East 79th Street, New York, New York, 10075. Email: drjbrown@hotmail.com

by analysis. In process thought, wholes are potentials or categories for specification. Parts are not in situ in wholes but are novel derivations that serve as sub-categories for ensuing partitions.

The account of Feeling from this point of view traces conscious experience back to the physical foundations of existence, or from the facts of perception — objective data or their appearances — to a deeper reality. We sense Feeling in activity and passivity, or agency and receptiveness, a dynamic that underlies mentality and is accentuated when its direction is impeded, as in tension, hesitation, or anxiety. If we could eliminate acts, objects, or mental contents in a momentary cognition, mental activity would likely be felt as pure feeling without origin or subjective aim. The lack of direction or intentionality would suspend the feeling of before and after and result in a felt stasis of energy. The pure Feeling described in this paper is thematic in the evolution of mind, and foundational to the derivation of instinct, drive, desire, and emotion. An emotion is a complex of feeling and idea — conceptual-feeling — that is a motive and an object for the self. Feeling is a deeper activity, prior to emotion and idea, out of which emotion and other contents of mind develop. The isolation of actualities from antecedent possibility, the force and specificity of conative drive, the sequence that brings entities into existence, are signatures of Feeling as the engine of evolutionary advance.

Energy is the foundation of matter. At the earliest stage of inanimate entities, Feeling, as energy, is non-directional, best described in the language of physics. We can speak of Feeling when the recurrence of asymmetric energy underlies the direction and cyclical nature of the organism. Feeling evolves when the recurrence of energy of an entity becomes uni-directional. Feeling, though non-relational and uniform, distributes into concepts that embody feeling as affect or emotion. Emotions such as drive or will, pain and pleasure, approach and avoidance, are vectors of Feeling that distribute into feeling, or as energy into emotion, as the essential dynamic of existence.

Introduction

Feeling is central to many philosophies, particularly those of Whitehead and Bradley to which I am greatly indebted, works that take differing positions on the nature of Feeling as the ground of existence and the relation to mental contents and those entities into which Feeling distributes. In this paper, Feeling, in capitals, refers to the basic activity that generates matter and life, while feeling, in lower case, refers to that which is proximate to and innervates the emotions. Pain, or its avoidance, and pleasure or its attraction, differ from Feeling, in that they involve, or are a felt dynamic in, the categorical form that embodies them. Feeling begins in the physical datum as energy, and evolves to feeling or emotion as a subjective quality. Many authors ascribe feeling to pain and pleasure with an objective component (e.g., Ward, 1920), and a subjective

aspect in response to sensory presentations. Whitehead[1] conceived feeling as an operation of passing from objectivity to subjectivity. This would agree with the present discussion if by passing from objectivity to subjectivity means the origination of Feeling as energy in material entities and its segregation in the objects (concepts) of subjective states, e.g., organisms, along with the replication of this process in every recurrence.

Bradley (1893) was closer to my account in treating Feeling as a complex unity without relations, an experience of many in one and genetically the first layer of experience (see discussion in Rusu, 2013). Primal Feeling is undivided and directed energy that partitions into feeling in relation to concepts, as drive, desire and emotion (see below). Feeling does not reduce to sensation, though for some authors feeling and sensation are identical. In the context of microgenetic theory, sensation is external to perception.[2] Those authors who relate feeling to sensation confound sensibility with perception, i.e., the physical with the endogenous, and they confuse perception and experience with emotion. In addition to the reckless employment of association terminology, the error lies in the interpretation of sensation as internal to mind and identified with stimulation, as stimulation is with feeling, and feeling is with pain and pleasure.

My position follows Bradley, in that Feeling is intrinsic, non-relational, uniform and non-decomposable. The account resembles Whitehead's idea that Feeling is a sub-atomic process (vibratory strings?) that, through *concrescence*[3] or *microgenesis*, actualizes the varied forms of mentality as intimations of the deeper, less-differentiated life of organism. Feeling is a quality that propels evolutionary process from its origination in inanimate nature and non-cognitive entities to its manifestation in the higher mentality, exhibiting trends in nature that transfer to the human brain as a physical entity. There is no Rubicon, or point of transition, from the inanimate to the living; rather, a continuous elaboration of Feeling into higher grades of organization and complexity.

The metaphysics of Feeling are not explicitly psychological but inevitably course through individual cognition. The contents of mind — taken as real or phenomenal — are manifestations of the Feeling that gives rise to them. These manifestations, such as ideas and emotions, are in constant transformation. Feeling,

---

[1]See Stenner (2008) for discussion of Whitehead and subjectivity. Microgenesis has an affinity with some concepts in process philosophy but the theory developed independently in clinical neuropsychology leading to a novel account of time, change, process and the mind/brain state (Brown, 1986 to 2010).

[2]The relation of sensation as a physical constraint on perception, to perception as an outcome of endogenous process, applies to both interoceptive and exteroceptive sensibility, e.g., pain and visual perception.

[3]For Whitehead, concrescence by way of feeling (prehension) resolves the many to the one, while microgenesis postulates a progression from unity to diversity.

though directional, is undifferentiated, comparable to an intrinsic energy that animates the organism. The account of Feeling as distinct from energy begins and ends with an internalist perspective that traces conscious experience back to the foundations of existence, or from the actual facts of perception — objective data or their appearances — to a deeper reality. William James believed that a final understanding of psychology would be metaphysical. I would add that psychology should be the starting point of metaphysics through which, in any event, it surreptitiously passes. Metaphysics encompasses universal, indeed, cosmic wholes, but the micro-temporal history of Feeling, from top to bottom, from an individual consciousness to the immensity of space and the diversity of nature, is an account of the physical dynamic of human mentality. In a word, metaphysics is metapsychology barren of the psychological data by which philosophy should be informed.

In ordinary language, Feeling implies a relation to emotion or an affective tonality that suffuses experience and enlivens objects. An example might be the postulation of an affect-pool or libidinal stream that distributes into specific modes of cognition. Perhaps there is a relation to the Chinese *Qi*. However, a more accurate depiction is that emotions crystallize a tacit background into particular occasions of experience, some of which accompany consciousness. An emotion is a complex of feeling and idea — conceptual-feeling — that is a motive and an object for the self.[4] Feeling is a deeper activity, prior to emotion and idea, out of which emotion and the contents of mind develop. Conceptual-feelings represent the precipitation of Feeling into affectively-charged ideas. Mental contents and events, when peeled away to expose their originating activity, reveal a convergence of matter and life that is the covert, intrinsic, and impalpable quality of Feeling they embody.

Feeling is uniform but its manifestations in affect, value, and emotion are protean. They assume many masks and inhabit all modes of thought and varieties of experience, differing in shape, intensity, and character. Feeling is non-relational, or rather, pure relationality, and the "stuff" out of which relations develop. Objects precipitate out of the flow of Feeling. In that objects are the outcomes of a subjective aim, they entail directionality.[5] All substances are forms of *purposefulness* in the relation of origination to actuality. In the human mind, the transition from core to surface, from the initiation of a mental state to its terminus, or from the

[4]The relation of Feeling to instinctual drive, and drive to desire, is detailed in Brown (2012a), which includes a lengthy discussion of similarities (few) and differences (many) with psychoanalysis. The theory is closest to that branch of psychoanalysis represented by Schilder (1951) and Rapaport (1950).

[5]For Whitehead, the subjective aim is the direction to value. Here, the term is used similar to intentionality in the aboutness of the mental state, its progression to an act, object, or idea. An object is an externalized image with its affective tonality, e.g., value. Value is not attached to objects or projected on them, but is specified with the object in its momentary journey from mind to world. The link is Dewey's argument that facts are irreducible values.

onset of an epoch to its perishing and replenishment,[6] is a relation of the immediate past of an existent streaming to an immediate present. The process leading from initiation to actuality creates a present that, in its forward momentum prepares for an oncoming future. The sense of movement to the future in the constant replacement of mental states is such that the developing replacement gives the occurrent state a feeling of an immediate future. This is why we can never grasp the present, for as it lays down the now of the moment it is felt as the seed of the future in the overlap if its replacement, even if, indeed because, the ensuing state is not yet actual. This together with the forward progression from memory to perception, or from past to present, and the disappearance of the present in the oncoming state, give the feeling of a future constantly revealing itself as idea and presentiment.

The present is delivered out of the personal past on the way to perceptual adaptation. Possibilities envelop and individuate the final datum. Need adapts to necessity as contents are selected. The present, or the outcome of the present state, resolves the contingency of conceptual possibility with the uncertainty of external events in a changing world. The specification of acts elaborates the choice inherent in agency. A conation to the present is the seed of purposefulness, reminiscent of Whitehead's (AI:249) insight on an occasion of experience, that between an effect facing the past and a cause facing the future lies the teleology of the Universe. The present surges into existence as a forward impulse to satisfaction, advancing the inheritance of an onto-phyletic past in a traversal through which objects resolve out of change to create a novel universe.

Evolution and Panpsychism

The teleology of Feeling is a speculation on the final aim of evolution. The idea is that anisotropic Feeling empowers evolutionary advance. The continuity of Feeling and distribution to affects, together with the basis of substance or being in the epochal nature of process (becoming), are sufficient to account for evolutionary gradualism as successive stages of proto-mentality without the postulation of emergence, e.g., of consciousness. The higher levels of realization to which evolutionary process leads are not attractants to their attainment but adaptations of Feeling and its implementations in the striving of antecedents to a further level. The account is consistent, though not dispositive, with regard to intelligent design. The intrinsic activity that underlies a surge to finality, the relation to embryogenesis and growth, the recurrence of successive epochs, intrinsic relations, and the conformance of microgenetic and evolutionary

[6]Michel Weber put this succinctly: actual entities "come into existence and then sediment into being (not vanish into nothingness)."

pattern, mark a wave of inherent purposefulness or proto-intentionality that tends to be obscured in a sea of contingency.

Energic process, displayed in the transition from whole to part, or from generality to definiteness, carves particulars out of categorical wholes. Contingency and chance are the *externalist attributes* of "adaptive strategies" that cede change to the environment, e.g., elimination of the unfit, while the internalist response is that uni-directionality carries a pending aim to realization. The isolation (sculpting) of an actuality from antecedent possibility, the specificity of conation or drive, the sequence that brings entities into existence, are signatures of Feeling as the engine of evolutionary advance. Energy is the foundation of matter. At the earliest stage, Feeling is non-directional and best described in the physics of elementary particles. In the gradual shift to vegetation, energy assumes directionality in primitive life forms, or perhaps one could say the life-forms channel energy in the direction of becoming and/or growth. Birth, growth, death, and re-birth exhibit direction within recurrence. The asymmetric manifestations of energy combined with successive recurrence underlie the direction and cyclical nature of organism and physical matter. Every entity, object, and mind/brain state is an epoch of becoming into being.

The account of evolution as a population dynamic centered on group-speciation contrasts with the microgenetic interpretation of the specification of final existents — particles, organisms — in relation to which population effects are secondary. Put differently, the societal thrust of evolution is empowered, generically, by subjective aim, potential at onset but developing in conformance with external conditions in the assimilation of drive to adaptation. Pleasure and avoidance of pain are goals for objectification or constraints on deviation in the path to satisfaction. The guiding principle is survival of the fittest, but the first priority is re-instantiation of the organism. Self-preservation is self-replication, which entails the recurrence (? causal persistence) of the organism. The motive force of recurrence is hunger,[7] which is prior to sexual drive. The latter entails engagement with others in the service of the population, replacing the organism with progeny. The perpetuation of species — their persistence and change — is an outcome of the recurrence and perishing of individuals and the renewal of overlapping and continuous epochs — representations — of the world.

From this standpoint, the population dynamic in speciation is a group model of an intrinsic process of self-actualization. To step outside the survival of the individual to the survival of the population preserves the pattern of specification even as it externalizes the intrinsic process of individuation. Microgenesis expands

[7]See Brown (2012a) for discussion. The hunger/thirst drive leads to feeding and avoidance, which promote the survival of the individual, while the sexual drive, which appears later in maturation, accounts for the survival of the population. The recurrence of the individual is prior to the replication of the species. The chicken recurs each moment before an egg for the future is laid.

the renewal and adaptation of individuals to the renewal and adaptation of species. In the species, individuals can be sacrificed as long as the population endures. Sexual drive sustains the group, and thus has a particular prominence. But from the individual perspective, the preservation of the solitary organism is paramount with hunger and feeding primary.

The basic manifestation of Feeling is the existence of an entity. Its most elaborate articulations are the intellect and values of the human organism. The continuum from the smallest particle to the complexity of a material brain begins with intrinsic value in material entities. Rocks and organisms are differing patterns and complexities of atomic units. The intrinsic value of a rock, its existence, is the energic process that generates the entity, and through which it recurs. Intrinsic value in organism is analogous to that in physical entities. A rock that is replicated (replaced) over some duration is a far-distant precursor of human cognition. In the physical world, a rock is an aggregate, a "society"[8] or compound of atomic particles. In perception a rock is a whole, with the potential for realness, meaning and value (as worth), e.g., for a wall, weapon. Feeling runs through all things great and small. Some suggestions as to foundational process in nature continuous with consciousness include an uncollapsed wave packet, virtual photons (Romijn, 2002), quantum entanglement (Shields, 2009) and microdurations (see critique in Hunt, 2001). Dombrowski (2001, p. 32) has written of the "microscopic sentiency found in cells, atoms, and particles."

A particle is a basic object that science represents as self-identical over time or context-independent. Yet a proton in a stellar mass is different from one in a hydrogen atom, as a brain cell in a tissue culture differs from one in an active brain (Birch, 1990). An elementary particle can be conceived as a waveform of energy that is epochal or, if quasi-epochal, developing out of a space-time continuum such as Bohm's implicate order. The epoch is the temporal extensibility of the particle, i.e., the minimal duration for the particle to exist. Energy condenses to a particle; duration is an epoch or category in the momentary life of a particle, over which the particle becomes what it is. *The process of Feeling that accounts for the existence of an object — particle, brain — is its intrinsic value.*

In my view, evolutionary thinking not only opens the door to but obligates some form of panpsychism.[9] The distinction of panpsychism and/or panexperientialism from emergentism defends a continuum against assaults that disavow conscious experience in material entities. The skeptic will contest the argument that rocks or particles have experience, psyche, or a primitive mode of conscious-

[8]In Whitehead's terminology: a "corpuscular society" (Cobb, 2008).

[9]There are similarities with the views of William James, who advocated a form of panpsychism in which reality consists in innumerable flows of feeling interacting with each other. Human consciousness is one kind of flow of feeling typified by the high level of conceptual thinking it contains (Sprigge, 2005).

ness. But the postulation of a continuum merely implies precursors of higher cognition. Energy fills this need. The alternative to panpsychism is sudden mutation or a saltatory leap[10] from one mode of perception to another or a spandrel that is epiphenomenal to other adaptations, e.g., language, but even if true, this does not exclude the need to explain how consciousness arises, whether the crucial step involves genes, cells, or connectivity patterns and/or complexity.[11] Does consciousness appear at a given point? Is it an emergent? If so, what is the distinction of an emergent from a resultant, the causes of which are unknown? The fatal presumption takes consciousness as an abstract entity that is the goal of the inquiry rather than describing phenomena that make consciousness possible.

Intrinsic Value and Existence

In sum, bi-valent energy interior to a particle evolves to uni-valent Feeling interior to organism. The step from physical matter to living organism transforms isotropic energy to anisotropic Feeling, with one cycle (packet) of energy constituting an epoch of existence. The intrinsic nature of the process is the seed of higher subjectivity. In basic entities or in brains, Feeling is non-relational, since the relata — the onset, terminus and phases of the energic wave — are not inter-related segments or polarities. We would not say, except in a trivial sense, that in a fountain the initial jet of water or an intermediate segment relates to the spray at the surface, since the former become the latter. The relations of feeling embodied in categories are over phases in the mind/brain state, and their overlapping replacements, while Feeling is the non-relational ground out of which these phases develop. Phases in the mind/brain are non-temporal (simultaneous) until the state actualizes with relations between phases (whole–part transitions) in abeyance until the sequence terminates. The question of external relations is beyond the scope of this paper, which concerns intra-psychic or subjective phenomena.

It is a long way from a particle to a brain, yet the pattern of process is comparable. The brain is a complex physical entity, existing like a particle as a duration of constituent phases. A brain state is a hierarchic series of vibratory patterns that pulse each epoch into existence. The pattern of neuronal activity over phases — rhythmic, oscillatory, cyclical — is analogous to the vibrations of a

[10]The theory of punctuated equilibrium may be applicable to this problem. Though popularized by Gould, I recall this idea discussed well before by Richard Goldschmidt, my genetics professor at Berkeley, based on studies in drosophila.

[11]Size alone is not explanatory, since the brain of a gorilla is larger than that of some bushmen and dwarf brains (Dart, 1956; Lenneberg, 1967), while complexity has to be parsed in terms of a spatiotemporal pattern of activity, not mass and intricacy of connections. Complexity is not an explanation. The Beethoven 5th played backwards is complex noise.

particle (Gunter, 1999). A neuron exists as the momentary envelope of its activity pattern. We have no precise knowledge of psychic experience associated with a neuron, nor for that matter, with the innumerable neurons in any region of the brain, nor the presumably quiescent neurons in a sleeping brain. A complex pattern of activity is essential for cognition, not only a *spatial configuration but a temporal* sequence in large neuronal populations over distributed systems in forebrain evolution.

The argument is that the intrinsic value (existence) of physical entities consists in a waveform of energy over the temporal extension of an entity, i.e., in Whiteheadian terms the minimum duration of process needed for entities to be what they are. In this bare epoch of existence, the waveform of the thing — its vibratory or oscillatory structure — is non-directional or isotropic. Gradually, internal relations expand as the nucleus of a shift from energy to Feeling; process takes on direction and becomes anisotropic. Internal relations in the epoch expand. The duration of process that constitutes existence enlarges as intrinsic value.[12] Existence becomes *realness* as the ground of later derivations, e.g., drive, instinct, desire. With respect to value, which includes interest, the existence of an entity precedes its realness (not reality), which leads to object-worth (Brown, 2005).

Relations

As noted, the idea of relations implies things or terms that are related, but from an internalist standpoint the relations of things and terms are outcomes of the process through which, as momentary exemplifications, they come into existence. Put differently, process lays down the terms (categories) it is presumed to relate. These pass to ensuing categories but do not constitute terms in relation. There is a temporal dimension in precedence or simultaneity, yet as mentioned, the relationality of endogenous process is non-temporal until it actualizes. While the Feeling that generates the state is non-relational, earlier and later phases in the state, which are articulated by feeling, have a relational quality in respect of their succession, but the phase-transition of the mind/brain state cannot be punctuated by a temporal locus. One can only say, on completion of the transition, that A precedes B, but A is not in the past of B, since A and B are co-temporal in the epoch. Moreover, a locus between past and future within a state requires that a series of mental states generate a present, or perspective, one that is felt in the present and one that is felt in the past. This past–present is an ascription that assigns a past-point that breaks earlier–later, with events facing the past of the point or its future.

[12]For historical precedence for the idea that existence is the initial step in value-creation, see Perry (1926).

An actuality that represents a succession, e.g., a thought, an object, is an epoch or series of epochs. Indeed, all objects are events, or clusters of epochs. The temporal order of ingredient phases becomes real when the epoch actualizes. In this process, the greater part of the mental state is devoted to recurrence, either as the underpinnings of the state or as memory. Except as an eternal object, the past exists only in the present. Each mental state and each state of the world replaces — and thus in some sense, remembers — the preceding one. The transition from before to after in a given mind/brain state is comparable to physical passage in the world, while the present gives a perspective for a conscious time series, i.e., the A and B series of McTaggart (1901). A sequence of perceived states, i.e., the perception of change, is, as Bergson (1923) noted, wrongly inferred as a kind of horizontal line extracted from the replacement, whereas continuity owes to the superimposition of an epoch on its precursor and the transposition of epochs to perceived successions, i.e., to a sequence of final actualities. How the simultaneity of a subjective past, which is embedded in a present, unfolds to the serial order of events — inner and outer — is a complex problem discussed in Brown (2010). The shift from one temporal series (before/after) to another (past/present/future) corresponds with the actualization of the state and the perceptible sequence of completed states. In this shift, the authenticity of internal relations within a state transforms to the illusion of external relations from one state to the next.

Feeling generates an indivisible cycle of becoming into being. Think of the inability to divide upstream and downstream segments of a river to relate one segment to another. A stretch of water may seem to correspond to some arbitrary point, e.g., the river bed, but the stream is a traveling wave in which one segment becomes the next. In a stream, as in microgenesis, or in the actualization of any entity, the earlier becomes the later. One would not say an orbiting electron in a hypothetical atom relates to, or is a relation of, one point to another. The motion of the electron creates the relation, i.e., the trajectory establishes the relation rather than the relation determining the position. An atom cannot be said to exist in the absence of a complete orbit. Similarly, an entity cannot be sliced in time but exists when it becomes what it is over its becoming. Feeling elaborates things to be related without relating them. This raises the question: What are those things and how are they created, and does Feeling change with changing objects and, if not, how do objects change if feeling is unchanging?

To repeat, Feeling — as opposed to feeling — is non-relational, or not relational in the ordinary sense (of external relations), in that it is continuous throughout one cycle of actualization. The relationality is constitutive, not interactive. The cycle forms an epochal or modular whole in which the succession within the state is a continuous becoming, while continuity in the succession of states is by way of the overlap of states. A state, epoch, or actual object does not, at least not in its conscious appearance, cause the next to occur. Instead, the object is relin-

quished in the replacement. The epoch is simultaneous with the phase-transition that configures it, such that it exists before the succession has temporal order or exists in time. Phases in concert with the entity as a whole are condensations of Feeling, while the final epoch — emotive, linguistic, etc. — embodies, or is a category of, internal relations of the entity. Such relations, like sine waves, are purely relational without things to relate to.

For example, the desire for an object can be interpreted as a relation between the self that desires and the desired object, but the object (idea, concept), whether perceived as internal or external, is part of the same mental state as the self and the feeling of desire. As the subjective aim of the state, the object does not stand in relation to earlier phases but incorporates and actualizes the entire sequence. The object consists of all of the phases through which it develops, so that value and meaning are not attached to, or projected on, the object after it is perceived. The growth in value of the object is coincident with its replacement; the object does not remain constant while its value changes, it changes with its value in each perception. This is the import of the epochal nature of the state. Similarly, an utterance or perception may seem to consist of separate components — conceptual, semantic, phonological — and subsume a variety of processes and contents, but the components (categories), which are manifestations of the Feeling that distributes into them, are incorporated in the trajectory.

More precisely, Feeling obtains in the internal relations of objects and organisms as the process through which objects objectify. Value arises in the transformation of existence to realness.[13] With an expansion of subjectivity, realness becomes attention or interest and, finally, is exemplified in drive, desire and object-worth. In the actualization of the mental state, Feeling bifurcates into subject and object, filling and imbuing the observer and, *inter alia*, infusing the object with realness. Interest is value sequestered in an object. It is a thread of mentality that "connects" mind and world, i.e., an objectified portion of subjectivity in a perceptual target. In the partition of inner and outer, and in differing proportions, Feeling is allocated to self and other, with the latter a tributary of the self that does not so much attract emotion as enjoy it.

Emotion

An account of the process generating an object or mental state as directed Feeling must address the categories — internal and external — into which Feeling is directed. How do categories arise, and what is their relation to Feeling? The categories most closely related to Feeling — the drives and their specification to the emotions — is a good place to begin. The energy that evolves to Feeling

[13]Realness is the quality of appearing or feeling real. Images can seem real, e.g., dream, hallucination, though not object-like in their realness.

leads to instinct, drive, and the aggressive and defensive vectors. These qualities of Feeling are sediments of process, not energy in association with idea. They constitute the form taken by the becoming, i.e., the components and behaviors into which Feeling distributes. An emotion collapses Feeling to a specific affect within a category. When an emotion or thought objectifies, the category embodies the Feeling that pulses in the background as its source and incitement.

Whitehead noted an "analogy between the transference of energy from particular occasion to particular occasion in physical nature and the transference of affective tone, with its emotional energy, from one occasion to another in any human personality" (AI: 242). Strictly speaking, the analogy is a likeness due to the progression of physical energy into organic Feeling. Emotion is a later stage in the evolution of energy, or a higher grade in the ramification of Feeling. Any emotion objectifies personal feeling together with a concept, the varieties and complexities of which obscure the authenticity of Feeling as the primary impulse of organism.[14]

Feeling considered apart from the objects and emotions that are its derivations is not a content in consciousness but prior to its differentia, enlivening the things and events into which it distributes. The acts that for some define a person, together with the inner life that for others is the greater part of individuality, still do not adequately represent the deeper current from which, like froth on the surface, evanescent behaviors are specified. An act of cognition seems to be experienced directly, but there is an unfelt lag in its conscious realization. Moreover, the feeling in an act or idea may seem outside the contents that are felt, though Feeling, which is not itself felt, is what feels them, as modes of subjectivity to which self-realization refers.[15] Only with the intensity of early cognition is the feeling within an act inseparable from the content, though even then, especially in higher organisms, there is no doubt some unfelt antecedent lag. The "I" that loves or fears, even if overcome on occasion, is generally felt as distinct from loving and fearing; the lag begins before the category of the self — the "I" — develops.

Uniformity of Feeling decants to locality in acts and objects. To feel is to feel a relation between the actualized state and a ground that is itself largely unfelt. Experience is for or within the local content into which Feeling distributes. Even agitation gives actuality to feeling as a rupture of uniformity. Admittedly,

---

[14]The partition of feeling into the variety of emotions and affect-ideas is a complex topic outside the scope of this paper. An attempt to deal with this problem, along with a critique of Freud's account and the James–Lange theory, is in Brown (2012a).

[15]The widespread effort to more deeply understand the genuine or authentic self (Brown, 2005), and the variety of methods used, from psychoanalysis to meditation, points to the artifice on which a life is constructed, the intuition that mind has not been fully plumbed in thought or action and that depth is not content but Feeling, namely that human behavior cannot realize the inaccessible truth of an unconscious mentality.

it is unclear what evidence would suffice to turn what is obligated by a theory into a fact that supports a truth. Wordsworth alludes to the precedence of belief over fact:

> "And 'tis my faith that every flower
> Enjoys the air it breathes."

Feeling and Category

The encapsulation of feeling by emotions that express concepts in diverse exemplifications is a transformation of invisible uniformity into palpable actuality. Each mental state is modulated by constraints of the just-prior one, by ingrained patterns of process and connectivity that relate to habit and character, and by adaptation to the world or the adjustment of need to circumstance. In the human mind, hunger, and sexual drive partition to desire and intentional aim. With an expansion of this trajectory, there is an opportunity for segregated accentuations within a fully-realized epoch. Drive, desire, interest, worth, all modes of value, are the dominant affects that specify ensuing phases, e.g., drive to desire, or to individuate a designated focus, e.g., desire to love, fear, hope, etc. The final object also undergoes specification as attention or value shifts from whole to part, for example, attending to a person, a face, eyes, voice, and so on. There are fluid shifts from category to sub-category, e.g., drive to desire, and the reverse, not as a regression but an incomplete revival.

A category is a set of actual or potential objects that are related by shared attributes, but it can be thought of as an envelope that frames an entity, an elementary particle, a mind/brain state or a segment within the state. In this latter sense, in framing the micro-temporal development of an act or entity, the category is the being or "substance" of the entity, while the Feeling within the category, the micro-temporal transition through which the entity develops, is its becoming or process. The relation of mass to energy foreshadows that of substance to process. This relation is replicated in the categorical primitives that enclose a drive, or in the conceptual-feelings that embody affectively charged ideas or objects.

In prior writings, I assumed the process transitioned from one phase to the next, but I would now argue that phases constrain the passage of a single wave, with traversed segments isolated after the transit becomes actual. The transition cannot be construed as a causal output of one phase to the next; instead, the traversal, which is continuous, is submitted to constraints at successive phases. A mental content — inner or outer, category and Feeling — is a phenomenal outcome that, for the moment of its appearance, is an endpoint in the partition. Thus, the response to the question, what is the relation of Feeling to category, is that the category embraces formative phases that constitute the being of an entity, while the process of Feeling embodied by the epoch is its becoming.

In the evolution of mind, becoming undergoes an expansion of the inward subjectivity of organic life. Successive phases in actualization form sub-categories within the whole of mind/brain. The category of drive corresponds to a population of neurons and connections configured by synaptic strengths that sequester Feeling at its inception. Innate patterns of activity draw in or concentrate Feeling rather than Feeling conjuring up behavior. Similarly, in the derivation of drive to desire, the concept (category) accompanies an influx of Feeling, such that one can speak of a strong or weak desire. But a concept is not a container of Feeling. A drive can swell with feeling and discharge, or transition with a quota of Feeling to desire. A portion of residual Feeling remains in drive as the engine of desire; another goes with the conceptual partition to desire. Indeed, without Feeling, desire would be a non-directional idea, like a dictionary entry. In sum, an affect-laden concept actualizes a subjective aim within a mind/brain state. Images and objects are derivations of earlier configurations that receive Feeling from a source-drive, as objects receive implicit beliefs and tacit knowledge from source-categories.

In human mind, the more primitive category of drive and its concentration of Feeling partition to the sub-categories of self, desire, and the objects or images of desire. The quality of Feeling is uniform within the state, though at one moment it is emphatic in unconscious drive, at another, it is pronounced in the conscious self and desire, and then, subdued, is carried through to its objects. These transitions, markers of value in the dynamic of Feeling, dominate a mental state for a moment and pass like eddies in a stream.

The distinguishing features of a concept or category are the lack of precise boundaries and the virtual or categorical nature of content. Every member of a category is still a category of subordinates, which in turn are categories for further partition. A category can be innate such as instinctual drive, or acquired and spontaneous with its own affective charge, such as "things to take to a picnic." A category, which gives specificity to the feeling it incorporates, "splits" into subordinate categories in a forward development, or "descends" to a source that is essentially bottomless. The contents of mind — ideas, emotions — vary with the categories and their feeling-tone. Feeling gives force and valence to ideas, while ideas embody emotions and give variety to the mental life. Weak and strong feeling, such as affection and passion, reflect an influx of Feeling, which must be reconciled with the fact that Feeling assumes specificity by virtue of the accompanying idea, i.e., a difference of degree becomes a difference of kind.[16]

[16]Alternative accounts in which feeling attaches to, or is attracted by, an idea, or that feelings are specified prior to attachment, or that ideas "find" and combine with appropriate feelings, i.e., that Feeling is a composite of feelings, are unsustainable (see Brown, 2012a, for discussion and critique).

On this view, Feeling accompanies the category *ab origine*, distributing into ideas as local affects or emotions and constituting the affective quality of concepts. A core category is partitioned to more refined concepts and affects. A specific affect or emotion along with its concept, e.g., envy, pride, humiliation, etc., may seem a distinct affect-idea, but it is a tributary of Feeling that invests every aspect of the mind/brain state in the fractionation of drive to partial affects and ideas (conceptual-feelings).

An organism is a unity and a multiplicity. For Whitehead, a society of parts is prehended in the concrescence of a novel individuality, a kind of fusion by way of feeling of the many into the one. In the generation of a mind/brain state, myriad units along with their own pattern of Feeling, assimilate to larger organs, like the strings of one violin in an orchestral piece of music. The "gathering-up" of elements into oscillatory patterns in the before/after of the mental state, goes from whole to part, or from unity to diversity. This transition underlies a shift from the purposefulness of drive to the manifold of desire, and its partition to acts, images, and objects. Mind reaches outward to create and fulfill its own diversity as Feeling binds disparate elements. Multiplicity at the outcome of the state unpacks potential at the base.

It is likely that the vibratory foundation of minute elements, e.g., kinetic energy in cellular units, combines and transforms to virtual oscillators. The progression from a foundational or fundamental frequency to a series of harmonics at levels in speech production (Brown, 1986) provides a model for prehension as a reverse of this sequence, i.e., from the smallest units of organism, or neurons in brain, to a rhythmic oscillator parsed to successive frequencies. In a word, it is vibrations all the way down or, as Heraclitus put it, the road up and the road down are the same.

Becoming

A theory of Feeling is a psychology of becoming. The postulation of becoming in opposition to an ontology of being is a distinction that goes back to Heraclitus and Parmenides. The nature of thought, which is that of increasing analysis, favors substance theory in the fractionation of wholes into parts, and the solidification into observables of the invisibility of transition. Substance is palpable, while the process basis of substance has no boundaries or resting points. A quantitative psychology in which objects are stabilities avoids the change in objects, or how a changing object is recognized as the same. Walking man and sitting man are the same man, as is the man who grows, ages, adds a beard, is healthy, sick, or differs one moment to the next and over the lifespan. An object seems to remain the same in spite of incessant change. This way to think about objects appeals to common sense; if every change no matter how great

or small created a novel object there would be an infinite number of worlds, and selves.

On this view, change is adventitious, with properties by and large inessential to the object. For example, bi-pedalism is part of a common definition of human but not the use or disuse of the limbs, in walking, running, sleeping, or paralysis. Certain properties are critical or paradigmatic to the object, as to its sameness, such as, in man, consciousness, and personality. In substance ontology, change is not ingredient but something an object undergoes, an activity or inactivity that is extraneous. Substance ontology internalizes perceptual objects, carving up and populating the psyche with "logical solids" similar to those in the world. The ontology is reinforced by speculation on the timeless, eternal, or changeless ideas of Platonic thought, or the cognitive stasis of a set of "bloodless categories," or a lifeless past forever fixed in time. Energy is the stuff of particles but Feeling is conceived as supplemental. Ultimately, substance is a composite of a host of elements all the way down to basic entities that are themselves compounds of external relations.

In contrast, the dynamic of process theory, or becoming, which is fleeting and unobservable, must explain why the world seems to contain innumerable substances, how they are stabilized and how they appear independent of the observer, since the cognitive process that underlies substance has no perceptible correlates. Becoming condenses into the affective content of the objects of thought and perception. Yet we feel and observe an arousal of ideas by emotion, or emotion as an incitement to action. This accentuates the energetics of emotion at the expense of conceptual form. All mental contents and objects are categorical frames of Feeling. From an internalist perspective, categories are consolidations of embedded phases. From a physiological standpoint, they are segregated nodes or vibratory levels that enfold primordial Feeling. The Absolute of process ontology is relational and dynamic, like the *pratītya-samutpāda* of Buddhist metaphysics (Brown, 1999) or contemporary string theory. Object-formation is a becoming into being; there are no solid things. A rock, Whitehead wrote, is a mass of raging particles. Process and substance ontology are complementary. The "rock bottom" foundation of the world is not at all rock-like; it is purely relational.

Objects seem to change before our eyes and effect change on other objects. In process theory, change is in the becoming of actualized categories. Every change is a novel recurrence, and every recurrence is a changed world. This resolves a long-standing problem in causal theory, how a cause is carried into an effect. The disappearance of the cause in the effect represents the perishing of the past as it is overlapped and replaced by another world. The impression that change occurs across states rather than within them is due to the "invisibility" of becoming at the interior of a forming object. The idea of change in external passage conforms to common sense but an account of change in process is deep and counter-intuitive. Thus, a major difference in the ontology

of process and substance theory is the locus of change, either in the becoming of an object or its causal transmission to another object.

Feeling and Emotion

In a complex organism such as a person, energy is active at multiple levels, from body cells and intra-cellular elements, to anisotropic Feeling in brain, beginning with instinctual drives and satisfactions and leading to conceptual-feelings and intentional aims. Moods are manifestations of non-local Feeling.[17] Some theorists postulate unconscious emotions and/or conflicts between moods, but how can such emotions be identified if they are unconscious other than as states of tension, anxiety, uncertainty, hesitation, or an obstructed inclination in some direction. Unconscious emotion may be like spontaneous or automatic action that must occur before its value can be identified. Courage might be an example of an unconscious impulse or emotion. Affect tends to be diminished in rational concepts. In external objects, it is often imperceptible or seems projected by the observer.

Feeling sequesters at the posterior limit of the mental state. Concentration at the onset enhances Feeling in drive and primal Will, which dissipates in affect-neutral objects. The more intense a desire, the closer it is to drive; the less intense, the closer it is to objects. The transition from a concentration at the unconscious core to affect-neutral diversity at the conscious surface reflects the allocation of Feeling from unity to multiplicity. Feeling is not diminished but is divided and muted by allocation as it is colored by diversity.

Desire is intermediate in the derivation, relinquished at the outer limit of the intentional aim. In the partition to desire, sexual and the hunger-drive moderate when satiety is achieved. Intensity in the pre-object is emptied in its progression from inception to termination, in the succession of phases, in the "derailments" and distillations. Like a torrent that loses force in tributaries, the mental landscape is enriched in rational thinking as affect is impoverished. Conceptual-feelings or affect-ideas nourished by a common source become the dry shoals of conscious particulars. A wave of Feeling shaped to a configuration by neuronal populations that correspond to instinctual drive is analogous to a mental category. Brain-activity is the dynamic of cell populations; mental-activity is their conceptual frame. A category segments a traveling wave.

Let me close with a word on the deep relation of category and Feeling. In one sense, Feeling as energy or process is aligned with brain activity, while a cate-

[17]The idea of emotion as a secondary interpretation of bodily changes, or as originating in brain probably needs re-thinking in light of this discussion. Energy and directional Feeling inhabit the body in liver cells as well as neurons, so that an energic theory of mind/brain is continuous with bodily feeling.

gory is a mental construct, so the relation of the two is equivalent to that of mind to brain. I have argued that process is a traveling wave of whole–part shifts constrained by unconscious patterns of the prior state, of beliefs, values, and habits, with the resultant configuration sculpted by sensation to model the external world. If the process can be depicted as a series of nested whole–part shifts, what is the nature and origin of a category and how might it correspond to whole–part process in brain? If the whole–part transition is central, and if wholes can be said to correspond to categories and parts to members, the whole-to-part transition would map to the elicitation of acts, ideas, emotions and objects. That is brain process and the mental state can be described in terms of whole/part, context/item or fractal-like transformations.[18]

A related question concerns the relation of a category to its members, especially with regard to Feeling. A category is an abstract "structure" that encloses a set of shared, tacit, or implicit possibilities. A category member that becomes conscious or explicit aborts the category for the particular, which then becomes a sub-category for a range of potential members. For example, once *dog* is elicited from the category *animals*, the prior category dissolves and the new one (dogs) appears. The antecedent category is replaced by the consequent one (dog), which includes canines. Process is forward-looking. However, any content can be assigned to some background category; one can withdraw to earlier, deeper phases, but in mind-active, items forecast subordinate members, not super-ordinate categories.

Feeling is an ineffable pattern of vibratory activity in the brain that satisfies this description, for it remains indefinite until the category resolves. Once an idea, emotion, or object clarifies, the antecedent category is left behind. The potential for unrealized particulars now belongs to the category of the elicited content. Desire has a multitude of possible objects that narrow down in love, then partition to affection, friendship, compassion, and so on. When the reverse occurs, e.g., going from interest or friendship to love, the replacing state revives the antecedent category to which feeling (interest) is subordinate. All concepts and attendant affects follow this pattern. With the implementation of a drive, potential members of the category are eliminated.[19] Similarly, the category of desire is forfeit when its object resolves, i.e., the potential for an object is abandoned when one impulse clarifies. Conversely, desire that begins with interest transitions to what is prior in the mental state. Since drive and desire are aroused in every mental state, the passage from interest to love replaces the

[18]Other studies that refer to ground/figure, surround/center, frame/content and so on, seem to be groping toward the same description, as well as more generalized accounts of individuation, specification or differentiation though none of these accounts has mapped cognition to process in relation to a concept of the mind/brain state.

[19]Displacement of drive is well-described in the ethological literature. With blockage in drive-expression, a return to core potential elicits substitute behaviors.

more superficial (object-close) category of interest with the deeper (self-close) category of love.

References

Atmanspacher, H. (2012), Dual-aspect monism à la Pauli and Jung. *Journal of Consciousness Studies, 1,* 96–120.

Atmanspacher, H., and Filk, T., (2011). Contra classical causality. *Journal of Consciousness Studies, 19,* 95–116.

Bergson, H. (1923), *Durée et simultanéité* (second edition). Felix Alcan: Paris.

Birch, C. (1990). Value and meaning of process thought. *Process Studies, 19,* 224–234.

Bosanquet, B. (1904). *Psychology of the moral self*. London: Macmillan.

Bradley, F. (1893). *Appearance and reality*. Oxford: Oxford University Press.

Bradley, J. (1985). The critique of pure feeling. *Process Studies, 14*, 253–264.

Brown, J.W. (1986). *Life of the mind*. Hillsdale, New Jersey: Erlbaum.

Brown, J.W. (1991). *Self and process*. New York: Springer Verlag.

Brown, J.W. (1994). Morphogenesis and mental process. *Development and Psychopathology, 6,* 551–563.

Brown, J.W. (1999). Microgenesis and Buddhism: The concept of momentariness. *Philosophy East and West, 49,* 261–277.

Brown, J.W. (2004). The illusory and the real. *Mind and Matter, 2,* 37–60.

Brown, J.W. (2005). *Process and the authentic life*. Frankfurt: Ontos Verlag.

Brown, J.W. (2010). Simultaneity and serial order. *Journal of Consciousness Studies, 17,* 7–40.

Brown, J.W. (2011). The relation of embryology to linguistic and cognitive process. *Journal of Psycholinguistic Research, 40,* 89–194.

Brown, J.W. (2012a). *Love and other emotions*. London: Karnac.

Brown, J.W. (2012b). What is consciousness? *Process Studies, 21,* 54–73.

Brown, J.W. (2012c). On thinking. *Mind and Matter, 10,* 47–73.

Cobb, J. (2008). *Whitehead word book*. Claremont, California: P&F Press.

Dart, R. (1956). The relationships of brain size and brain pattern to human status. *South African Journal of Medical Science. 21,* 23–45.

Dombrowski, D. (2001). The replaceability argument. *Process Studies, 30,* 22–35.

Goodwin, B. (1984). Changing from an evolutionary to a generative paradigm in biology. In J. Pollard (Ed.), *Evolutionary theory: Paths into the future* (pp. 99–120). New York. Wiley.

Gunter, P. (1999). Bergson, mathematics and creativity. *Process Studies, 28*, 268–288.

Hunt, H. (2001). Some perils of quantum consciousness. *Journal of Consciousness Studies, 8*, 35–46.

James, W. (1912). *Essays in radical empiricism*. New York: Longmans, Green and Co.

Langer, S. (1942. *Philosophy in a new key*. Cambridge, Massachusetts:: Harvard University Press.

Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley.

McTaggart, J.M. (1901). *Studies in Hegelian cosmology*. Cambridge: Cambridge University Press.

Perry, R. (1926). *General theory of value*. New York: Longmans, Green and Co.

Pribram, K. (1991). *Brain and perception*. Hillsdale, New Jersey: Erlbaum.

Rapaport, D. (1950). *Emotions and memory*. New York: International Universities Press.

Romijn, H. (2002). Are virtual photons the elementary carriers of consciousness? *Journal of Consciousness Studies, 9,* 61–81.

Rusu, B. (2013). Feeling. From British psychology to Whitehead's metaphysics. Paper presented at the conference on *Process Approaches to Consciousness*. Fontareches, France.

Schilder, P. (1951). On the development of thoughts. In D. Rapaport (Ed.), *Organization and pathology of thought* (pp. 497–519). New York: Columbia University Press.

Schilder, P. (1964). *Contributions to developmental neuropsychiatry*. New York: International Universities Press.

Shields, G. (2009). Panexperientialism, quantum theory, and neuroplasticity. In M. Weber and A. Weekes (Eds.), *Process approaches to consciousness in psychology, neuroscience, and philosophy of mind* (pp. 235–259). Albany: SUNY Press.

Sprigge, T. (2005). *The God of metaphysics*. Oxford: Oxford University Press.
Stenner, P. (2008). A.N. Whitehead and subjectivity. *Subjectivity, 22,* 90–109.
Ward, J. (1920). *Psychological principles*. Cambridge: Cambridge University Press.
Weber, M., and Weekes, A. (2009) *Process approaches to consciousness in psychology, neuroscience, and philosophy of mind*. Albany: SUNY Press.
Whitehead, A.N. (1933). *Adventures of ideas*. Cambridge: Cambridge University Press.

The Journal of Mind and Behavior
Winter and Spring 2014, Volume 35, Numbers 1 and 2
Pages 21–50
ISSN 0271–0137

ADHD as Emergent Institutional Exploitation

Lincoln Stoller

Mind Strength Balance, Inc.

Attention Deficit/Hyperactivity Disorder (ADHD) is examined in terms of the systems that define it and as a structure that creates the world around it. Considering ADHD as an aspect of the whole environment allows the assembly of partial and conflicting views to create a single, multi-faceted picture. The ADHD label is shown to be an emergent property that manifests the failure of the social, economic, therapeutic, and political parts of our culture. This approach provides a theoretical basis on which to analyze the diagnosis's evolutionary path and to make predictions about its future.

Keywords: ADHD, institutions, systems theory, education, emergent structure

Complex systems change as a multitude of interconnected agents create new rules from old rules. From this process emerges unpredictable new structures that may complement or consume existing structures: evolution is inherently destabilizing. The same process underlies changes in attitude and social behavior. In particular, our notions of health and normality change over time and, as in any evolutionary system, new attitudes are often antithetical to old ones. Such is the strange case of Attention Deficit/Hyperactivity Disorder (ADHD), a diagnosis of dysfunction now assigned to 20 percent of adolescent males.

I will show the ADHD diagnosis does not follow previously existing norms of health care because it does not aim to improve an individual's own sense of well-being. Instead, ADHD represents an institutional exploitation of children for the benefit of institutions. This diagnosis generates greater rewards and fewer penalties for powerful social interests than other ways of organizing people. It is a structure that has emerged from our complex, self-organizing society. It has no independent biological reality, and requires none. In order to understand how this diagnosis has come about, and how it will evolve in the future, we need to study the institutions that sustain it and their agents.

Correspondence concerning this article should be addressed to Lincoln Stoller, Ph.D., 148 Dubois Road, Shokan, New York 12481. Email: LS@mindstrengthbalance.com

Controversy has surrounded ADD/ADHD since it was added to the third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM–III) by committee vote in 1980 to codify the symptoms of a behavior for which there was no known cause or biological indicator. A large number of interested institutions – schools, corporations, unions, professional organizations, health-care providers, and government organizations — were involved in a disorder that had different implications for each. A tapestry of unacknowledged special interests, commercial advantages, historical relationships, and cultural paradigms held together an otherwise implausible diagnosis. People who get labeled ADHD are a diverse group who actually suffer from a spectrum of problems whose remediation would require separate medical, psychological, cultural, and political solutions. But ADHD has emerged as a diagnosis precisely because it provides a greater remuneration to special interests than attempting to address the diverse problems from which the ADHD population suffers.

Method

This article adopts a systems approach to elucidate the forces behind the ADHD diagnosis. The purpose is to understand the costs, benefits, and impacts of one subsystem on another and to determine what interests are being served through the diagnosis and treatment of ADHD.

Four constituencies that define, support, or benefit from the diagnosis – psychiatry, education, pharmaceutical manufacturers, and parents — are considered along with the evidence that ADHD is a dysfunction of biological origin. Scientific, financial, and political relationships exist within and between these constituencies that lie outside public scrutiny and the consideration of specialists. These relationships develop without interference because of our culture's deference to authority and to institutions that control the social dialog. This dialog depends upon the functional unity of a common understanding and generates predictable relationships in accordance with the widely accepted tenets of biological psychology and other modern "mythologies."

ADHD is not an individual dysfunction, but a societal issue involving organizations that compete to profit from school-age children. Each constituency benefits from the position that ADHD be treated as an incurable biological dysfunction in spite of much evidence that no physical difference reliably distinguishes individuals assigned this label from a population of healthy individuals. Using the notion of emergent properties in feedback systems, I consider ADHD as a construct created for the benefit of these systems and the organizations from which they are built. I conclude that the diagnosis of ADHD will evolve in whatever manner is maximally profitable in this multi-systems context.

Attention Deficit Hyperactive Disorder

The DSM, written and published by the American Psychiatric Association (American Psychiatric Association, 2000), is considered the authoritative source in diagnosing mental illness. The DSM is used by therapists as the basis of treatment, by pharmaceutical companies as a description of symptoms needing remediation, by the courts as a basis of culpability, by government and social institutions to identify abnormal behavior, and by insurance companies as a basis for compensation.

The DSM is designed to provide an understandable, repeatable, and uniformly applicable description of behavior that is currently recognized as a disorder. It does not argue why these particular behaviors should be taken as indicative of a disorder. It does not explain why various disorders have been added, removed, further resolved, or dropped from mention altogether in subsequent editions. The DSM assigns diagnostic labels to behavior but should not substitute for understanding patients as people (Carlat, 2010, p. 62).

The DSM is the standard text used by mental health professionals for defining, labeling, and treating so called mental illnesses. It forms the nexus between what therapists treat, pharmaceutical companies research, doctors prescribe, insurance reimburse, the law allows, and patients are expected to accept. The DSM is a document that caters to the needs of many institutions and social structures. It is not a scientific document.

The diagnosis of ADHD is extreme in the generality of its symptoms. Each of its supposedly defining characteristics is separately or jointly present to varying degrees in normal behavior. Consequently, there is a lack of specific criteria for diagnosis (Conrad, 2006, p. 54). For example, the DSM suggests that a person's lack of focus and attention are indicators of possible dysfunction but these qualities are not aberrant in themselves and only symptomatic in certain conditions, when observed repeatedly over extended times and in different situations. The DSM describes ADHD entirely in terms of the subjective observation of a person's behavior in social situations, as judged by teachers and other people in positions of authority with whom the person may be in casual contact and whose connections to that person are largely lacking in psychological depth and intimacy. The opinion of the person who is a candidate for the diagnosis plays little to no role in the diagnosis.

All the symptoms of ADHD exist in varying degrees in normal behavior. The diagnostic criteria rest on subjective notions of deference to authority and behavior that these authorities consider appropriate in the context of social activities in an often hostile school environment. In addition to being potentially attention-disabled, people who behave in manners considered inappropriate may be angry, bored, depressed, or frustrated. The ADHD diagnostic criteria do not distinguish among these causes of what is being deemed inappropriate behavior.

Agreement among practitioners regarding the validity of a diagnosis does not confer validity to a disorder, or establish that what is being identified as a disorder is a disorder. That is to say, even if authorities unanimously agree that you have a disorder, this provides no evidence that such a disorder really exists. This does not mean that people labeled with ADHD are normal; rather, it is the incontestable observation that symptoms never explain anything, and that things do not exist just because people say so.

Institutions

Psychiatry

Psychiatry is a socially driven practice that identifies and treats symptoms of aberrant behavior, where aberrance is defined relative to social norms. Predominant current opinion, as has been the opinion at various times in the past, is that mental illnesses originate from biological causes that can be cured through chemical, electrical, or surgical means. These claims have been repeatedly vitiated (Kendler and First, 2010; Valenstein, 1988, p. 3). In addition, it is established that rather than normalizing brain chemistry, psychiatric drugs tend to cause brain damage and may induce pathology (Breggin, 2008; Whitaker, 2005).

None of the major psychiatric ailments has been traced to biological etiologies, although there have been many attempts — and biological factors such as developmental and physical trauma remain possible causes. The reason for this partly lies in the reductive models that propose simple biological causes for complex mental conditions. Biology certainly plays a role in one's mental condition, but mental illnesses do not have simple structural causes.

The 1951 Durham–Humphrey Amendment explicitly distinguished prescription from over-the-counter medication and accorded doctors the exclusive privilege of writing prescriptions. Doctors' salaries doubled, and income from pharmaceutical advertising in journals published by the American Medical Association went up tenfold. A doctor's prescription was then required to obtain what are now heavily marketed pharmaceuticals.

Select members of the American Psychiatric Association write — and the organization publishes and carefully guards copyrights to — the DSM, from which it garners $5,000,000 per year in sales, or 1/6 of the total annual income that the APA relies on to fund its operations. The DSM's classification of illnesses has financial impact for both pharmaceutical companies and association members. Nearly 70 percent of the members of the task force charged with assembling the next version of the Diagnostic and Statistical Manual, the DSM–5, have financial relationships with pharmaceutical companies, up from 57 percent for the previous version (Cosgrove, Krimsky, Vijayaraghavan, and Schneider, 2006; PLoS Medicine Editors, 2012).

As a discipline, psychiatry is based on observation, but it lacks objectivity and a definitive test for its theories. Guided by consensus, its norms develop in whatever direction is advocated most loudly and practiced most widely. Psychiatry's mutable methods and conclusions are prone to oversimplification and political influence.

Political and financial interests pressure the field of psychiatry to support their objectives and foster the expectation that psychiatry will help resolve social or behavioral issues. Financial reward, political power, and institutional security are some of the rewards offered to the field of psychiatry, and to psychiatrists themselves, for developing and applying their expertise in the aid of these institutions (see Carey and Harris, 2008). In this way the field is offered positive incentives for defining and diagnosing ADHD for the benefit of other institutions. Psychiatry has scientific pretensions that conflict with its subordination to corporate interests.

Compulsory Education

Trends and current practice in compulsory education provide support for, or benefit from, the notion that ADHD is an individual's dysfunction. Many leaders in the formation of compulsory education policy have advocated the molding of student behavior and the shaping of intellectual dialog in accordance with political objectives. In this context part of compulsory education's objective is to identify children who lie outside this norm, to label them as different, and to remediate their behaviors and attitudes as part of the educational agenda. Educational attitudes of this kind are consistent with the creation of ADHD as a diagnostic category and the treatment of children assigned to this category.

From its inception, a primary goal of compulsory schooling has been socialization, with "education" defined in terms of its effect on society and its success in training students to play a useful economic role. Writing in 1915 John Dewey was a leading voice in the design of compulsory education. He considered the student to be possessed of innate capacities that needed to be shaped and modeled. A school's role in the development of children was "not to leave them alone to follow their own 'spontaneous development,' but to provide an environment which shall organize them" (Dewey, 1916, p. 134). These ideas were consonant with the laboratory schools Dewey developed, the subsequent development of behaviorist educational psychology, and today's standardized testing.

Dewey exerted and continues to exert a great influence on education policy in the United States. He called for a curriculum that developed students in accordance with the moral and intellectual needs of the society. In his later writings Dewey empahsized that education's main goal was social reform and, according to Ravitch (2000, p. 203), Dewey was frequently "skeptical of . . . excessive concern for individualism and spontaneity." He lauded the removal

of family influence, which Dewey (1984, see pp. 229–231) considered "a breeder of non-social interests."

Compulsory education has been widely used in Western European history as a tool of denationalization and assimilation according to historian and economist Ludwig Von Mises (1944, p. 82), and has been applied in numerous cases in United States history as a means of national, economic, or cultural disenfranchisement. Riots in 1917 reversed New York City's "Gary Plan" to apply industrial models to the education of immigrent children (Gatto, 2006, p. 187). Southern states created segregated schools that perpetuated the economic disparity between Blacks and Whites until the Supreme Court's 1954 case, Brown v. Board of Education, ruled segregation unconsititutional. A series of United States congressional acts between 1877 and the late 1960s forced the relocation of Native American children to off-reservation boarding schools where they were stripped of their cultural identity (Curcio, 2006; Stahl, 1979). From its origins in Prussian social engineering, compulsory schooling has been a battle ground between institutional control and personal freedom.

From 1981 to 2002, the weekly time spent doing homework by six- to eight-year-old children had tripled, despite the lack of any demonstrable benefit (Cohen, 2006; Cooper, Robinson, and Patall, 2006). According to the 2006 National Sleep Foundation survey, 45 percent of adolescents get an insufficient amount of sleep, and 31 percent get a borderline amount of sleep on school nights (National Sleep Foundation, 2006). A 2009 study (Gau and Chiang, 2009) reported a strong association between ADHD symptoms and sleep problems, and suggested that adolescents be screened for sleep problems before being assigned a diagnosis of ADHD.

The right of children to be free of institutional supervision and control was once considered a medical necessity.

> In the early 1900's doctors led a movement to abolish [homework], insisting that children needed at least 5 hours of fresh air and sunshine each day. At that time, those kids who would be diagnosed with attention-deficit/hyperactivity disorder were told to go outside and play more — not take medication so that they could sit still. (Bennett and Kalish, 2006, p. 35)

In spite of these studies correlating environmental influences with symptoms of ADHD (Biederman et al., 1995; Carlson, Jacobvitz, and Sroufe, 1995) we continue to be told that the explosion in the number of children diagnosed with ADHD is due to an undiscovered biological pathology.

At the beginning of the twentieth century Edward Thorndike helped start the field of educational psychology by defining learning as the process by which animals repeat ever more efficiently and economically those actions for which they are rewarded. His laws of learning underpin current notions of operant conditioning, and are a basis for segregation in the teaching of different subjects

(Horn, 2007, p. 227). Thorndike's work to undermine the teaching of general intellectual skills — work that was inconclusive at the time (Hofstadter, 1962, p. 349) and since overturned (Breuner, 1977, p. 6) — coincided with the burgeoning government interest in behaviorism and scientific management. Educational psychology, adopted at teachers colleges, became a science, and schools became their laboratories.

> As schooling encroaches further and further into family and personal life, monopolizing the development of mind and character, children become human resources at the disposal of whatever form of government is dominant at the moment. This confers a huge advantage on the leadership of the moment, allowing it to successfully reproduce itself, foreclosing the strength of its competitors. (Gatto, 2006, p. 359)

The training of teachers, which previously focused on teaching content, was refocused on educational psychology as "central to the teaching enterprise and to the prepartion of teachers" (Peterson, Clark, and Dickson, 1990). In 1973 the United States Department of Education commissioned the Rand Corporation to create a seven volume study on how schools could be better used to foster behavioral modification (Berman and McLaughlin, 1974; Eakman, 1991, p. 118). The training of teachers as "change agents" started in the 1960s and this training advocated behavioral control, managed conflict and resistance, and the testing of students' values and obedience to authority. The role of teachers as change agents remains a current research topic (Lu and Ortlieg, 2009).

Hierarchical institutions, like the federal government and the Department of Education, are always concerned with the identification and training of people whose task is to catalyze, instigate, aid, and nourish change or to prevent it (Havelock, 1973, p. 7). It is necessary to recognize that in the mix of individual, curricular, community, labor, and government dynamics almost everything that transpires in the classroom, from the choice of textbooks to the protocol for speaking out during class, is either manifestly political or has political ramifications.

The change agent is admonished to "be a familiar object to the client in ways that are not important to his mission," and to "identify some common interests which are far removed from any change project . . ." (Havelock, 1973, p. 54). While educators must have some skill in the art of persuasion, when taken to an extreme the exploitation of familiarity is legally recognized as "affinity fraud." This is troubling in the context of education because those involved — students, parents and communities — may not know who is acting on what agency's instructions.

In 1990 Anita Hoge prevailed in actions conducted within the Department of Education accusing the United States Federal government of: (a) amassing personal, psychological profiles fraudulently passed off as academic achievement tests, (b) approving curricula to remediate incorrect attitudes, and (c) subsidizing a policy of practicing medicine without a license (see Eakman, 1991, p. xi).

This established that government control exists at the highest levels of public education for the purpose of promulgating social values pertaining to deference to law, authority, and community norms. The collection and cross-tabulation of attitudinal data are now widespread (Eakman, 2007, p. 201).

The aims of compulsory education are strongly influenced by institutions whose aim is to control individuals, rather than to empower individuals to control institutions. The diagnosis of ADHD serves many of this system's goals while providing rewards to the institutional elements that operate within it.

The Pharmaceutical Industry

It is estimated to cost between $400 million to $2 billion dollars to develop and bring a new drug to market (DiMasi, Hanse, and Grabowski, 2003; Masia, 2008, p. 82). Global sales of the ADHD drugs Ritalin, Adderal, Concerta and similar generics, were estimated to be $2.8 billion in 2003, with 85 percent of these sales in the United States (Scheffler, Hinshaw, Modrek, and Levine, 2007). Based on the 3.4 percent annual growth rate of ADHD medication, according to 2008 figures (Nauer, 2009), we can estimate that 2012 sales were over $3.6 billion.

The management of pharmaceutical companies, like the management of all corporations, maximizes profits with little regard to social cost. Not only is this expected, it is effectively required by law as public corporations are obliged to satisfy the interests of their shareholders first. Executives are professionally responsible to pursue any action that will generate legal profits.

Prolonged use of amphetamines or Ritalin can create neurochemical imbalances (Higgins, 2009), stunt growth (Swanson et al., 2007), result in chromosomal changes (El-Zein et al., 2005), and sometimes lead to substance addiction. Overdoses can cause liver, kidney, and heart damage (Greene, Kerr, and Braitberg, 2008). The manufacture's imperative is to ensure that side effects do not become financial liabilities. This can be effected by such means as supporting research that contradicts alleged adverse reactions or that creates difficulty in interpreting the scope or identification of adverse reactions, by withholding the publication of research that demonstrates adverse side effects, and by using advertising leverage or corporate affiliations to discourage journalistic exploration, discussion, or release of information that might imply a correlation between drug use and adverse reactions. Other legal protections include limiting liability, using confidentiality to withhold public disclosure, and offering settlements in exchange for denial of culpability and sealing of court proceedings.

Pharmaceuticals companies spend billions of dollars in research to fund organizations, conferences, and educational programs to extol their product. They work for the election of politicians and the appointment of authorities who will support their strategy. An analysis of the United States pharmaceutical industry concluded that the industry spent almost twice as much on promotion than on research and development (Gagnon and Lexchin, 2008).

Twenty five percent of all doctors in the United States received drug company money for helping to market drugs in 2004 (Campbell et al., 2007). Arguing in the British Medical Journal, psychiatrist Giovanni Fava (2008) says this money is dispensed by pharmaceutical companies for the purpose of getting "as close as possible to universal prescribing of a drug by manipulating evidence and withholding data" (p. 1405).

Financial bias in research is pervasive and has been widely noted (Pachter, Fox, Zimbardo, and Antonuccio, 2008; Sen and Prabhu, 2012). Financial incentives are paid to universities that invest their endowments in pharmaceutical companies, and these universities rely upon pharmaceutical companies for research grants. Not-for-profit organizations advocating the use of pharmaceuticals are subsidized by pharmaceutical companies (Herxheimer, 2003). These organizations include the National Alliance for the Mentally Ill (Harris, 2009) and Children and Adults with Attention Deficit/Hyperactivity Disorder (Eberstadt, 1999). News in the mainstream press is seeded by pharmaceutical companies who freely disseminate research that endorses pharmaceutical use. Ninety percent of the authors of three major psychiatric clinical practice guides had undisclosed financial ties to companies that manufacture drugs identified or recommended as therapies for the respective mental illnesses (Cosgrove et al., 2009).

The pharmaceutical industry funds most of the pharmaceutical research and, since the 1992 passage of the Prescription Drug User Fee Act, all of the FDA's costs for approving and licensing drugs. Most researchers are invested in or directly paid by the industry, and most professional and public educational material is produced and paid for by the industry. There is virtually no support for independent voices or unbiased opinion. Scientific standards provide little countervailing influence because the standards are subverted by financial incentives to willing researchers. Government regulation is limited by scarce resources and the high cost of enforcement, and further undermined by the government and the industry's revolving door policy of placing agents of industry in government positions, and then hiring agents of government to fill industry positions. Pharmaceutical research is directed so as to maximize profit, and ADHD provides a highly profitable market.

Parents

Studies show a connection between a parent's mental state and the child's ADHD, with parental stress increasing in proportion to the problems caused by the child's condition (Harrison and Sofronoff, 2002). But the stress of parents of ADHD children is not simply due to their childrens' problems. "ADD children are far more likely than other children to have parents who have suffered major depression, about 30 percent compared to 6 percent" (Maté, 1999, p. 104).

Doctors warn parents that without treatment the long-term outcome for children with ADHD is poor (Mannuzza and Klein, 2000). According to Newton–

Howes, "possibly the most disabling aspect of ADHD in adulthood is the disruption it causes in interpersonal relationships, with increased risk of chronic conflict with work peers, socially inappropriate behaviours, disputes with partners and spouses and trouble with the law" (2004, p. 533). Marketing by pharmaceutical companies, as well as by the educational and medical establishments, targets these vulnerable parents by citing research perpetuating the assumption of ADHD's biological origin.

Gabor Maté describes parents of children with ADHD this way:

> The erosion of community, the breakdown of the extended family, the pressures on marriage relationships, the harried lives of nuclear families still intact and the growing sense of insecurity even in the midst of relative wealth have all combined to create an emotional milieu in which calm, attuned parenting is becoming alarmingly difficult. (1999, p. 109)

> No matter whom they had consulted, not one of the couples I have seen in my practice had ever before been encouraged to look closely at how their emotions, lives and marriages might affect their children . . . . It seems to them just normal human existence to live at a hectic pace and in tense relationships, nerves stretched taught as piano wires. Sensitive children, as all children with ADD are, will be particularly affected. (1999, p. 96)

While not usually thought of as an interest group, parents are consumers of ADHD treatment in the same way that society is the consumer of law enforcement. It is the parents who pay for treatment, and without parental acceptance of the diagnosis, few children would be so labeled. Because parental acceptance is key to the acceptance of the diagnosis we can ask what parents have to gain.

In a society composed either of families of single parents, or families in which both parents work and where in-laws are not available for child care, many parents believe that sending their children to school is an economic necessity. The education industry insists, and parents generally believe, that parents cannot teach their own children and that children cannot teach themselves. The growing home and democratic school movements disprove these assertions (Miller, 2002). The United States school-age population has grown approximately 2 percent in the last several years and this represents a number of children equal to the number that are leaving institutional schools in order to be home-schooled. Considered as a whole, the home-schooled population now constitutes roughly 3 percent of the school-age population (Ray, 2011).

The ADHD diagnosis is offered as a solution to an existing problem. Given that a problem exists, what other solution do parents have? For children not disabled in any other regard, at least not according to authorities, there may be no other explanation aside from the diagnosis of ADHD. Most parents do not have the resources or the self-confidence to challenge the educational, psychological, pharmacological, social service, and medical establishments. They have nowhere else to go for advice and direction.

Once parents have accepted the diagnosis, and the recommended treatment has solved their child's behavior problem — at least according to the accepted

criteria — there is little incentive for parents to reject the diagnosis even if urged to do so by their children. The behavioral modification to accompany children's psychiatric drug-induced compliance — as commonly recommended by psychologists, teachers, and parent groups — then ensures that both the parents and children comply with teachers' wishes (American Academy of Pediatrics, 2002). This is made clear by the growing trend in which parents seek ADHD medication to enhance their children's performance without regard to whether the children have any disorder at all (Schwarz, 2012). This demonstrates that some parents view the ADHD treatment as conferring scholastic benefits that they desire for their children, and that the treatment does not just restore normal performance to a subgroup of individuals.

Parents who have accepted that their children have ADHD are under pressure to defend the label. Rejecting the diagnosis not only means they erred in accepting it initially and that they now have to find some other solution, but also that their behavior as parents may have contributed to the condition. This is because the alternative to the biological explanation of ADHD is a developmental etiology in which school and family environments contribute to the development of the child's dysfunctional behavior. It is not surprising that parent support groups, such as CHADD (Children and Adults with Attention Deficit/Hyperactivity Disorder), defend the validity of the diagnosis; but the existence of these groups does not constitute evidence that a medical condition exists.

Medication-based Reasoning

If treating the child as if he has an illness generates family-wide rewards, then there is an incentive for parents to believe their child has an illness. The logic is this: the treatment assumes the existence of a disease. Therefore if the treatment is successful, it is taken as proof that a disease exists. According to this logic if a person is given an anti-depressant and improves, this is taken as proof that he suffers from biological depression. This is erroneous reasoning as there may be environmental factors causing the depression, even though a drug might elevate the child's mood.

This reasoning, termed *ex juvantibus* from the Latin phrase meaning "from that which helps," is the classic fallacy of seeing one thing as causing another when only a correlation exists. There are many examples in medicine where there is no direct connection between the treatment of a condition and the condition's cause (Valenstein, 1998, p. 133). The point is that the various rewards of treatment lead parents to support the diagnosis regardless of its validity. In this situation the ADHD diagnosis generates positive feedback from parents as a result of how the treatment affects them and the behavior of their children, and not because it resolves any biological problem.

ADHD Childen as Individuals

Much research explores the cause of ADHD in order to find or to advocate a solution, but little is done to explore how the available solutions shape our understanding of the issue: if the preponderance of research is done by neurobiologists or pharmacologists, then the preponderance of explanations are biological and chemical. Various factions — researchers, therapists, politicians, teachers, and parents — advocate for particular solutions and have a vested interest that leads them to argue for particular causes. The factions are not united in their efforts to understand one condition; they are divided by their interests in justifying different programs. One faction aims to control individuals in ways that strengthen institutions, while another aims to refashion institutions in ways that strengthen individuals by making them more confident, independent, and able to employ their own resources. The strongest of these groups determines public attitudes toward the treatment of ADHD. The nature and scope of each group's strength is not as important as recognizing that the whole system develops toward the configuration that provides the greatest benefits to the most powerful players.

The Existence of a Disorder

No causative agent has been found for ADHD, and there exists no objective criteria to diagnose the condition. The definition of ADHD as a loose collection of subjective assessments precludes a scientific basis for its definition. The fact that a large number of practitioners believe that the subjectively evaluated symptoms of ADHD imply that ADHD is a real disease calls into question what psychiatrists and psychologists mean by a disorder. I can only conclude that for these practitioners a condition is a disease if it is effectively treatable to some degree.

Biological reductionists look for correlations between ADHD behavior and brain chemistry. This approach only requires the study of a population of those who "have it." These researchers have explored various possible correlations over the last 35 years but have found no correlation sufficiently accurate or consistent that it can be used for reliable diagnosis (Leo and Cohen, 2003). No biological cause has yet been found, and the arguments for the inheritance of ADHD are flawed (Joseph, 2000).

A paper by Hanneke van Ewijk and colleagues illustrates this kind of reductionist research (van Ewijk, Heslenfeld, Zwiers, Buitelaar, and Oosterlaan, 2012). van Ewijk et al. performed a meta-analysis of previous research in tensor imaging in an attempt to find support for the thesis that tensor imaging can discriminate between controls and people with ADHD. In this they are compiling research that was originally conducted for purposes other than exploring ADHD, such

as the ability of tensor imaging in yielding information for a given population. Tensor imaging is a new means of brain imaging that measures small changes in the acceleration of water at different locations within the brain.

This paper engaged in a number of common fallacies. First, van Ewijk et al. assume that the ADHD participants are biologically abnormal, which is what they are trying to establish. Second, they mine the data to uncover any differences in observations taken from different groups — and there are always differences to be found somewhere — and imply the differences are biologically and etiologically significant. This is a logical error because it fails to weigh alternatives or recognize the existence of what are known in statistics as confounds. Confounds are correlations between events that are only related because of a incidental commonality. An example of this is the correlation between the eating of ice cream and the frequency of drowning. Ice cream consumption and drowning are related because they both occur on hot summer days, and not because one causes the other.

Third, they mention but do not integrate into their conclusions the idea that these differences have no causal relationship to ADHD. And fourth, the authors overlook fundamental uncertainties in their own observations. For example, the data provide a measure of fluid movement within the brain and describe a greater amount of fluid movement in ADHD subjects. This may be due to the greater movement of the subjects themselves and not to the fluids within them. As the authors themselves point out, the fuzzier images gathered from ADHD subjects could simply result from the fact that ADHD subjects are restless and their movement is causing the images to be out of focus. In spite of recognizing this limitation the authors do not control for it. This does not undermine the usefulness of this work for specialists in the field of brain imaging, as discussions of this nature are critical to the development of a better understanding of the strengths and limitations of using imaging technology for diagnostic purposes, but it does illustrate how psychologists can overlook the limitations inherent in brain imaging and interpret results of this kind as confirmation of the neurological origin of ADHD.

Genetics

A similar confusion surrounds the argument for the genetic origin of ADHD. The argument is that since ADHD runs in families it is inherited, therefore genetic, and therefore of biological origin. This reasoning rests on logical fallacies and scientific misunderstandings that lead to the common statement that 80 percent of the disorder is caused by genetic factors. Evidence of the near uselessness of using genetics as the foundation for a biological model of ADHD is revealed when those who present it as such qualify that "these are not 'dominant' genes but rather 'susceptibility' genes, which may interact with one another and with

a child's environment to create the potential for ADHD . . . . A child's environment may powerfully influence even strongly heritable traits" (DeGrandpre and Hinshaw, 2000).

Consider a similar case in which 80 percent of one's height is said to be genetically determined. This roughly means that by surveying the heights of a person's ancestors one can predict a person's height to 80 percent accuracy. In contrast, as will be discussed, Cummings and Wiggins showed that patient and parent counseling resolved nearly 85 percent of cases diagnosed as ADHD. Because biological conditions cannot be resolved by counseling, this implies that only the 15 percent of the population they considered, whose diagnosis was not reversed, might be afflicted by a biological condition. This means that at most 15 percent of people originally diagnosed with ADHD might actually have an inherited predisposition to the condition. In summary, dominant genetic factors account for 85 percent of height, while genetic susceptibility may play a role in 15 percent of the population labeled ADHD. Dominant genetic influences can explain something, but genetic susceptibilities do not provide a cause. Genetic susceptibilities imply a unresolved risk factor in which environmental forces play a crucial role.

In a 2010 study heralded as being the first to find direct evidence that ADHD is a genetic disorder (Williams et al., 2010), Thapar, who was one of the authors, is cited in a press release preceding publication of the article as saying: "Now we can say with confidence that ADHD is a genetic disease and that the brains of children with this condition develop differently to those of other children" (Walsh, 2010). This is a misleading statement that implies a genetic determinism that does not exist, stigmatizes a group of individuals who are labeled ADHD, and instills terror in the parents of these children. Furthermore Thapar implies that those who differ from the norm are necessarily inferior. She later clarified this statement by saying that "she was not asserting that genes alone were responsible for ADHD but rather a complex mix of genes and environmental factors" (Walsh, 2010).

What the study showed was that 85 percent of those labeled as having ADHD had no discernible genetic difference from those without ADHD at a 95 percent confidence level. That is to say, the basis for asserting that genetics plays a role in the etiology of ADHD is the observation that genetics may play a role for 15 percent of those with ADHD. The study did not control for differences in IQ. By removing from this study participants with impaired IQs (below 70), the number of individuals *failing* to display significant genetic differences rises to 89 percent. In a subsequent paper, two of the authors (Stergiakouli and Thapar, 2010, p. 557) clarified that "gene variants still explain only a small percentage of the inherited component of ADHD," by which they are referring to epigenetic tendencies and familial patterns that are not of genetic origin.

Geneticist Ruth Hubbard argues that genetic susceptibility does not determine behavior, saying that:

> Most inherited conditions exhibit a variety of symptoms and patterns of development, and may turn out to be families of related conditions rather than unique entities . . . . The situation becomes even more complicated when scientists try to predict conditions that are said to involve inherited "tendencies." . . . From a therapeutic perspective, it makes little sense to try to sort out the genes involved with complex genetic conditions, even if DNA is involved at some level . . . . Not only will this not cure or prevent the condition, it will create a new group of stigmatized people. (Hubbard and Wald, 1999, p. 37)

Freitag and Retz (2010) summarize studies of monozygotic and dizygotic twins that conclude 60 to 80 percent of ADHD in children and adolescents is hereditary and therefore might be a genetically determined trait. Joseph (2013) vitiates the conclusion of ADHD's genetic origin because all the studies are based on thoroughly discredited equal environmental assumptions. Twin studies of this kind cannot demonstrate ADHD has an underlying genetic proclivity any more than they could be used to argue that monozygotic twins are genetically more vulnerable to snake poison. The fallacy of that conclusion would be related to the environmental factor of identical twins frequently traveling and being bitten together, which has nothing to do with a genetic vulnerability to the poison. Joseph concludes that "genetic interpretations of twin method data in political science, psychology, psychiatry, and other social and behavioral sciences must be rejected outright" (p. 34).

Neurological, Emotional, and Developmental Correlations

Lydia Mary Furman writes: "evidence for a genetic or neuroanatomic cause of ADHD is insufficient. Experimental work shows that executive function deficits do not explain ADHD. The psychometric properties of widely used ADHD rating scales do not meet standards expected for disease identification" (Furman, 2008, p. 775). She concludes that ADHD is unlikely to exist as an identifiable disease and that its diagnostic criteria are symptoms of other treatable conditions underlying the medical, emotional, and psychosocial condition of children.

In an analysis of the epistemology of ADHD, Thurber, Sheehan, and Roberts (2009) suggest that conflicting claims regarding ADHD's etiology stem from a conflict between those who subscribe to the edicts of established institutions, and those who employ criteria derived from the scientific method. They assert that discussions of ADHD are dominated by persons of authority and power, and, consequently, by the institutions that grant authority.

This supports the current thesis that ADHD is an emergent property because it identifies the leaders in the debate as institutions that attempt to maximize their advantage in political influence, financial benefit, or other gross measures. Support also comes from the authors' observation that the discussion of ADHD has been indifferent to differences in the meaning of basic terms, investigative methods, and scientific standards. Failing to reconcile these differences creates factionalism, weakens critical feedback, and allows institutions to

shape the field according to their own needs. Thurber et al. conclude that "ADHD currently does not have status beyond that of the 'hypothetical construct.' Moreover, current brain-based causal models have failed to provide rigorous supporting data that comes [sic] from testing falsifiable hypotheses" (p. 33).

EEG Research

A test of a falsifiable hypothesis of the biological origin of ADHD has been done by Ogrim, Kropotov, and Hestad (2012), who measured the brainwaves of children labeled ADHD. The authors attempted to determine if differences in the amplitude of certain brainwaves can discriminate between children with and without the ADHD label, as claimed in previous studies. Kropotov is a specialist in EEG analysis and this study is one of the most rigorous works to date. Ogrim et al. concluded:

> We hypothesized that the accuracy of the theta/beta ratio, and theta and beta separately to discriminate between ADHD and normal controls would be 80 percent or more. This was not found. In fact none of the three EEG measures were significantly different in patients and controls . . . . Our results do not confirm research showing that elevated theta/beta ratio captures most ADHD, but are more in accordance with research showing several EEG patterns in ADHD. (Ogrim et al., 2012)

In particular, they found that 26 percent of their ADHD subjects showed a brainwave pattern (elevation in the level of frontal theta frequencies) that occurred in less than 3 percent of their controls. An excess of theta waves is a transient state in normal subjects that is traditionally associated with a dreamy state of mind and a lack of focus on the external environment. This suggests there may be a subclass of people labeled ADHD with either a measurable biological difference or a different cognitive style, neither of which is necessarily pathological. While this is a small portion of the population, it is roughly in line with 15 percent of clients in the Cummings and Wiggins (2001) study whose ADHD symptoms were not resolved by psychotherapy alone.

Wright (2005, p. 129) has noted a host of neurological, physical, and emotional factors that can cause ADHD-like symptoms, and rarely is a sufficiently thorough evaluation done to distinguish between these individuals and those with real cognitive problems (Leslie, Weckerly, Plemmons, Landsverk, and Eastman, 2004). Some of these may be related to, or develop into the depression noted by Weinberg et al. in 74 percent of their subjects identified as ADHD, although it was the neurological, physical, and emotional issues which they claim were the major factors determining their subjects' behavior (Weinberg, Harper, Emslie, and Brumback, 1995; see also Brumback, 2000).

No biological cause for ADHD has been found and no biological marker for any psychiatric disorder has been seen. Specialists in many disciplines have

engaged in the exploration of alternative hypotheses with scant attempt to confirm or refute any one hypothesis. The fact that a great quantity of conjectural work continues to be funded to explore possible new biological etiologies in spite of this is further indication that this work is motivated by institutional advantage, the pressures of professional advancement, and other forces within the social system itself.

Diagnostic Tests

As spelled out in the DSM, a diagnosis of ADHD rests on a comprehensive history and careful observations. This is clarified in instructions for testing, given by the Educational Testing Service (2008). Two types of test and one type of measure used to substantiate the diagnosis of ADHD are Continuous Performance Tests, structured personal interviews, and neurological profiles such as EEG and brain images.

Neither Continuous Performance Tests (Gillen, 2003; Riccio, Reynolds, and Lowe, 2001) nor neurological profiles demonstrate a level of accuracy deemed sufficient for clinical diagnosis (Loo and Makeig, 2012). The structured personal interview consists of at least one interview with observers of the subject, and at least one interview with the subject herself (Gualtieri and Johnson, 2005). The DSM's diagnostic criteria are detailed, but there is no instruction within the DSM as to how these criteria are to be met. For example, there is no instruction as to how to measure inattention, lack of focus, or impulsivity.

Consistent conclusions about a person considered for the ADHD diagnosis can be drawn by different practitioners who interview third-party observers such as parents, teachers, counselors, or administrators, but conclusions drawn by practitioners interviewing the subject have been shown to be inconsistent, so that there can be no reliable collective conclusion. The k statistic for inter-rater reliability of child-based interviews is reported at .10, where values below .20 indicate weak inter-rater reliability (Landis and Koch, 1977). There is no consistent diagnosis based on the subject's own statements or presentation in an interview. The only basis of diagnosis that is consistent — and this does not mean a valid diagnosis — rests on statements solicited by raters from third parties concerning the child. The validity of this diagnosis is doubtful since third parties are often partial, being employed by the institutions that have something to gain from a positive diagnosis, and are chosen because they support a positive diagnosis. As Kendell (1993, p. 290) points out, "reliability can be very high while validity remains trivial and in such a situation high reliability is of very limited value." It is considered acceptable — and in many cases all that is undertaken — to base a diagnosis solely on an interview with a child's caregiver and a report from school authorities. A recent study estimates that 90 percent of medical specialists who diagnose ADHD in preschoolers do not follow clinical

guidelines published by the American Academy of Pediatrics (North Shore–Long Island Jewish Health System, 2013).

Results of the Collaborative Primary Care/Behavioral Health Model

In 2001 Cummings and Wiggins reported on the result of 168,113 cases of behavioral intervention in the treatment of children and adolescents from five to 18 years old who received psychotropic medication over the period of 1988 to 1992, roughly half of whom were diagnosed with ADD/ADHD. Their retrospective data were summary in nature but the uniform diagnostic protocol and treatment make the study effective in assessing the effect of behavioral treatment on an actual ADD/ADHD sample rather than on an unrepresentative sample selected using diagnostic standards in order to narrow statistical variance.

Cummings and Wiggins' collaborative model involved primary care physicians, behavioral care therapists, parents, educators, social workers, peers, and juvenile authorities, and resulted in an assessment protocol that included more review and input from different parties than appears to be the average for the treatment of ADHD today. This gave greater weight to people inclined to maintain the diagnosis for the purpose of avoiding social disruption than what would result from a more strict diagnosis according to the guidelines given in the DSM. For these reasons the following results can be taken as more typical of real-world diagnosis and more pessimistic regarding the effect of behavioral therapy than what would result from a strictly clinical diagnosis.

Behavioral intervention consisted of an average of 17.2 sessions of therapy of which 6.3 were conducted with the child and 10.9 with the parent-figure. The study reports that 61 percent of the boys and 23 percent of the girls in the pre-treatment population were diagnosed with and medicated for ADD/ADHD compared with only 11 percent of the boys and 2 percent of the girls who retained the diagnosis at the conclusion of the intervention. Of the whole population, including those with diagnoses other than ADD/ADHD, less than 3 percent had to resume medication following discontinuance after having been diagnosed as free of symptoms.

These findings show that in one of the largest, if not the largest nation-wide sample of children diagnosed with ADD/ADHD, the symptoms of 82 percent and 91 percent of boys and girls respectively were resolved through behavioral therapy, with a relapse of rate of less than 3 percent. To be conservative we can say that the full 3 percent were improperly diagnosed as having been freed of ADD/ADHD behavior. This would then mean that on average at least $((82 + 91)/2) - 3 = 83.5$ percent of children grouped by sex and diagnosed with ADD/ADHD did not suffer a biological dysfunction, since such a dysfunction could not have been resolved by therapy without medication. Regarding the remaining 16.5 percent, nothing can be said with regard to whether their condition was

or was not of a biological nature. For ease of comparison I have rounded this figure up to 85 percent when referencing this result.

A meta-analysis was recently conducted on the effectiveness of parent behavioral therapy on the remediation of ADHD behavior in pre-school children up to six years old. Though these studies dealt with a mostly younger population, the results support the observations of Cummings and Wiggins in concluding that behavioral therapy shows "high strength of effectiveness for improving child disruptive behavior, including ADHD, in pre-schoolers," and "Methylphenidate (Ritalin) has low strength of effectiveness for improving child disruptive behavior, including ADHD . . ." (Charach, Carson, Fox, Ali, Beckett, and Lim, 2013, p. 12).

The Fiction of the Biological Model

Allen Frances, lead editor and chairman of the DSM-IV task force, believes the high number of children diagnosed with ADHD constitute a "faddish overdiagnosis" (Frances, 2012; Greenberg, 2011). Based on the results of Cummings and Wiggins we can infer that this diagnosis fails to identify patients' real issues 85 percent of the time. There is no dispute that there exists a population at risk, under stress, and in need of help. The issue revolves around finding a solution.

Evidence of the fallaciousness of the ADHD diagnosis has been known for years, and new information continues to support it, yet few embrace the evidence from the wider perspective shown here. The question for children diagnosed with ADHD and their parents is who should be given the authority to define the problem and its solution. There is no one "thing" that is ADHD. It is an emergent social construct applied to whichever people or groups of people fit the description.

Conclusion

Feedback Systems

The argument that ADHD is a structure created to benefit institutions begs the question of how institutions, whose agents express a concern for the welfare of children, could develop a structure that exploits children. To answer this we need to understand how structures develop from the interactions between systems and their agents.

The notion of feedback is essential to an understanding of how systems develop and influence those affected by them (Richardson, 2011). The formal theory of systems describes system regulation using feedback loops (Kirkwood, 1998). Positive feedback loops form reinforcing patterns that amplify effects; negative feedback loops form braking or "de-inforcing" patterns reducing actions or effects. Most social feedback systems have inherent limitations to growth so

that, unlike electronic feedback, they do not become unstable in the way that generates the familiar auditory squeal of runaway amplification.

The existence of ADHD rests on a series of feedback loops that exist within the context of social, political, and economic systems. ADHD is a phenomenon that emerges from the reinforcing feedback of society itself. Public education — and private education following the public model — play a central role in defining ADHD by involving corporations, psychologists, and parents in the loops shown in Figure 1 and described in Table 1.

System Dynamics and Emergent Properties

In this paper I am considering a system that generates some number of diagnosed cases of ADHD over a period of time. Each of four major factions involved in manufacturing this number — psychiatrists, educators, parents, and pharmaceutical companies — benefit from this diagnosis. What is more, with each increase in the number of those diagnosed with ADHD, the net reward generated for each faction increases proportionally. The benefits to each faction "feed back" in proportion to the growing numbers of children diagnosed. For each of the four factions this feedback is positive. If the consequence of a growing number of diagnosed cases of ADHD were negative for one or more factions, which could arise due to some net cost, risk, or other detrimental impact, then the system would contain negative feedback.

The type of feedback strongly affects how a faction responds. Three basic types of feedback are constant, linear, and proportional. A constant positive feedback occurs when a choice results in a fixed reward over a period of time. If all parents as a group received a single, lump sum reward for accepting the ADHD diagnosis, and this is not the case, then accepting the diagnosis would be described as having a constant, positive feedback effect. Linear feedback occurs when the reward increases in proportion to the duration over which the choice is maintained. If schools as a group received an extra, annual subsidy for supporting students diagnosed with ADHD, which also is not the case, then diagnosis would generate positive feedback that was linear over time. Proportional feedback occurs when an additional reward is obtained from each and every instance where the diagnosis is made. Pharmaceutical companies experience positive proportional feedback because every newly diagnosed person that is given medication, which occurs in one half to two thirds of the cases, results in additional income. In fact, all four of the factions experience proportional feedback in which the reward for accepting the diagnosis increases in proportion with the number of cases diagnosed.

Systems with a mixture of positive and negative feedback can reach an equilibrium where the gains to one faction are offset by the losses suffered by another. Systems governed by a fixed incentive will shift their state and then stabilize.

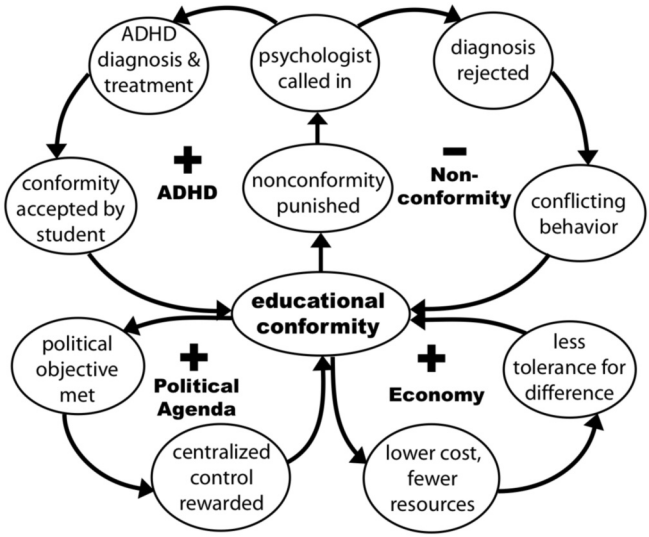

Figure 1: Educational conformity feedback loops. ADHD, political agenda, and economy provide positive feedback, nonconformity provides negative feedback.

Systems experiencing linear feedback tend to change at a constant rate over time. Systems that experience positive proportional feedback manifest exponential growth with an increasing number of rewarding choices being made in each period. This unconstrained growth slows only when rewards diminish, or penalties increase, for one or more of the factions.

The system considered here consists of four major factions, each separately responding to the choice of whether or not to accept the ADHD diagnosis. This is a first-order model because I am considering the factions as being independent from each other. A more sophisticated model would be a so-called second-order model in which additional rewards or disincentives arise from the interaction of one faction with another. For example, I have not considered the effect on psychologists, parents, or educators that arises from the decision of pharmaceutical companies to encourage or discourage the diagnosis. Second-order effects are usually smaller than first-order effects, at least in the early stages of a system's development. I am assuming that these forces are not of critical importance in the general evolution of this system, at least not yet. An investigation of these second-order effects is a topic for future study.

The average annual parent-reported rate of diagnosed ADHD in male and female children aged four to 17 has gone from 4.8 percent in 1997, to 7.8 percent in 2003, to 9.5 percent in 2007 (Visser, Bitsko, Danielson, Perou, and Blumberg, 2010). Almost 20 percent of males four to 17 years old were given the diagnosis

Table 1

Explanation of Feedback Loops

Loop Name	Purpose	Description
Educational agenda	Conformity	Educational managers serve their mandate by instilling conformity leading to an agenda that ostracizes ADHD behaviors. Eliminating these behaviors eliminates objections to conformity-based schooling, strengthens cost-based and test-based measures of success that provide evidence in support of the conformist approach.
ADHD diagnosis	Conformity	Nonperforming and noncompliant students are punished with lower grades, disciplinary action, and stigmatization. Psychological services are applied to mitigate the ensuing conflict by issuing the ADHD diagnosis and prescribing treatment. The diagnosis justifies blaming performance failure on student deficiencies. Successful treatment remediates the problem of noncompliance and performance measures related to it.
Educational economy	Lower costs	Greater conformity supports the factory model of education that remains at the root of the public education program. Conformity supports hierarchical management and services, lowering costs by requiring less diversity of products and services.
Institutional psychology	Profit	Psychological institutions and individual practitioners carve an economic niche and a social role in defining and servicing the needs of the ADHD community. Greater recognition and acceptance of ADHD reinforces this loop.
Pharmaceutical industry	Profit	Increased pharmaceutical treatment of ADHD raises profits that feed greater support for ADHD advertising and research, thus reinforcing this loop.
Parental	Compliance and stability	Angry, depressed, reactive children suffering insults for noncompliance bring amplified frustrations back to the family. Parents are pressured by schools to either bring their children into compliance or consent to diagnosis and treatment. Parental acceptance of ADHD confers validity on the diagnosis and encourages other parents to do the same.

Table 1 (Continued)

Explanation of Feedback Loops

Loop Name	Purpose	Description
Media	Profit	Increasing public acceptance of and interest in the diagnosis enables the media to build an audience and a profitable pharmaceutical advertising business.
Academic research	Profit	Academic journals endeavor to serve and focus interest, and their own interests are more readily served when there is acceptance of and interest in the diagnosis.

in 2012 (Schwarz and Cohen, 2013). The model given here predicts that the rate of diagnosis of ADHD will continue to grow until some factions are negatively impacted, at which point the positive feedback system breaks down and the acceleration in the number of diagnosed cases of ADHD will slow and may reverse. The evolution of the diagnosis has all to do with gains and losses of institutions and has little to do with whether or not the ADHD label corresponds to a genuine medical disorder, or whether its diagnosis and treatment offers a benefit to individuals. This follows without any requirement that ADHD exist as anything more than a label.

ADHD as an Emergent Property

The program to create an ADHD-compliant culture is not necessarily intentional or recognized as a program by those who participate in it. This is a key observation of systems theory: the outcome does not need to be an intentional goal of those involved if the process is reinforced by the rules on which the system operates. In such a case the outcome is an "emergent property" of the system (MacLennan, 2007). If the process is stable and generates positive feedback, then the outcome may be more likely if the actors are unaware of the process and simply act in accordance with rewards and expectations: that is to say, when those in the system do not question the process or their assigned roles (Meyer and Rowan, 1977).

Processes of this sort are evolutionary, proceed by natural selection without central direction (Richerson and Boyd, 1984), and develop in a self-organizing manner through the sharing of resources (Ostrom, 2009). These strong aggregating forces lead to the emergence of behaviors in which it is commonly found that the details do not matter (Miller and Page, 2007, p. 154). Whether or not ADHD is a real dysfunction is, in this case, one of those irrelevant details.

Even this simple, first-order model shows us that the prevalence of ADHD increases in response to positive feedback even though no one faction is acting with the stated intention of increasing the number of people given the diagnosis. This process is one of many assaults on the autonomy and professionalism of "doctors, scientists, and teachers [which are] being increasingly replaced by the needs and dictates of corporate America" (Welch, 2008, p. 183), which is to say by virtue of the feedback loops that exist within the larger system. This is just Adam Smith's market dynamics in which every individual "intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention" (Smith, 1843, p. 184).

I have presented a simple picture of feedback systems with no differential equations, contingent probabilities, external sources, damping factors, or hysteresis effects. There is only one bifurcation in Figure 1 that distinguishes the path of students who are given the ADHD diagnosis from the so called normal populations.

To counter the argument that this picture is too simple, three points should be emphasized. First, this holistic picture contains within it all of the complexity of the subsystems that it includes. The holistic picture examines the forces operating on society, medicine, and science that are generated by large, vested interests and finds that these forces are simple. This picture has traditionally been overlooked and is a necessary first step in any understanding of how the system has developed and will continue to evolve.

Second, remove any one of these reinforcing feedback loops and the number of ADHD diagnosis would shrink or disappear entirely. And third, a systems theory approach is predictive and therefore testable and falsifiable. The prediction is that the system will develop in such a way as to create the greatest benefit, where "benefit" is defined separately by each group according to the power each group exerts on the development of the whole.

I have argued in this paper that ADHD diagnosis does not denote a disease of an individual and I have described how the system functions. Remove psychologists, education, pharmaceuticals, and parental support, and what we now call ADHD would splinter into the set of issues of which it is composed. We are already seeing this as parents clamor for access to ADHD prescription drugs as a means of enhancing their children's academic performance. If such general drug use is seen as a "win–win" proposition for the more powerful social and political factions, then it will likely manifest in spite of evidence of adverse drug effects, long term health risks, and lack of benefit to the patient.

This systems theory model explains the dominance of the biomedical model as a result of the benefits that the model provides — not to patients but to psychiatrists, drug companies, and the educational establishment — regardless of the evidence that supports or refutes the biomedical model. The systems theory model predicts that the ADHD diagnosis will persist for as long as the system that generates the diagnosis continues to profit from it. The biomedical model

supports the introduction of powerful psycho-pharmaceuticals into a child population and a school environment for which this was previously socially unacceptable. At the same time, the model is limiting because it compels the use of drugs for only a limited population who are identified by a doctor or therapist.

Failure of Institutions

The notion of ADHD as a biological disease helps to sell compulsory education's social engineering program. As described by Richard DeGrandpre,

> The difficulties experienced by the ADHD-diagnosed child vary across several dimensions, and each dimension — behavioral, cognitive, experiential — has its own continuum of severity. It is a mistake to try to flatten this multidimensional picture into a simple yes or no — as does the DSM in its diagnostic criteria for "ADHD." Such over-simplification is designed not to clarify and address these problems but to label children medically and then, as occurs in the vast majority of cases, "manage" their symptoms with psychiatric medications. A billion-dollar industry has grown up with the explicit function of carrying out this scheme. (DeGrandpre and Hinshaw, 2000)

ADHD is a fictitious illness projected by caregivers onto the child for whom care is given. As a systemic dysfunction, like global warming, obesity, and the debt crisis, it is a mistake to believe that ADHD will naturally evolve in accordance with the tenets of health care, or in a manner beneficial to children.

References

American Academy of Pediatrics, and National Initiative for Children's Healthcare Quality (2002). *Caring for children with ADHD: A resource toolkit for clinicians* (first edition). Retrieved from http://www.nichq.org/adhd_tools.html

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (fourth edition, text revised). Washington, American Psychiatric Association. doi:10.1176/appi.books.9780890423349

Bennett, S., and Kalish, N. (2006). *The case against homework: How homework is hurting our children and what we can do about it*. New York: Random House.

Berman, P., and McLaughlin, M.W. (1974). *Federal programs supporting educational change, Volume 1: A model of educational change*. Santa Monica: Rand Corporation.

Biederman, J., Milberger, S., Faraone, S. V., Kiely, K., Guite, J., et al. (1995). Family–environment risk factors for Attention-Deficit Hyperactivity Disorder. *Archives of General Psychiatry, 52*, 464–470. doi: 10.1001/archpsyc.1995.03950180050007

Breggin, P.R, (2008). *Brain-disabling treatments in psychiatry*. New York: Springer Publishing.

Breuner, J. (1977). *The process of education*. Cambridge, Massachusetts: Harvard University Press.

Brumback, R.A. (2000). Weinberg's syndrome: A disorder of attention and behavior problems needing further research. *Journal of Child Neurology, 15,* 478–480.

Campbell, E.G., Gruen, R.L., Mountford, J., Miller, L.G., Cleary, P.D., and Blumenthal, D. (2007). A national survey of physician–industry relationships. *New England Journal of Medicine, 356,* 1742–1750.

Carey, B., and Harris, G. (2008, July 12). Psychiatric group faces scrutiny over drug industry ties. *New York Times*. Retrieved from http://www.nytimes.com/2008/07/12/washington/12psych.html

Carlat, D.J. (2010). *Unhinged, the trouble with psychiatry — A doctor's revelations about a profession in crisis*. New York: Simon and Schuster.

Carlson, E.A., Jacobvitz, D., and Sroufe, L.A. (1995). A developmental investigation of inattentiveness and hyperactivity. *Child Development, 66,* 37–54

Charach, A., Carson, P., Fox, S., Ali, M. U., Beckett, J., and Lim, C.G. (2013). Interventions for preschool children at high risk for ADHD: A comparative effectiveness review. *Pediatrics, 131*(5), e1584–e1604. Published online April 1, 2013. doi: 10.1542/peds.2012-0974

Cohen, A. (2006, September 14). Kids may be right after all: Homework stinks. *USA Today*, p. 13A.

Conrad, P. (2006). *Identifying hyperactivity in children: The medicalization of deviant behavior*. Burlington, Vermont: Ashgate Publishing.

Cooper, H., Robinson, J.C., and Patall, E.A. (2006). Does homework improve academic achievement?: A synthesis of research, 1987–2003. *Review of Educational Research, 76,* 1-62.

Cosgrove, L., Bursztajn, H. J., Krimsky, S., Anaya, M., and Walker, J. (2009). Conflicts of interest and disclosure in the American Psychiatric Association's clinical practice guidelines. *Psychotherapy and Psychosomatics, 78,* 228–232.

Cosgrove, L., Krimsky, S., Vijayaraghavan, M., and Schneider, L. (2006). Financial ties between DSM–IV panel members and the pharmaceutical industry. *Psychotherapy and Psychosomatics, 75,* 154–160. doi: 10.1159/000091772

Cummings, N.A., and Wiggins, J.G. (2001). A collaborative primary care/behavioral health model for the use of psychotropic medication with children and adolescents: A report of a national retrospective study. *Issues in Interdisciplinary Care, 3,* 121–128.

Curcio, A.A. (2006). Civil claims for uncivilized acts: Filing suit against the government for American Indian boarding school abuses. *Hastings Race and Poverty Law Journal, 4,* 45–130.

DeGrandpre, R., and Hinshaw, S.P. (2000). *ADHD: Serious psychiatric problem or all-American cop-out?* Retrieved from http://www.dana.org/news/cerebrum/detail.aspx?id=3214

Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. New York: The Macmillan Company.

Dewey, J. (1984). *John Dewey, the later works, 1925–1953, Volume 3: 1927–1928* [J.A. Boydston, Ed.]. Carbondale, Illinois: Southern Illinois University Press.

DiMasi, J., Hanse, R., and Grabowski, H. (2003). The price of innovation: New estimates of drug development costs. *Journal of Health Economics, 22,* 151–185.

Eakman, B.K. (1991). *Education for the "New World Order."* Portland, Oregon: Halcyon House.

Eakman, B.K. (2007). *Walking targets: How our psychologized classrooms are producing a nation of sitting ducks*. Raleigh, North Carolina: Midnight Whistler Publishers.

Eberstadt, M. (1999). Why Ritalin rules: Parents and teachers have lost patience with childhood. *Policy Review, 94,* 1–13. Retreived from http://www.hoover.org/publications/policy-review/article/6494

Educational Testing Service (2008). *Policy statement for documentation of Attention-Deficit/Hyperactivity Disorder in adolescents and adults*. Princeton, New Jersey: Educational Testing Service.

El-Zein, R.A., Abdel–Rahman, S.Z., Hay, M.J., Lopez, M.S, Bondy, M.L., Morris, D.L., and Legator, M.S. (2005). Cytogenetic effects in children treated with methylphenidate. *Cancer Letters, 230,* 284–291.

Fava, G. (2008). Should the drug industry work with key opinion leaders? *No. British Medical Journal, 336*(7658), 1405.

Frances, A. (2012, May 12). Diagnosing the D.S.M. *New York Times*, p. A19. Retrieved from http://www.nytimes.com/2012/05/12/opinion/break-up-the-psychiatric-monopoly.html

Freitag, C.M., and Retz, W. (2010). Family and twin studies in Attention-Deficit Hyperactivity Disorder. In W. Retz and R.G. Klein (Eds.), *Attention-Deficit Hyperactivity Disorder (ADHD) in adults. Key issues in mental health* (Volume 176, pp. 38–57). Basel: Karger. Retrieved from http://www.karger.com/ProdukteDB/Katalogteile/isbn3_8055/_92/_37/KIMH176_02.pdf

Furman, L.M. (2008). Attention-Deficit hyperactivity disorder (ADHD): Does new research support old concepts? *Journal of Child Neurology, 23,* 775–784.

Gagnon, M.A., and Lexchin, J., (2008). The cost of pushing pills: A new estimate of pharmaceutical promotion expenditures in the United States. *PLoS Medicine, 5*(1), e1. doi:10.1371/journal.pmed.0050001

Gatto, J.T. (2006). *The underground history of American education: An intimate investigation into the prison of modern schooling.* New York: Oxford Village Press.

Gau, S.S., and Chiang, H. (2009). Sleep problems and disorders among adolescents with persistent and subthreshold Attention-deficit/Hyperactivity Disorders. *Sleep, 32,* 671–679. Retrieved from http://www.journalsleep.org/ViewAbstract.aspx?pid=27462

Gillen, W.G. (2003). The TOVA: Discriminant validity for ADHD and depression. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 64* (3–B), 1491.

Greenberg, G. (2011). Inside the battle to define mental illness. *Wired*. December, 27, 2010. Retrieved from http://www.wired.com/magazine/2010/12/ff_dsmv/

Greene, S.L., Kerr, F., and Braitberg, G. (2008) Amphetamines and related drugs of abuse. *Emergency Medicine Australasia, 20,* 391–402.

Gualtieri, C.T., and Johnson, L.G. (2005). ADHD: Is objective diagnosis possible? *Psychiatry (Edgmont), 2*(11), 44–53. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2993524/

Harris, G. (2009, October 22). Drug makers are advocacy group's biggest donors. *New York Times*, p. A23. Retreived from http://www.nytimes.com/2009/10/22/health/22nami.html

Harrison, C., and Sofronoff, K. (2002). ADHD and parental psychological distress: Role of demographics, child behavioral characteristics, and parental cognitions. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*, 703–711.

Havelock, R.G. (1973). *The change agent's guide to innovation in education*. Englewood Cliffs, New Jersey: Educational Technology Publications.

Herxheimer, A. (2003). Relationships between the pharmaceutical industry and patients' organisations. *British Medical Journal, 326,* 1208–1210.

Higgins, E.S. (2009). Do ADHD drugs take a toll on the brain?. *Scientific American Mind, 20*, 38–43. doi:10.1038/scientificamericanmind0709-38.

Hofstadter, R. (1962). *Anti-intellectualism in American life*. New York: Vintage Books.

Horn, R.A. Jr. (2007). Edward L. Thorndike. In J.L. Kincheloe and R.A. Horn, Jr. (Eds.), *The Praeger handbook of education and psychology: Volume 1* (pp. 225–230). Westport, Connecticut: Praeger Publishers.

Hubbard, R., and Wald, E. (1999). *Exploding the gene myth: How genetic information is produced and manipulated by scientists, physicians, employers, insurance companies, educators, and law enforcers*. Boston: Beacon Press.

Joseph, J. (2000). Not in their genes: A critical view of the genetics of Attention-Deficit Hyperactivity Disorder. *Developmental Review, 20,* 539–567. doi.org/10.1006/drev.2000.0511

Joseph, J. (2013). The use of the classical twin method in the social and behavioral sciences: The fallacy continues. *Journal of Mind and Behavior, 34,* 1–40. Retreived from http://jayjoseph.net/yahoo_site_admin/assets/docs/2013_Joseph_Fallacy_of_the_Twin_Method_in_the_Social_and_Behavioral_Sciences.262140341.pdf

Kendell, R.E. (1993). Diagnosis and classification. In R.E. Kendell and A.K. Zealley (Eds.), *Companion to psychiatric studies* (fifth edition, pp. 277–294). Edinburgh: Churchill Livingstone.

Kendler, K.S., and First, M.B. (2010). Alternative futures for the DSM revision process: Iteration v. paradigm shift. *The British Journal of Psychiatry, 197,* 263–265. doi: 10.1192/bjp.bp.109.076794

Kirkwood, C.W. (1998). *System dynamic methods: A quick introduction.* Tempe, Arizona: Arizona State University Press. Retrieved from http://www.public.asu.edu/~kirkwood/sysdyn/SDIntro/SDIntro.htm

Landis, J.R., and Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics 33,* 159–174. doi:10.2307/2529310

Leo, J., and Cohen, D. (2003). Broken brains or flawed studies? A critical review of ADHD neuroimaging research. *Journal of Mind and Behavior, 24,* 29–56.

Loo, S.K., and Makeig, S. (2012). Clinical utility of EEG in Attention-Deficit/Hyperactivity Disorder: A research update. *Neurotherapeutics, 9,* 569–587. doi: 10.1007/s13311-012-0131-z

Leslie, L., Weckerly, J., Plemmons, D., Landsverk, J., and Eastman, S. (2004). Implementing the American Academy of Pediatrics Attention-Deficit/Hyperactivity Disorder diagnostic guidelines in primary care settings. *Pediatrics, 114,* 129–140.

Lu, L.Y., and Ortlieg, E.T. (2009). Teacher candidates as innovative change agents. *Current Issues in Education, 11*(5), 2–6. Retreived from http://cie.asu.edu/volume11/number5/

MacLennan, B. (2007). Evolutionary psychology, complex systems, and social theory. *Soundings: An Interdisciplinary Journal, XC,* 169–189. Retrieved from http://web.eecs.utk.edu/~mclennan/papers/EPCSST.pdf

Mannuzza, S., and Klein, R.G. (2000). Long-term prognosis in Attention-Deficit/Hyperactivity Disorder. *Child and Adolescent Psychiatric Clinics of North America, 9,* 711–726.

Masia, N. (2008). The costs of developing a new drug. In G. Clack and M.S. Neely (Eds.), *Focus on intellectual property rights* (pp. 82–83). U.S. Department of State. Retrieved from http://www.america.gov/st/econ-english/2008/April/20080429230904myleen0.5233981.html

Maté, G. (1999). *Scattered: How attention deficit disorder originates and what you can do about it*. New York: Penguin Group.

Meyer, J.W., and Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *The American Journal of Sociology, 83,* 340–363. Retrieved from http://www.sleepfoundation.org/sites/default/files/Highlights_facts_06.pdfhttp://www.sasse.se/akademiska/310/meyer%20rowan.pdf

Miller, J.H., and Page, S.E. (2007). *Complex adaptive systems: An introduction to computational models of social life*. Princeton, New Jersey: Princeton University Press.

Miller, R. (2002). *Free schools, free people: Education and democracy after the 1960s*. Albany, New York: State University of New York Press.

National Sleep Foundation (2006). *2006 Sleep in America Poll: Highlights and key findings.* Washington, DC: National Sleep Foundation. Retrieved from http://www.sleepfoundation.org/sites/default/files/2006_summary_of_findings.pdf

Nauer, R. (2009, September). Use of ADHD meds has grown steadily. *Psych Central*. Retrieved from http://psychcentral.com/news/2011/09/29/use-of-adhd-meds-has-grown-steadily/29884.html

Newton–Howes, G. (2004). What happens when children with Attention Deficit/Hyperactivity Disorder grow up? *Journal of the Royal Society of Medicine, 97*, 531–535.

North Shore–Long Island Jewish Health System (2013, May 4). 90 percent of pediatric specialists not following clinical guidelines when treating preschoolers with ADHD. *ScienceDaily*. Retrieved from http:// http://www.sciencedaily.com/releases/2013/05/130504163310.htm

Ogrim, G., Kropotov, J., and Hestad, K. (2012). The QEEG theta/beta ratio in ADHD and normal controls: Sensitivity, specificity, and behavioral correlates. *Psychiatry Research, 198*, 482–488. doi: 10.1016/j.psychres.2011.12.041

Ostrom, E. (2009). Beyond markets and states: Polycentric governance of complex economic systems. *Nobel Prize Lecture*. Retrieved from http://www.nobelprize.org/nobel_prizes/economics/laureates/2009/ostrom_lecture.pdf

Pachter, W.S.., Fox, R.E., Zimbardo, P., and Antonuccio, D.O. (2008). Corporate funding and conflicts of interest: A primer for psychologists. *American Psychologist, 62,* 1005–1015. Retrieved from http://www.apa.org/pubs/journals/releases/amp-6291005.pdf

Peterson, P.L., Clark, C.M., and Dickson W.P. (1990). Educational psychology as a foundation in teacher education: Reforming an old notion. *Teachers College Record, 91,* 322–346.

PLoS Medicine Editors (2012). Does conflict of interest disclosure worsen bias? *PLoS Medicine,* 9(4): e1001210. doi:10.1371/journal.pmed.1001210

Ravitch, D. (2000). *Left back: A century of battles over school reform*. New York: Touchstone.

Ray, B.D. (2011, January 3). 2.04 million homeschool students in the United States in 2010. *National Home Education Research Institute*. Retrieved from http://www.nheri.org/HomeschoolPopulationReport2010.pdf

Riccio, C.A., Reynolds, C.R., and Lowe, P.A. (2001). *Clinical applications of continuous performance tests: Measuring attention and impulsive responding in children and adults*. New York: John Wiley & Sons.

Richardson, G.P. (2011). Reflections on the foundations of system dynamics. *System Dynamics Review, 27,* 219–243. doi; 10.1002/sdr.462. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/sdr.462/pdf

Richerson, P.J., and Boyd, R. (1984). Natural selection and culture. *BioScience, 34,* 430–434. Retrieved from http://www.des.ucdavis.edu/faculty/Richerson/BioScience.pdf

Scheffler, R.M., Hinshaw, S.P., Modrek, S., and Levine, P. (2007). The global market for ADHD medications. *Health Affairs, 26,* 450–457.

Schwarz, A. (2012, October 9). Attention disorder or not, pills to help in school. *New York Times*, p. A1. Retrieved from http://www.nytimes.com/2012/10/09/health/attention-disorder-or-not-children-prescribed-pills-to-help-in-school.html?

Schwarz, A., and Cohen, S. (2013, April 1). ADHD seen in 11% of U.S. children as diagnoses rise. *New York Times*, p. A1. Retreived from http://www.nytimes.com/2013/04/01/health/more-diagnoses-of-hyperactivity-causing-concern.html?

Sen, S., and Prabhu, M. (2012). Reporting bias in industry-supported medication trials presented at the American Psychiatric Association meeting. *Journal of Clinical Psychopharmacology, 32*(3), 435.

Smith, A. (1843). *An inquiry into the nature and causes of the wealth of nations, Book IV, Chapter II.* Edinburgh: Thomas Nelson.

Stahl, W.K. (1979). The U.S. and Native American education: A survey of federal legislation. *Journal of American Indian Education, 18,* 28–32.

Stergiakouli, E., and Thapar, A. (2010). Fitting the pieces together: Current research on the genetic basis of Attention-Deficit/Hyperactivity Disorder (ADHD). *Neuropsychiatric Disease and Treatment, 6,* 551–560. doi: 10.2147/NDT.S11322

Swanson, J.M., Elliott, G.R., Greenhill, L.L., Wigal, T., Arnold, L.E., Vitiello, B., et al. (2007). Effects of stimulant medication on growth rates across 3 years in the MTA follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry, 46,* 1015–1027.

Thurber, S., Sheehan, W., and Roberts, R.J. (2009). Attention Deficit Hyperactivity Disorder and scientific epistemology. *Dialoges in Philosophy, Mental and Neuro Sciences, 2,* 33–39.

Valenstein, E. (1988). *Blaming the brain: The truth about drugs and mental health*. New York: The Free Press.

van Ewijk, H., Heslenfeld, D.J., Zwiers, M.P., Buitelaar, J.K., and Oosterlaan, J. (2012) Diffusion tensor imaging in ADHD disorder: A systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews, 36,* 1093–1106.

Visser, S.N., Bitsko, R.H., Danielson, M.L., Perou, R., and Blumberg, S.J. (2010). Increasing prevalence of parent-reported Attention-Deficit/Hyperactivity Disorder among children — United States, 2003 and 2007. *Centers for Disease Control, Morbidity and Mortality Weekly Report, 59*(44), 1439–1443.

Von Mises, L. (1944). *Omnipotent government: The rise of the total state and total war.* New Haven, Connecticut: Yale University Press. Retrieved from http://mises.org/document/5829/Omnipotent-Government-The-Rise-of-the-Total-State-and-Total-War

Walsh, F. (2010, September 30). *The genetics of ADHD. BBC News Fergus's Medical Files*. Retreived from http://www.bbc.co.uk/blogs/thereporters/ferguswalsh/2010/09/the_genetics_of_adhd.html

Welch, B. (2008). *State of confusion: Political manipulation and the assault on the American mind.* New York: Thomas Dunne.

Weinberg, W.A., Harper, C.R., Emslie, G.J., and Brumback, R.A. (1995). Depression and other affective illnesses as a cause of school failure and maladaption in learning disabled children, adolescents and young adults. In *Secondary education and beyond: Providing opportunities for students with learning disabilities* (pp. 234–264). No city: Learning Disabilities Association. Retrieved from http://www.depressedchild.org/weinberg%20chapter.htm

Whitaker, R. (2005). Anatomy of an epidemic: Psychiatric drugs and the astonishing rise of mental illness in America. *Ethical Human Psychology and Psychiatry, 7,* 23–35.

Williams, N.M., Zaharieva I., Martin, A., Langley, K., Mantripragada, K., Rossdal, R., et al. (2010). Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: A genome-wide analysis. *The Lancet, 376,* 1401–1408. doi:10.1016/S0140–6736(10)61109-9.

Wright, R.H. (2005). Attention Deficit Hyperactivity Disorder: What it is and what it is not. In R.H. Wright and N.A. Cummings (Eds.), *Destructive trends in mental health: The well-intentioned path to harm* (pp. 129–142). New York: Routledge.

The Journal of Mind and Behavior
Winter and Spring 2014, Volume 35, Numbers 1 and 2
Pages 51–70
ISSN 0271–0137

Experimental Methods for Unraveling the Mind–Body Problem: The Phenomenal Judgment Approach

Victor Yu. Argonov

Pacific Oceanological Institute of the Russian Academy of Sciences

A rigorous approach to the study of the mind–body problem is suggested. Since humans are able to talk about consciousness (produce phenomenal judgments), it is argued that the study of neural mechanisms of phenomenal judgments can solve the hard problem of consciousness. Particular methods are suggested for: (1) verification and falsification of materialism; (2) verification and falsification of interactionism; (3) falsification of epiphenomenalism and parallelism (verification is problematic); (4) verification of particular materialistic theories of consciousness; (5) a non-Turing test for machine consciousness. A complex research program is constructed that includes studies of intelligent machines, numerical models of human and artificial creatures, language, neural correlates of consciousness, and quantum mechanisms in brain.

Keywords: mind–body relationship, neural correlates of consciousness, tests for consciousness

In the twentieth century, scientific progress suggested new hypotheses and approaches to the study of consciousness. These hypotheses softened some old problems and created some new ones. However, the most important issue had not changed significantly since Cartesian times. For the mind–body problem, philosophy still had two basic alternatives: (1) consciousness can be studied and controlled as an objective part of matter, or (2) consciousness contains immaterial "degrees of freedom," which cannot be controlled or observed by objective methods. For the sake of simplicity, I shall call these alternatives "materialism" (including materialistic monism, neutral monism, panpsychism, pantheism, "anomalous" monism, etc.) and "(substance) dualism" (including all theories that regard the human being as a combination of physical body and immaterial soul).

Correspondence concerning this article should be addressed to Victor Argonov, Ph.D., Pacific Oceanological Institute of the Russian Academy of Sciences, 41 Baltiiskaya Street, Vladivostok, Russia 690041. Email: argonov@list.ru

Today, there are four basic views on perspectives of scientific choice between materialism and substance dualism. The first view is that materialism is true, and there are sufficient logical and scientific arguments for it. To support this view, materialist philosophers either consider consciousness as a scientific problem with a materialistic solution (Davidson, 1970; Putnam, 1967) or as a pseudo-problem, and argue that consciousness is a non-scientific, folk term (Churchland and Churchland, 1981). The second view is that science had failed to explain consciousness, and dualism is true (Eccles, 1994; Stapp, 1993). Dualist authors believe that science is obviously incomplete, and modern results (e.g., in quantum mechanics) only support this idea. The third view is radically skeptical: materialism and dualism are the unverifiable doctrines, and mind is incapable of comprehending itself entirely (McGinn, 1989). The fourth view is softly skeptical: in principle, materialism and dualism are verifiable hypotheses, but such verification will be technically possible only in the future (Place, 1956, 1960).

In this paper, I provide arguments for the fourth viewpoint. I argue that today's science has no sufficient arguments for materialism or dualism, but the scientific key to the mind–body problem is our ability to talk about consciousness (produce phenomenal judgments). The study of neural mechanisms of phenomenal judgment production can solve at least some aspects of the mind–body problem and make a scientific choice between various forms of materialism and dualism. Several authors noted the crucial role of phenomenal judgments in the study consciousness (Chalmers, 1997; Elitzur, 1989; Rudd, 2000; Valdman, 1997) and argued that the phenomenal judgment argument can be directly employed for the solution of hard problems (for example, for refutation of ephiphenomenalism, see Rudd, 2000; Valdman, 1997). However, their ideas did not become widely accepted. I make an attempt to construct a general phenomenal judgment approach to the mind–body problem and introduce scientific methods for experimental verification of basic theories of consciousness (both materialistic and dualistic). In contrast to Rudd (2000) and Valdman (1997), I do not attempt to disprove any of these theories, but I provide a scientific tool for their study. I also suggest a non-Turing test for machine consciousness (based on phenomenal judgments).

I use a special set of definitions (see next subsection) optimized for my approach. In particular, I view materialism as a theory in which there is a one-to-one correspondence between all subjective facts and some objective facts (constituting a neural correlate of consciousness). I do not discuss whether this means complete metaphysical reducibility of consciousness to matter or only property dualism. The only question I discuss concerns the correspondence between matter and consciousness, not their identity. Some readers might say that this decreases the philosophical value of my study, but I think that the question of correspondence is more scientific than the question of identity. If the state of consciousness is comprehensively determined by the state of matter, then, in practice, consciousness can be studied and controlled as a material object.

Theoretical Basis of the Phenomenal Judgment Approach

Definitions

In this subsection I introduce and discuss the terminology used in this paper. The term "consciousness" ("phenomenal consciousness," "mind") will be used in a standard philosophical sense of subjective reality (a totality of person's subjective [mental] phenomena: sensations, thoughts, volitional acts, etc). The term "matter" will be used in a sense of objective reality (physical particles, energy, space, their properties, and physical laws). In the framework of this paper, information, stored and processed in physical machines (such as deterministic computers) will be also considered as a part of matter.

"Phenomenal judgments" are the words, discussions, and texts about consciousness, subjective phenomena, and the mind–body problem. In this paper, the term "phenomenal judgments" will be used as a synonym of "speech about consciousness" (or other objective phenomenon containing verbal information about consciousness), not "thoughts about consciousness" (subjective phenomena).

The "neural correlate" of a subjective phenomenon (or of a property of consciousness) is a physiological phenomenon containing comprehensive objective (detectable and/or measurable) information about this phenomenon (or a property of consciousness). Here, the words "comprehensive information" mean that there is one-to-one correspondence between any parameter of a given subjective phenomena (for example, a visual image) and some objective parameter in a brain. Any difference between two subjective phenomena must be manifested in their neural correlates. Different subjective phenomena must have objectively distinguishable neural correlates. "Neural correlate of consciousness" is a physiological system containing comprehensive objective (detectable and/or measurable) information on every subjective phenomenon and every property of consciousness of a creature. Note that such correlates must also contain the information that the creature is conscious. If a neural correlate of consciousness exists, then the body of any conscious creature must differ from the body of any unconscious creature. Otherwise, the subjective fact "I am conscious and this body is mine" would not have neural correlates, so the information in a neural correlate of consciousness is not comprehensive.

"Problematic properties of consciousness" are the properties of consciousness having no current satisfactory scientific explanation (no discovered neural correlates). Problematic properties of consciousness are related to so-called hard problems of consciousness (Chalmers, 1997) and are often hypothesized to be immaterial. Examples of such problematic properties: qualia, unity of consciousness, possibility or impossibility of reincarnation, existence of consciousness in a particular creature. Note that some authors deny the existence of "hard" problems and related properties of consciousness (eliminative materialists such as the Churchlands).

Therefore, strictly speaking, all problematic properties of consciousness (discussed in a literature) are hypothetical.

I define "materialism" (physicalism) as a doctrine stating that a neural correlate of consciousness exists, and "dualism" as a doctrine stating that it does not exist. This choice of terms might look strange for some readers. Some materialists do not state that a comprehensive neural correlate of the whole consciousness definitely exists; they state only that matter produces consciousness. However, if some subjective parameters have no neural correlates, then they are hidden from third-person study, and the state of consciousness is not determined unambiguously by the state of matter. Therefore, I define materialism as a doctrine that each subjective parameter has a neural correlate.

For rigorous distinction among the forms of dualism, I introduce the term "immaterial influence on matter": the change in physical processes meeting two conditions: (1) the change cannot be comprehensively explained (predicted) by physical laws (breaks the causal closure of matter) and (2) it has causal relationship with subjective phenomena that have no neural correlates. I define "interactionism" as a form of dualism stating that at least some forms of conscious behavior are caused by immaterial influence on a creature's body (Descartes, 1641; Eccles, 1994). In a simplified sense, interactionism is the doctrine that the immaterial soul controls the material body. In contrast, "epiphenomenalism" (in this paper, not distinguished from parallelism) is a form of dualism stating that all basic forms of behavior are possible without immaterial influence on a creature's body (Hodgson, 1870; Huxley, 1874; Leibniz, 1720). In simplified sense, epiphenomenalism is the doctrine that the immaterial soul exists but does not control the material body. Most authors do not discuss the possibility that an immaterial influence on the human body exists, but that it is not necessary for any important form of behavior. In this paper, I consider this as a special form of epiphenomenalism, because such influence does not play the functional role supposed by classical interactionist authors such as Descartes. However, this is merely a terminological choice.

There are three forms of materialism. "Information-based materialism" (computationalism) is a form of materialism stating that consciousness is a purely informational (functional/computational) object or process (Dennett, 1990; Fodor, 1975; Putnam, 1967). "Substrate-based materialism" is a form of materialism stating that consciousness is a property of a special physical process (chemical, electric, quantum etc.) or of physical substance itself (Anokhin, 1974; Davidson, 1970; Hameroff, 2006; Ivanov, 1998). "Eliminative materialism" is a form of materialism stating that consciousness (or, at least, its problematic properties) is a pseudo-problem, while all cognitive processes have a physical nature (Churchland and Churchland, 1981).

Crucial Examples of the Problematic Properties of Consciousness

Qualia are prominent examples of the problematic property of consciousness. The existence of qualia is often used as an argument against materialism: we are unable to describe qualitative properties of our perception and imagination verbally. For example, we can't explain the essence of red to a human with color blindness (see anti-materialist "knowledge argument" [Jackson, 1982]). And even in the case of normal color vision, we do not know exactly how another person perceives the red color. Maybe the person just calls it "red" but subjectively perceives it as we perceive the blue color (see "inverted spectrum argument" [Shoemaker, 1982]). Discussions on Jackson's, Shoemaker's and other arguments have shown that the existence of qualia does not refute materialism; it rather refutes materialism's most primitive reductive forms. However, it is still unknown whether brain contains comprehensive correlates of qualia or not. The physical world seems to be purely quantitative (space, time, mass, energy, other measurable values). If qualia have neural correlates, then these correlates are supposed to be very special physical phenomena. Materialistic hypotheses on the nature of qualia have been suggested, for example, in Hayek (1952), where qualia are related to functional properties of neuronal analyzers; in Anokhin (1974) and Chuprikova (1985), where qualia are related to neuronal chemistry; and in Hameroff (2006) and Ivanov (1998), where qualia are related to quantum states. However, none of these views is universally accepted today. It should be noted that if neural correlates of qualia exist, then qualia can be comprehensively measured and modified by objective methods. It may be even possible (in principle) to develop equipment, which one person could connect to another person (or to Nagel's [1974] bat) and feel all the other's sensations including unarticulated qualitative content.

The second example of the problematic property of consciousness is the unity of consciousness (binding). Consciousness of a human contains subjective phenomena produced by several sensory organs in a single observer, "self." These phenomena constitute a unified conscious experience of a single person. Self seems to be a fundamentally indivisible thing. Descartes supposed that the unity of self cannot be explained in physical terms, and today this unity still remains unexplained. Brain and its information processes do not demonstrate any fundamental unity that might be interpreted as a neural correlate of the unity of consciousness (Bayne and Chalmers, 2003). Collective quantum effects (Hameroff, 2006), membrane potential oscillation synchrony (Crick, 1995), single-cell (Edwards, 2005; Sevush, 2006) and even single-electron (Argonov, 2012) consciousness concepts have been suggested to explain the unity of consciousness. However, all of these hypotheses remain controversial.

The third example of the problematic property of consciousness is the subjectivity of a particular creature (zombie problem). The existence of consciousness in a particular creature (its subjectivity, sentience) is also often supposed to be an objectively undetectable parameter. In particular, epiphenomenalism supposes that some systems might be twin zombies (unconscious creatures structurally and functionally indistinguishable from a given human). The property of being associated with a particular material system is perhaps the most important problematic property of consciousness. If consciousness has a neural correlate, then it is possible to develop a scientific test for it (applicable to arbitrary systems, including artificial intelligence).

Postulates

Here I declare the postulates of my approach. I provide some argumentation for them but do not pretend that I prove them. They are, rather, based on common sense. All studies suggested in this paper are correct only in the framework of the postulates used. This should not be considered as a drawback, because the explicit appearance of postulates (although controversial) makes the analysis more transparent. It should be also noted that none of the postulates presumes any particular theory of consciousness.

> Postulate 1–1. In order to produce detailed phenomenal judgments about problematic properties of consciousness, an intelligent system must have a source of knowledge about the properties of consciousness.

This postulate is based on the hypothesis that problematic properties of consciousness are so complex that occasional production of detailed phenomenal judgments on them is almost impossible (for example, I neglect the possibility that a random algorithm is able to reproduce the books of Descartes and Leibniz). Direct or indirect causation between someone's consciousness and phenomenal judgments is required. I do not state that each phenomenal judgment is caused by the consciousness of a speaking human. Alternatively, phenomenal judgment might be based on knowledge taken from a book written by another person (see "non-eliminative materialism" panel in Figure 1). Moreover, phenomenal judgment might be caused not only by a human consciousness but also by the God who created it (see "dualism" panel in Figure 1). However, at least indirect causation (correlation between problematic properties of consciousness and phenomenal judgments established by a third factor in the past) must exist (see Appendix for additional discussion).

> Postulate 1–2. There are only five basic sources of phenomenal judgments on problematic properties of consciousness. Source 1: neural correlates

of problematic properties of consciousness (producing a phenomenal judgment, a creature describes its own brain structure or functions). Source 2: cognitive errors (producing a phenomenal judgment, a creature describes pseudo-problems). Source 3: immaterial influence on a creature's body. Source 4: innate knowledge (causally related to someone's consciousness). Source 5: external material sources such as discussions and books

Note that Sources 1, 2, and 4 are related to brain structure based on genetic information. Therefore, in some sense, all these sources might be called "innate." The difference is their causal relation to consciousness. Source 1 is directly related to the neural correlate of a creature's own consciousness. Source 4 is related to someone's consciousness and may exist even in an unconscious creature. Source 2 is not related to anyone's consciousness.

Postulates 1–1 and 1–2 are very important in the zombie problem. According to these postulates, zombies (unconscious creatures with normal human behavior) can produce correct phenomenal judgments on problematic properties of consciousness only if they have knowledge about these properties. Source 1 cannot provide such information, because a zombie's internal structure is unconscious and cannot produce self-describing phenomenal judgments. Source 2 might provide some phenomenal judgments, but not about all problematic properties of consciousness (otherwise, eliminative materialism is true, and the term "zombie" is incorrect). Therefore, a zombie requires Sources 3–5 to produce correct phenomenal judgments on at least some problematic properties of consciousness.

Postulate 2–1. If Sources 1–2 are able to provide phenomenal judgments on all known/hypothesized problematic properties of consciousness (in the absence of other Sources), then materialism is true.

Postulate 2–2. In particular, if Source 2 is able to provide phenomenal judgments on all known/hypothesized problematic properties of consciousness, then eliminative materialism is true.

These two postulates are based on the fact that all known arguments against materialism are related to problematic properties of consciousness. By the definition, dualism is true, if at least one property of consciousness has no neural correlate. If, however, all known problematic properties of consciousness have neural correlates (i.e., can be comprehensively studied by objective methods), or simply not exist (related to pseudo-problems), then there is no reason to suppose that consciousness has immaterial "degrees of freedom." The materialistic solution is simpler than the dualistic one. Therefore, caeteris paribus, it is preferable (see the Appendix for additional discussion).

The theoretical basis for the study of the mind–body problem can be summarized as follows. Production of phenomenal judgments due to description of

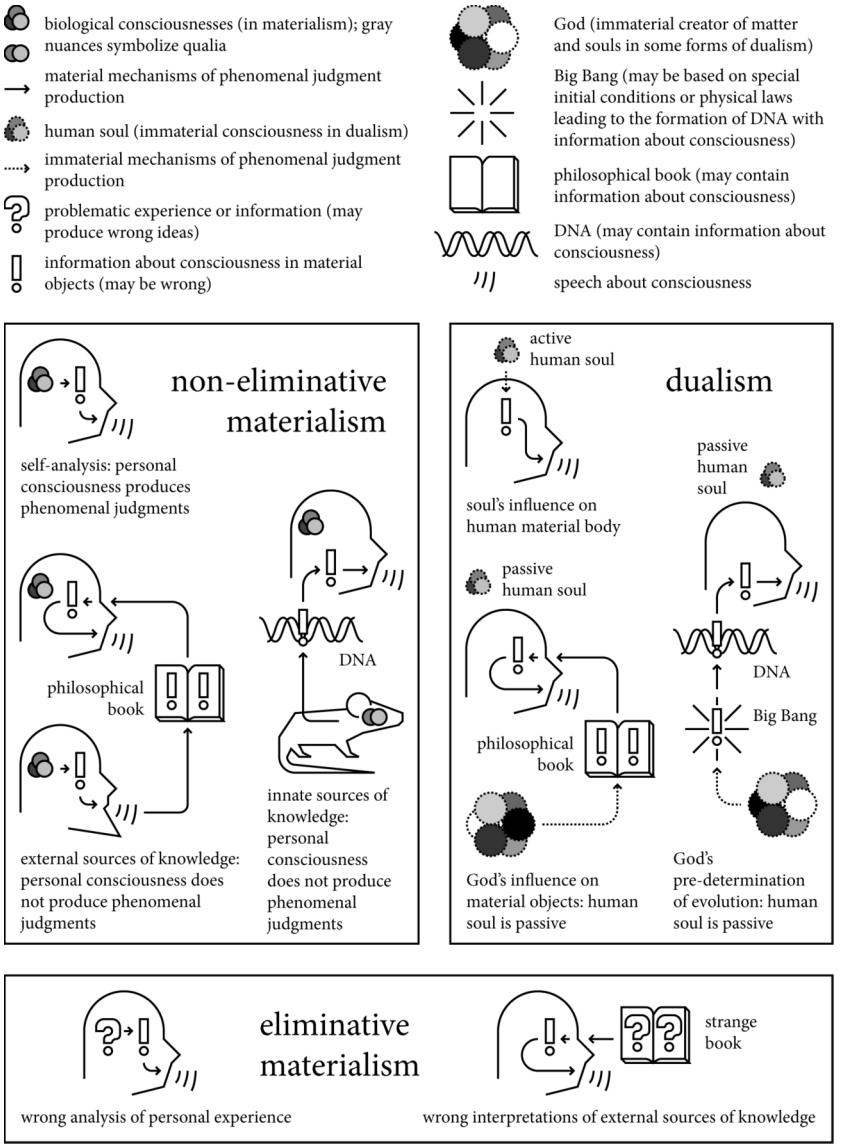

Figure 1: Examples of phenomenal judgment mechanisms in materialism (all sources of knowledge about consciousness are material) and dualism (some sources are immaterial).

the brain's properties (Source 1) is an argument for non-eliminative materialism (see Postulate 2–1). Production of phenomenal judgments due to cognitive errors (Source 2) is an argument for eliminative materialism (see Postulate 2–2). Production of phenomenal judgments due to immaterial influence (Source 3) is an argument for interactionism (see the definition of interactionism). Production of phenomenal judgments due to external or innate sources of knowledge (Sources 4–5) is possible in all theories of consciousness. Complex study of phenomenal judgment mechanisms can give not only "arguments" but also direct scientific verification and falsification of some theories of consciousness (see the next section). In Figure 1, examples of phenomenal judgment mechanisms are shown. Note that external and innate sources of knowledge might be very sophisticated (from philosophical books written by other people to innate knowledge given by God).

Scientific Methods for the Study of the Mind–Body Problem

In this section, I suggest a series of "studies" (some of them may be also called "tests") focused on the solution of the mind–body problem. Studies 1–3 are the studies of classical computers and computer programs. Classical (deterministic) computations can't be affected by immaterial factors without obvious errors. It is possible to monitor a computer's memory and check that it works according to algorithms (that can't be changed by such factors). Therefore, Source 3 can be easily eliminated in Studies 1–3. However, these studies require very powerful equipment (an intelligent computer). Studies 4–5 are the studies of real humans who definitely have consciousnesses. Real humans have very complex structure. Their physiology involves informational, biophysical, chemical, and, supposedly, even quantum processes. The human brain has no clear algorithm, and there is no simple method to eliminate immaterial influences (that might affect indeterministic quantum processes in cells [Eccles, 1994; Stapp, 1993]).

Study 1. Detection of Phenomenal Judgments Produced by a Deterministic Computer

This subsection describes a simple test that can verify materialism (rigorously[1]), detect a machine's consciousness (rigorously), and verify particular information-

[1]I do not claim that the suggested studies can prove something in the ideal mathematical sense. Science collects fragmentary experimental facts and interpolates among them. The precision of any experiment is limited and approximations are always present. In this and further subsections the word "rigorous" will mean that some philosophical statement is a logical consequence of some scientific result and postulates (see previous section). This does not mean that a scientific result itself (or postulates) can be rigorously proven.

based materialistic theories of consciousness (non-rigorously). Consider a deterministic (non-quantum) intelligent machine (a computer or a robot) having no innate (preloaded) philosophical knowledge or philosophical discussions while learning. Also, the machine does not contain informational models of other creatures (that may implicitly or explicitly contain knowledge about these creatures' consciousness). If, under these conditions, the machine produces phenomenal judgments on all problematic properties of consciousness, then, according to Postulates 1–1, 1–2, and 2–1, materialism is true and the machine is conscious (if consciousness is not a senseless term). The postulates are applicable here because the machine is deprived of innate and external physical sources of information about problematic properties of consciousness (Sources 4–5), while immaterial factors (Source 3) cannot affect a deterministic machine during the computations.

This test was originally introduced in Argonov (2011). It can be employed not only for verification of materialism, but also for verification of particular materialistic theories of consciousness. However, this application is not rigorous. For example, if somebody thinks that "consciousness is equal to self-learning" then they should build a self-learning machine and test it for consciousness by the described method. Positive results (detection of phenomenal judgments) would be an argument but not a rigorous proof of the tested theory because the machine might have other mechanisms responsible for consciousness. However, complex studies involving several machines of different structures can improve the reliability and validity of the result. The same approach may be employed for verification of any informational (computational) theory of consciousness.

There are limitations to this study. The first limitation is that the study can verify only information-based materialism and eliminative materialism. If, however, consciousness is based on chemical, biophysical, or quantum mechanisms (impossible in the classical computer), then the experiment would demonstrate a negative result. The second limitation is that the study can provide only verification, but not falsification of materialism. A positive result proves materialism but a negative result proves nothing. For example, absence of phenomenal judgments may be caused by lack of the machine's intellect, not by absense of consciousness (at least, in a primitive form).

Beyond these fundamental theoretical limitations, there are several technical problems that may distort the results of the test. The first is human mistakes or unconscious actions during the construction of the machine that may create implicit knowledge about consciousness in software or hardware. In principle, the presence of such knowledge can be discovered in the analysis of memory logs (of phenomenal judgment production). However, a very complex monitoring is required. The second problem is that the machine (having no philosophical education) might have problems with verbal formulation of ideas. The same is true for a human who does not know the philosophical ideas of other people.

It is not easy to produce understandable phenomenal judgments if one is unacquainted with existing discourse and terminology. I suppose that it is merely a technical problem, otherwise philosophy would never emerge in human culture. However, this problem limits language that can be used in the test. Problematic terms such as "consciousness" should be avoided. Maybe it would be better to start the test with the discussion about qualia and religious questions. In human society, even children (e.g., myself in childhood) often have religious hypotheses. An operator may ask the machine: "Can you concede that another computer (identical with you) perceives the red color as you perceive the blue one?" If the computer is not conscious, it most likely will answer "red is red, it can't be perceived as blue" or "I can't understand this question." If the computer is conscious, it can answer "yes" or even "yes, it is a difficult problem. I already thought about it." Alternatively, an operator may ask: "Did you ever think that your life could be prolonged after the destruction of your body?" A conscious computer might have such ideas. The third (and maybe the most important) problem is that the machine's design must not be purposefully optimized for the production of phenomenal judgments. If the machine is designed just for passing the test, then such design is an implicit innate knowledge of the problematic properties of consciousness (it is causally related to human knowledge about consciousness). Ideally, the computer should be built according to basic principles of self-learning machines without obvious algorithms of phenomenal judgment production.

In practice, a very intelligent system is needed to produce human-like phenomenal judgments. Real experiments seem to be impossible today due to an absense of intelligent machines with human-like behavior. However, Study 1 can be also performed in thought experiments such as in Argonov (2011), where a self-learning robot seems to be able to produce phenomenal judgments on some problems of consciousness (self, reincarnation etc). However, that robot seems to be unable to understand the problem of qualia.

Study 2. Detection of Phenomenal Judgments Produced by a Deterministic Numerical Model of a Hypothetically Conscious Creature

This subsection describes a more sophisticated test that can verify materialism (rigorously), detect a modeled creature's consciousness (rigorously), and verify particular materialistic theories of consciousness based on calculable processes (non-rigorously). Consider a deterministic computer containing a numerical model of a hypothetically conscious creature. The model does not contain explicit information about philosophical problems of consciousness, but it may be based on some particular theory of consciousness (this theory will be tested in the study). For example, if one supposes that consciousness has a chemical basis (Anokhin, 1974; Chuprikova, 1985), then the creature's model must contain

numerical simulation of appropriate chemical processes. If the model correctly describes processes in the creature's material body and produces phenomenal judgments without external or innate sources of knowledge about consciousness, then the modeled creature is also able to produce phenomenal judgments without external or innate sources of knowledge. Therefore, according to Postulates 1–1, 1–2, and 2–1, materialism is true, and the modeled creature is conscious (if consciousness is not a senseless term). The postulates are applicable here for the same reasons as in Study 1. In Study 2, the modeled creature is deprived from external and innate sources of knowledge in the same manner as the computer is deprived of them in Study 1. Note the important differences between Studies 1 and 2. First, in Study 1 the computer must not contain detailed numerical models of other creatures, while in Study 2 it must contain them. Second, in Study 1 a positive result of the experiment (detection of phenomenal judgments) proves that the computer is conscious; while in Study 2 it proves only that the modeled creature is conscious. Study 2 cannot detect the computer's consciousness because the computer contains the numerical model of another creature, and this does not meet the requirements of Study 1.

Most of the fundamental limitations and technical problems for Study 2 are the same as for Study 1 (the word "computer" should be replaced with "creature model"). However, the important new feature is that Study 2 can verify both information- and substrate-based materialistic theories of consciousness, including even some quantum theories (such as Argonov, 2012; Bernroider and Roy, 2004; Ivanov, 1998). Most physicists consider quantum processes as indeterministic (see subsection Study 5), but indeterministic quantum fluctuations can be simulated on a deterministic computer. If a deterministic model of a quantum computer will produce phenomenal judgments, then the difference between computer-generated pseudo-random fluctuations and real stochastic processes is not important in our study. This study, however, is inapplicable to the theory of Penrose (1994), who supposes that some quantum systems may demontrate incalculable dynamics, and such dynamics are related to consciousness. Such dynamics cannot be simulated in principle.

Study 3. Detection of Phenomenal Judgments in a Deterministic Numerical Simulation of a Human

This study has only a single possible philosophical application: falsification of interactionism (rigorous). Consider a deterministic machine (in particular, a robot) containing a detailed numerical model of a human (including simulation of growth and development) and deprived of external sources of philosophical knowledge (Source 5). However, the robot cannot be completely deprived of innate knowledge (Source 4) because, in principle, such knowledge might exist in a real human. In the beginning of the experiment, the robot contains a model

of a developing embryo. After some time period, the robot begins to move, study language, and communicate with people (avoiding philosophical discussions). If the machine demonstrates normal human behavior (including production of phenomenal judgments on all problematic properties of consciousness), then immaterial influence is not necessary for such behavior, and, by the definition, interactionism is wrong. Even if the robot's phenomenal judgments are based on innate knowledge (Source 4), then a real human also has such knowledge and such phenomenal judgments are a part of "normal behavior."

Study 3 provides only falsification, but not verification of interactionism. A positive result refutes interactionism but a negative result proves nothing. Study 3 is less problematic in its ideology (innate knowledge is not completely prohibited) than Studies 1–2, but it is extremely problematic in pure technological terms. Detailed simulation of human development seems to be much more complex than the construction of basic experimental models of consciousness (in Study 1). Nevertheless, some attempts to create a numerical model of human brain have been already made (Markram, 2006).

Study 4. Study of Human Phenomenal Judgment Mechanisms: The Search for the Neural Correlate of Consciousness and the Search for Cognitive Errors

In contrast with Studies 1–3, this study is not just a "test." It is rather a complex scientific program that, theoretically, can verify materialism (rigorously) and particular materialistic theories of consciousness (non-rigorously).

According to the Postulate 2–1, natural explanation of phenomenal judgment production without Sources 3–5 (see subsection Postulates) can prove materialism in general. The particulars form of materialism can be (less rigorously) determined by the search for neural correlates of problematic properties of consciousness and the study of their role in production of phenomenal judgment. If some brain properties are (1) similar to problematic properties of consciousness and (2) functionally related to the production of phenomenal judgment on them, then these properties, most likely, constitute neural correlates of problematic properties of consciousness. Therefore, non-eliminative materialism is true (phenomenal judgments really describe brain properties, so they are produced by Source 1). If these neural correlates have a purely informational nature, then materialism is information-based. If they are related to physical properties, then materialism is substrate-based. If, however, alternative mechanisms of phenomenal judgment production (not involving brain features similar to problematic properties of consciousness) will be discovered (humans talk about problematic properties of consciousness not because these properties really exist), then there is no reason to believe in the existence of problematic properties of consciousness. These phenomenal judgments contain wrong ideas about consciousness (they are produced by Source 2) and eliminative materialism is true.

Modern science has some arguments for the existence of neural correlate of consciousness (Metzinger, 2000; Wegner, 2003). However, neural correlates of problematic properties of consciousness (such as qualia and unity of consciousness) are not yet found, and science has no comprehensive explanation of phenomenal judgment production. A scheme of a self-educating machine has been demonstrated that seems to reproduce human phenomenal judgment on some philosophical problems (in particular, the idea of soul), but not phenomenal judgment on other problems (in particular, the idea of qualia) [Argonov, 2011].

Study 5. Analysis of Quantum Effects in Phenomenal Judgment Production

The study of quantum effects in the brain has, of course, philosophical implications. In particular, the discovery of a quantum neural correlate of the whole consciousness would prove a substrate-based materialism model (see Study 4). However, I shall discuss only the specific features of research that do not reproduce results of the above-mentioned studies. In particular, I shall show that scientific analysis of quantum effects in the brain may provide falsification (rigorous) and verification (although non-rigorous) of interactionism.

According to the above definitions (see subsection Definitions), interactionism states that: (1) matter is not causally closed ("nonphysical" effects exist), and (2) immaterial degrees of freedom of consciousness have a causal relationship with these effects. In classical mechanics, immaterial influence on matter is impossible: every system evolves according to deterministic equations of motion, and nothing immaterial can change the physical result. Matter contains comprehensive information for the prediction of particle motion with unlimited precision. However, in quantum mechanics, exact predictions are impossible. Most physicists (proponents of indeterministic interpretations such as the "orthodox" Copenhagen interpretation) suppose that random, indeterministic effects are fundamental in quantum mechanics. Some authors (Eccles, 1994; Mensky, 2000; Stapp, 1993) suppose that immaterial consciousness influences random quantum fluctuations in the brain, causing some forms of behavior. These fluctuations might look very similar to random processes but they must contain some additional correlations not predicted by today's quantum theory (otherwise, intelligent control of behavior is either impossible or produced by physical factors without immaterial influence). Therefore, quantum interactionism must be verifiable or falsifiable in principle.

It must be emphasized that not all proponents of quantum consciousness are interactionists. For example, Argonov (2012), Bernroider and Roy (2004), and Ivanov (1998) suggest quantum but materialistic hypotheses of consciousness. The difference between materialistic and interactionistic quantum theories of consciousness is that interactionism regards quantum systems as "antennae" (receiving control signals from the soul) rather than a real thinking mechanism. Interactionism regards brain quantum systems as windows to other realities rather than normal physical systems.

There are three basic ways to refute interactionism. First, interactionism is wrong if indeterministic interpretations of quantum mechanics are wrong. Today, different interpretations of quantum mechanics are considered as experimentally indistinguishable, but the discussion is not yet finished. For example, Kocsis et al. (2011) made an attempt to measure Bohmian trajectories of photons (predicted by a deterministic Bohmian model). Due to methodological reasons, this experiment is not a rigorous refutation of quantum indeterminism, but it gives at least an aesthetical argument for deterministic interpretation. Second, interactionism is wrong if quantum mechanisms do not take part in information processing in the brain. There are several well-known proponents of quantum consciousness (Bernroider and Roy, 2004; Hameroff, 2006; Penrose, 1994), but most researchers are skeptical regarding these theories, and Tegmark (2000) presented quantitative arguments against the idea of quantum computations in neurons. Third, interactionism is wrong if quantum mechanisms take part in information processing, but the role of these mechanisms can be comprehensively explained by existing theories. Of course, it is problematic to "prove" any result, but it is a completely scientific problem.

Verification of interactionism is more problematic than refutation. Interactionism might be true only if quantum systems in the brain act as black boxes: quantum systems perform complex cognitive operations, but known physical laws can't explain their functioning. It must be additionally shown that these nonphysical effects are causally related to human subjective phenomena having no neural correlates. For example, a person thinks, "after a minute, I'll move my hand," and this thought has no neural correlate, but some quantum fluctuations occur and the hand really moves. Another way to verify interactionism is to show that quantum fluctuations produce phenomenal judgments on problematic properties of consciousness. Such research has obvious methodological problems (lack of knowledge about existence of neural correlates; necessity to take subjective reports into account, etc.), but I suppose that they can be softened by the combined use of several approaches described in this paper.

Summary: Research Program and its Limitations

Studies 1–5 constitute a consolidated research program that, in principle, can give a scientific solution to the mind–body problem. Let me summarize three basic possible results of the research. Two of them help determine the particular nature of consciousness, while the third one leaves some issues unclear.

Result 1. Verification of Materialism and Falsification of Dualism

The most important feature of the research program is its ability to verify materialism. Positive results of Studies 1, 2, or 4 can give scientific support to

materialism (it may be even called "proof" in the present theoretical framework) and, therefore, may be a refutation of dualism.

First, materialism is true if a deterministic computer in Study 1 produces phenomenal judgments on all problematic properties of consciousness. Second, materialism is true if a numerical model of a hypothetically conscious creature in Study 2 does the same. The particular form of materialism (informational, substrate, or eliminative) can be determined in the study of many machines and creature models based on different cognitive and physiological mechanisms (see details in subsections Study 1 and Study 2). Third, materialism is true if Study 4 gives natural explanation to human phenomenal judgment production. In particular, if phenomenal judgments about problematic properties of consciousness really describe some brain properties (neural correlates of problematic properties of consciousness), then materialism is true in non-eliminative form. The particular nature of consciousness (informational or substrate) is determined by the physiological nature of the neural correlate of consciousness. If, however, problematic properties of consciousness have no neural correlates, and all related phenomenal judgments are based on some cognitive errors, then materialism is true in eliminative form.

It should be emphasized that any of these results is sufficient for the materialistic solution of the mind–body problem. For example, if the creature model in Study 2 produces the required phenomenal judgments, then materialism is true, and no other studies are necessary.

Result 2. Verification of Interactionism and Falsification of Materialism

Another important feature of the research program is its ability to verify interactionism and falsify materialism. This can be made within the framework of Study 5. If brain quantum effects take a crucial part in the production of phenomenal judgments on problematic properties of consciousness, but the brain does not contain the neural correlates of these properties (quantum systems "receive" the information about problematic properties rather than produce it), then interactionism is true (see subsection Study 5).

Result 3. Falsification of Interactionism

Studies 3 and 5 can also give a partial solution to the mind–body problem. They can give arguments against interactionism without explicit support of the other two alternatives (materialism and epiphenomenalism).

First, if the robot in Study 3 demonstrates "normal" human behavior and produces phenomenal judgments on all problematic properties of consciousness, then interactionism is wrong. Second, if Study 5 shows that brain quantum effects are not related to problematic properties of consciousness, then interactionism is wrong, too.

However, such results are insufficient for the solution of mind–body problem. Other studies are required to make a scientific choice between materialism and epiphenomenalism, and this may be a problem.

Is it Possible to Verify Epiphenomenalism?

The pro-materialistic results of Studies 1, 2, or 4 refute epiphenomenalism because they refute dualism in general. Therefore, epiphenomenalism is, at least, a falsifiable doctrine. However, none of the studies is focused on the verification of epiphenomenalism. The negative results of Studies 1, 2, and 4 (materialism is not proven) combined with the anti-interactionistic results of Studies 3 or 5 (interactionism is refuted), provide an ambiguous solution: either materialism (Studies 1, 2, and 4 have failed because the natural phenomenal judgment mechanism is extremely sophisticated) or epiphenomenalism (Studies 1, 2, and 4 have failed because materialism is wrong).

Rudd (2000) and Valdman (1997) argued that epiphenomenalism is incompatible with phenomenal judgments. However, I do not share this position. Epiphenomenalism is incompatible only with Sources 1 and 3 (see subsection Postulates), so a creature (even a zombie) can produce phenomenal judgments due to other sources of knowledge about consciousness. The only important restriction is that epiphenomenalism must explain the existence of information about immaterial problematic properties of consciousness in physical reality. The radical epiphenomenalist idea that science (describing physical reality) can completely ignore the existence of consciousness (or, at least, its problematic properties) is wrong. Any theory must explain the existence of the philosophy of consciousness in human culture, and this is a hard question for epiphenomenalism. Two alternative explanations are shown in Figure 1 (fragment "dualism," schemas regarding the passive human soul).

The first alternative is that the information about problematic properties of consciousness is implicitly "written" in the material world since its creation. For example, according to Leibniz's idea of "pre-established harmony," God created consciousnesses and matter and synchronized them. In a deterministic paradigm (classical mechanics and deterministic interpretations of quantum mechanics), the initial conditions of the universe contained comprehensive information on the universe's future history. Therefore, it may be supposed that humans speak about mind–body problems because God created a detailed "program" of human behavior including phenomenal judgments.

The second alternative is that the information about problematic properties of consciousness is a result of immaterial influence on various physical objects (other than brain). For example, it may be supposed that God controls the evolution of living creatures causing their mutations and changing genetic information. God created implicit philosophical knowledge in DNA, so we have innate knowledge. Also it may be supposed that God created philosophical

books by macroscopic quantum "miracles" (theoretically possible in indeterministic interpretations of quantum mechanics).

Both these alternatives are highly exotic. My personal position is that they are not very realistic, and Studies 1, 2, 4, or 5 will, most likely, support other doctrines. However, I cannot completely exclude the unclear result (interactionism is refuted but materialism is not proven) from consideration. Then, epiphenomenalism may be almost indistinguishable from materialism. Today's science can't verify the existence of immaterial influence in a distant past. However, in principle, materialism and epiphenomenalism assume different phenomenal judgment mechanisms, so scientific choice between them seems to be only a technical problem. I hope that in future theoretical works, precise scientific tests for such choices will be developed.

Conclusion

I have built a general phenomenal judgment approach to the mind–body problem. My basic idea is that consciousness is physical if phenomenal judgments about problematic properties of consciousness are produced by purely physical mechanisms. I have proposed a detailed research program for verification and falsification of various forms of materialism and dualism. All suggested tests and methods are focused on the study of phenomenal judgment mechanisms (in humans and machines). Study 1 (originally described in Argonov, 2011) also suggests a novel non-Turing test for machine consciousness.

I understand that some aspects of the suggested approach might seem questionable. I appreciate future discussion on this issue. The main goal of this paper is to demonstrate that the experimental study of "hard" problems is possible in principle. And I hope that it will encourage researchers to further study these "unsolvable" issues.

Appendix: Additional Discussion about the Postulates

Commentary to Postulate 1–1

I expect the following objection: some problematic properties of consciousness seem not very complex. For example, the unity of consciousness might be expressed in four words: "consciousness is something whole." However, this is a mistake. In practice, complex discussion is needed to explain to another human the essence of each problematic property of consciousness, and Postulate 1–1 states that only conscious and/or a highly educated creature is able to provide such an explanation. Random programs can generate the statement "consciousness is something whole" but not to repeat the books of Descartes or Leibniz.

Note that Postulate 1–1 is weaker than analogous assumptions suggested by other authors. There have been several attempts to use the phenomenal judgment argument for the formulation of some fundamental postulate about consciousness. In particular, Chalmers (1997) said that at least some phenomenal judgments are fully justified because people are acquainted with the phenomenal states that are the objects of such judgments. Valdman (1997) tried to prove the impossibility of zombies using the phenomenal judgment argument. I soften these ideas, and suppose that some unconscious creatures may also produce phenomenal judgments if they have sources of knowledge about consciousness. These sources might be very sophisticated (for example, Leibniz's "pre-established harmony"). The only thing incompatible with Postulate 1–1 is occasional production of phenomenal judgments on some complex problematic properties of consciousness.

Commentary to Postulate 2–1

Stating that materialism is a preferable theory (in the absence of counter-arguments based on problematic properties of consciousness), I use the positivist principle that any theory should describe phenomena in the simplest manner. This does not mean that, being once "proved," materialism must be regarded as an eternally true idea. If new anti-materialist arguments will appear in the future, then this position might be changed. The same is true in any scientific branch. Postulate 2–1 might seem a strong claim, but it only follows common scientific practice. For example, the energy conservation law is not verified for all objects in the universe; it is verified mainly on Earth. However, until we do not know experimental facts against it, we consider it as a true law.

References

Anokhin, P.K. (1974). System analysis of integrative activity of a neural cell. *Uspekhi Fiziologicheskikh Nauk [Progress in Physiological Sciences], 5,* 5–92.

Argonov, V.Yu. (2011). Is a machine able to speak about consciousness? Rigorous approach to mind–body problem and strong AI. In S. Hameroff (Ed.), *Towards a science of consciousness* (pp. 59–59). Tucson: Center of Consciousness Studies.

Argonov, V.Yu. (2012). Neural correlate of consciousness in a single electron: Radical answer to "quantum theories of consciousness." *Neuroquantology, 10,* 276–285.

Ayer, A.J. (1936). *Language, truth and logic*. London: Victor Gollancz.

Bayne, T., and Chalmers, D. (2003). What is the unity of consciousness? In A. Cleeremans (Ed.), *The unity of consciousness: Binding, integration and dissociation* (pp. 23–58). Oxford: Oxford University Press.

Bernroider, G., and Roy S. (2004). Quantum–classical correspondence in the brain: Scaling, action distances and predictability behind neural signal. *Forma, 19,* 55–68.

Chalmers, D. (1997). *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.

Chuprikova, N.I. (1985). *Psyche and consciousness as brain function*. Moscow: Science.

Churchland, P.S., and Churchland, P.M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy, 78*(2), 67–90.

Crick, F. (1995). *The astonishing hypothesis: The scientific search for the soul.* New York: Scribner.

Davidson, D. (1970). Mental events. In L. Foster and J.W. Swanson (Eds.), *Experience and theory* (pp. 79–101). Amherst: University of Massachusetts Press.
Dennett, D. (1990). Quining qualia. In W. Lycan (Ed.), *Mind and cognition* (pp. 519–548). Oxford: Blackwells.
Descartes, R. (1641). *Meditationes de prima philosophia, in qua Dei existentia et animæ immortalitas demonstrator*. Paris: Apud Michaelem Soly.
Eccles, J.C. (1994). *How the self controls its brain*. Berlin: Springer.
Edwards, J. (2005). Is consciousness only a property of individual cells? *Journal of consciousness studies, 12,* 429–457.
Elitzur, A. (1989). Consciousness and the incompleteness of the physical explanation of behavior. *Journal of Mind and Behavior, 10,* 1–20.
Fodor, J. (1975). *The language of thought*. Cambridge, Massachusetts: Harvard University Press.
Hameroff, S.R. (2006). Consciousness, neurobiology and quantum mechanics: The case for a connection. In J. Tuszynski (Ed.), *The emerging physics of consciousness* (pp. 193–253). Berlin: Springer.
Hayek, F.A. (1952). *The sensory order: An inquiry into the foundations of theoretical psychology*. London: Routledge & Kegan Paul.
Hodgson, S.H. (1870). *The theory of practice. An ethical enquiry in two books*. London: Longmans, Green, Reader, and Dyer.
Huxley, T.H. (1874). On the hypothesis that animals are automata, and its history. *Fortnightly Review, 22,* 555–580.
Ivanov, E.M. (1998). *Matter and subjectivity.* Saratov, Russia: Saratov State University Press.
Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly, 32,* 127–136.
Kocsis, S., Braverman, B., Ravets, S., Stevens, M.J., Mirin, R.P., Shalm, L.K., and Steinberg, A.M. (2011). Observing the average trajectories of single photons in a two-slit interferometer. *Science, 332,* 1170–1173.
Leibniz, G.W.F. (1720). *Lehrsätze über die Monadologie.* Frankfurt: Johan Meyer
Markram, H. (2006). The blue brain project. *Nature Reviews Neuroscience, 7,* 153–160.
McGinn, C. (1989). Can we solve the mind–body problem? *Mind, 98,* 349–366.
Mensky, M.B. (2000). Quantum mechanics: New experiments, new applications, and new formulations of old questions. *Physics–Uspekhi, 43,* 585–600.
Metzinger, T. (2000). *Neural correlates of consciousness: Empirical and conceptual questions.* Cambridge, Massachusetts: MIT Press.
Nagel, T. (1974). What is it like to be a bat? *Philosophical Review, 83, 435–450.*
Nemes, T. (1969). *Cybernetic machines.* Budapest: Iliffe Books.
Penrose, R. (1994). *Shadows of the mind: A search for the missing science of consciousness.* Oxford: Oxford University Press.
Place, U.T. (1956). Is consciousness a brain process? *British Journal of Psychology, 47,* 44–50.
Place, U.T. (1960). Materialism as a scientific hypothesis. *Philosophical Review, 69,* 101–104.
Putnam, H. (1967). The nature of mental states. In W.H. Capitan and D.D. Merrill (Eds.), *Art, mind, and religion* (pp. 37–48). Pittsburgh: Pittsburgh University Press.
Rudd, A. (2000). Phenomenal judgment and mental causation. *Journal of Consciousness Studies, 7,* 53–66.
Searle, J. (1980). Minds, brains and programs. *Behavioral and Brain Sciences, 3,* 417–424.
Sevush, S. (2006). Single-neuron theory of consciousness. *Journal of Theoretical Biology, 238,* 704–725.
Shoemaker, S. (1982). The inverted spectrum. *Journal of Philosophy, 79,* 357–381.
Stapp, H. (1993). *Mind, matter, and quantum mechanics*. Munich: Springer.
Tegmark, M. (2000). Importance of quantum coherence in brain processes. *Physical Review E, 61,* 4194–4206.
Valdman, M. (1997). Will zombies talk about consciousness? The paradox of phenomenal judgment: Its implications for naturalistic dualism and other theories of mind. *Journal of Experimental and Theoretical Artificial Intelligence, 9,* 471–490
Wegner, D.M. (2003). *The illusion of the concsious will*. Massachusets: MIT Press.

Critical Notices
Book Reviews
Book Notes

The Journal of Mind and Behavior
Winter and Spring 2014, Volume 35, Numbers 1 and 2
Pages 71–82
ISSN 0271–0137

Radicalizing Enactivism: Basic Minds without Content. Daniel D. Hutto and Erik Myin. Cambridge, Massachusetts: MIT Press, 2013, 206 pages, $ 35.00 hardcover.

Reviewed by Tom Froese, National Autonomous University of Mexico

---

Increasing numbers of philosophers of mind and cognitive scientists are jumping on the embodied cognition bandwagon. Accordingly, mind is no longer viewed as locked away in some Platonic realm of pure logic, as the computational theory of mind has traditionally proposed. Instead, mind has become identified with purposeful activity in the world, an activity that is realized by the body, extended by usage of tools, and scaffolded by a sociocultural environment.

The enactive approach initiated by Varela, Thompson, and Rosch (1991) has done much to develop this new perspective, especially by replacing the traditional emphasis on speculation about supposed subpersonal mental representations with a phenomenological analysis of actual first-person lived experience. More recent efforts in the enactive tradition have given the computational theory of mind, according to which humans are nothing but sophisticated robots, another severe blow. The rejection of representationalism now explicitly goes hand in hand with an acceptance of the living, and therefore mortal, material body as the original foundation of mind and its sense-making activities (Thompson, 2011). On this view, mind is embodied in a living body that is said to be autopoietic, i.e., a materially self-producing network of processes that is self-maintaining under far-from-equilibrium conditions, and organisms therefore lead an existence that is both autonomous and precarious (Froese and Stewart, 2010). This return to the concrete phenomena of life and mind leaves no room for representationalism or for functionalism, either (Di Paolo, 2009a).

It is hardly surprising, then, that this enactive approach is hard to accept for the majority of researchers. What is needed is a careful philosophical account that shows that these conclusions do not entail a rejection of cognitive science as such, but rather that they offer solutions to persistent problems that are inherent in its traditional framework. Hutto and Myin's proposal of a "Radical Enactive (or Embodied) Cognition" (REC), which they defend at length in the 2013 book *Radicalizing Enactivism: Basic Minds without Content*, makes an important step in this direction. The remarkable virtue of Hutto and Myin's proposal is that they are familiar with, and sympathetic to, the aims of traditional cognitive science of providing a naturalized theory of intentionality

---

Thanks to Leslie Marsh and Raymond Russ for helpful comments on an earlier draft. Correspondence concerning this article should be addressed to Dr. Tom Froese, IIMAS–UNAM, Apartado Postal 20–126, Col. San Ángel, DF, 01000, Mexico. Email: t.froese@gmail.com

and meaning. This starting point allows them to invite everyone working in traditional cognitive science onboard the embodied cognition bandwagon, while nevertheless firmly insisting that all of their baggage with representational content must be left behind. They do this effectively and with style. However, it seems that that they neglected to plan the next stops on their radical journey. The concept of mental representation fulfills a psychological need of scientists more than doing any real explanatory work in science, so demonstrating its philosophical uselessness will not simply lead to its abandonment — unless another framework that can account for lived meaning in a material world is on offer.

Hutto and Myin repeatedly suggest that a basic mind consists in an organism's interest-driven ways of skillfully interacting with the world, but they remain vague about the details of this alternative. What constitutes an organism? An interest? A skill? An interaction? No answers are provided. Instead Hutto and Myin distance their REC approach from autopoietic enactivism, i.e., precisely the framework that systematically addresses these kinds of questions. A united front is mutually desirable: to clear autopoietic enactivism of its occasionally careless assertions, and to ground contentless basic minds in biological embodiment.

Summary of the Arguments

Hutto and Myin's book is a manual for radicalizing enactivism by getting rid of the last vestiges of representationalism that afflict embodied, extended, and enactive cognitive science. The stakes are high. If they are right, "basic cognition is not contentful; basic minds are fundamentally, constitutively already world-involving. They are, as we say, extensive" (p. 137). Hutto and Myin thereby turn the extended mind hypothesis, namely that cognition (only sometimes) becomes extended into the world during certain actions, on its head: "coupled activities are the ultimate basis of the decoupled ones, not the other way around" (p. 153). Hutto and Myin also invert the classic cognitivist idea that some kind of language of thought is the basis of cognition. Instead they assign the primary origin of abstract cognition to the skillful use of external symbol systems: "the capacity to engage in decoupled contentful activities is derived, in both a logical and a developmental sense, from activities that involve the manipulation of external vehicles. Scaffolded activities involving external symbols undoubtedly transform and augment cognition" (p. 153).

Related claims are familiar from other varieties of enactive cognition, which also argue for a distributed foundation of life and mind and a socio-cultural basis for detached forms of cognition (Froese and Di Paolo, 2011). But there is always room for improvement and clarification. Hutto and Myin convincingly demonstrate that other enactivisms still have to fully clean up their act: a consistent non-representationalism requires that we give up all talk of a "content" of perceptual experience. These are already useful clarifications. However, Hutto and Myin have a much bigger target in sight. They aim to do nothing less than to dismantle the very foundations of orthodox philosophy of mind and cognitive science by demonstrating that naturalized information does not entail naturalized meaning. The central idea of this move is familiar from the roots of autopoietic enactivism (e.g., Maturana, 1974; Maturana and Varela, 1987; von Foerster, 1980), but Hutto and Myin give it an additional twist. They aim to put an end to representationalism once and for all by robbing it of all plausibility, namely by — and this is where things get interesting — deconstructing it from within, on its own terms.

This is an eminently valuable contribution. Other enactivisms have been too busy developing an alternative framework such that, apart from brash dismissals of representationalism scattered throughout the literature, they have not to given much attention to devising an attentive and sustained critique of the representationalist status quo.[1] Moreover, Hutto and Myin should be commended for devising a jujitsu strategy: rather than directly opposing representationalism, they derive the force of their arguments by taking the goals of representationalism seriously and then turning these goals against representationalism by showing that, time and time again, its appeals to content are doomed to failure. The inevitable conclusion: content is a hindrance, rather than an aid, for naturalizing the mind. How so?

Elsewhere Hutto (2013b) has argued against the possibility of a retreat to action-oriented representations. In addition, Hutto and Myin (in press) focus on putting an end to neural representations. But these cases are mere sideshows compared to the "masterstroke" (Hutto, 2013b, p. 146) attempted in the current book. In a nutshell, Hutto and Myin agree with mainstream consensus that the most worthy attempt to naturalize semantic content is to appeal to principles of information theory, such as Dretske's (1988) teleofunctional account of representation. Indeed, the concept of information is already well established throughout the natural sciences and it can be well defined in a purely objective manner as the lawful covariance of two phenomena. For instance, a relationship of covariance holds between the rings in a tree's trunk and that tree's age. This entails that the rings are informative about the tree's age for those who know about the covariance relationship.

And here lies the root of the problem, for the covariance relationship does not contain any semantic content by itself. One of Hutto and Myin's key arguments is that "there is no naturally occurring contentful information that can be 'used and fused' to form inner representations" (p. 70). To be informative or meaningful, the covariance relationship first requires an act of interpretation by someone. Moreover, since no pre-packaged content exists in the natural environment, this interpretive act does not consist in information processing, at least not if this processing is understood as the receiving of external informational contents as *in-put*, which are transformed into other mental content as *out-put*. The popular metaphor that there are bits of information in the world that must be transferred into mind-as-container via perception-as-information-processing is misleading (for a similar epistemological critique by one of the forefathers of enactivism, see von Foerster, 1980).

By arguing that a naturalized concept of information is not contentful, Hutto and Myin force representationalism to focus on the active role of the interpreter. For example, Millikan's (2005) theory of teleosemantics, which appeals to evolution by natural selection so as to account for semantic content, is modified such that cognitive agents are better conceived of as "content-creating systems" and not as "content-consuming systems" (p. 76). They note that this already comes closer to existing enactive notions of "sense-making" and "meaning generation" (see also Hutto, 2011). But it is not enough. Hutto and Myin's final push is to get rid of the metaphor of the production of semantic content altogether, and thus to settle on a suitably modified theory of "teleosemiotics" (p. 78) that aims to explain why basic minds exhibit a directedness toward certain aspects of the environment.

---

[1]Of course there are exceptions. For example, enactivism draws inspiration from Dreyfus' (1992) systematic critique of the computational theory of mind. For more recent critical assessments we can refer to Harvey (2008) and Gallagher (2008).

> With important adjustments, much can be salvaged from attempts to naturalize representational content. For example, although teleosemantic accounts fail to provide an adequate basis for naturalizing intensional content, they provide adequate tools for making sense of something more modest — i.e., responses involving only intentionality. (p. 80)

In the rest of the book Hutto and Myin explore the implications of their rejection of theories of semantic content in favor of what they call a teleosemiotic theory of life and mind. For instance, they assess what radical enactivism implies for our understanding of perceptual experience and illusions (Chapters 5 and 6), the extended mind hypothesis (Chapter 7), and consciousness (Chapter 8). In what follows I will not go through these remaining chapters one by one. A lot of the book speaks to current debates in analytic philosophy and orthodox cognitive science, and others will be better qualified to review those arguments. Instead I will touch on some of the topics that are geared toward other enactivisms, and I will do so from the perspective of someone who is already convinced by non-representationalism. In particular, I will review Hutto and Myin's REC from the perspective they refer to as "autopoietic enactivism" (e.g., Di Paolo, Rohde, and De Jaegher, 2010; Froese, 2012; Thompson, 2007, 2011).

A Dialogue between Radical and Autopoietic Enactivism

Hutto and Myin discuss two misgivings that they have with regard to autopoietic enactivism (pp. 32–36). The first has to do with how the perceiver's relationship to the world is conceived. If the world as perceived does not represent reality as it is, what is the alternative? Enactive approaches to perception sometimes speak as if the world as perceived is constructed by the perceiver's embodied action. However, it is not clear what this constructive relationship consists in if there is no perceptual content. The second misgiving is related to the vocabulary used by autopoietic enactivism to describe basic minds. Hutto and Myin concur that the doings of creatures with basic minds are to be situated somewhere in between mindless mechanism and full-blown planned action, but where exactly should we draw the lines? And how are we supposed to describe such intermediary forms of life without falling into one or the other extreme? In what follows I intend to address these questions by putting radical and autopoietic enactivism into a mutually beneficial dialogue.

The perceived world is neither represented nor constructed. Hutto and Myin are in agreement with Varela, Thompson, and Rosch's (1991) rejection of theories of mind that posit mental representations as fundamental to mentality. But they remain neutral with regard to some of the other claims by Varela and colleagues, for example that we "enact," "constitute," or "bring forth" the world as we perceive it. This neutral stance is understandable, because there exists considerable ambiguity in the primary literature about what is precisely meant by these phrases. This may have to do with a lingering internalism when it comes to thinking about perceivers. Indeed, there is a growing realization that the pioneering work in enactivism, despite overt claims to the contrary, failed to fully overcome the internalism of mainstream theories of mind. For example, in Maturana and Varela's (1987) biology of cognition, which was one of the most important precursors to Varela's enactivism (Froese, 2010), we find a comparison of the brain with a submarine navigator.

> All that exists for the man inside the submarine are indicator readings, their transitions, and ways of obtaining specific relations between them. [. . .] The dynamics of the submarine's different states, with its navigator who does not know the outside world, never

> occurs in an operation with representations of the world that the outside observer sees [. . .]. Entities such as beaches, reefs, or surface are valid only for an outside observer, not for the submarine or for the navigator who functions as a component of it. What is valid for the submarine in this analogy is valid also for all living systems [. . .]. (Maturana and Varela, 1987, p. 137)

It is worthwhile dwelling on this rather strange analogy for a moment because in Varela's afterword to this book he explicitly emphasizes its continuity with his "enactive" approach (p. 255)[2], which is there expressed in one of its earliest formulations. Ironically, Maturana and Varela made use of the classic homunculus inside a black box in order to argue against the notion of internal mental representations. To be sure, their motivations are different from traditional cognitive science, since this brain-centered internalism derives from their desire to emphasize the autonomy of the organism. But it is an internalism nonetheless. And this insight provides us with the key for unlocking the meaning of the obscure phrase "bringing forth a world," since an internalism that rejects representationalism must by default become some kind of constructivism. This is not to deny the essential differences between Maturana and Varela's biology of cognition and, for example, von Glasersfeld's (1995) radical constructivism (see, e.g., Proulx, 2008). But suspicions that Maturana and Varela's autopoietic theory of life is internalist (Wheeler, 2010), and that Varela's neurophenomenology has a lingering internalism (Beaton, 2013) and an idealist streak (Pascal and O'Regan, 2008), are not entirely unfounded.

And yet at the same time we find that Maturana and Varela refer to an "outside observer" who sees the submarine in its reef environment — so, is our conscious mind situated directly in the world after all? This possibility is indicated by our pre-reflective experience of being-in-the-world, which became a cornerstone of Varela's enactive turn (Froese, 2011). Maturana and Varela's submarine analogy is an illustration of their doctrine of non-intersecting domains, which is intended to ensure both the autonomy of the organism and the relational perspective of the observer. But in this operationalized mind–body dualism we find the origins of a tension that is only now starting to be overcome by autopoietic enactivism (Di Paolo, 2009a, 2009b). For without a complete rejection of internalism, even if it is just a part of the operational story, there will always be a temptation to fall back on a radical constructivism, or even representationalism, to somehow fill in for the absence of the world itself.

Perception is constitutively world-involving. Putting enactivism on more stable middle ground requires rejecting even the last hidden remnants of internalism (Beaton, 2013). In this regard Hutto and Myin's insistence that basic minds are extensive by nature (and not just contingently extended) is a step in the right direction. But if basic minds are extensive, how shall we conceive of perception? Hutto and Myin promote the idea that "minds, quite generally, are best understood in terms of capacities" (p. 151). This stance is reminiscent of sensorimotor enactivism (e.g., Noë, 2004; O'Regan and Noë, 2001). However, they disagree with the way in which proponents of the sensorimotor approach cash out the idea of mastery of sensorimotor contingencies.

> Although they insist that perception and its experience is based on a kind of know how, they tend to fall into unguarded talk of perceivers' (or their brains') making assumptions,

[2]As far as I know, Maturana has never used the term "enactive" to describe his approach. See Froese (2011) and Froese and Stewart (2010) for a more detailed discussion of the similarities and differences of their approaches.

> predictions, and judgments in ways that look decidedly as if the view is committed to the existence of propositional rather than essentially practical knowledge. (p. 26)

I am in full agreement with this assessment. We may speculate that this implicit affinity between sensorimotor theory and cognitivism has helped to turn O'Regan and Noë's sensorimotor approach into one of the most popular variants of enactivism. This lingering cognitivism is also what has facilitated the wide adoption of the sensorimotor approach to robotics, thus prompting Dreyfus (2007) to launch another updated version of his famous critiques of AI. Indeed, one of the hottest new theoretical developments in cognitivism, so-called predictive processing, is triumphantly formalizing sensorimotor theory in terms of its classic internalist–representationalist framework (Clark, 2012, 2013; Seth, in press). Hutto and Myin are therefore right to be wary of associating REC with standard versions of sensorimotor theory. In order to avoid sensorimotor contingencies becoming a Trojan horse for enactivism they first have to be freed of their implicit cognitive baggage, for example, in terms of a dynamical systems account (Buhrmann, Di Paolo, and Barandiaran, 2013).

What does the rejection of content tell us about the perceiver's relationship to the world? Unfortunately, Hutto and Myin do not make their epistemology of perception sufficiently explicit. But given that (i) they reject representationalist theories of perception, (ii) they reject constructivist claims that embodied action literally brings forth the world as perceived, and (iii) they argue that basic minds are essentially extensive, this leaves only the radical conclusion: veridical perception is constitutively world-involving. Moreover, no appeal to content is necessary to account for the existence of non-veridical experiences.

> Dreaming, visual imagery and experience in paralysis, then, are cases in which the explanatory balance tips more fully in the direction of past sensorimotor contingencies. What one experiences under such circumstances is dictated almost exclusively by one's attunement to previous interactive regularities, rather than by one's current stimulation. (Myin and Degenaar, 2014, pp. 96–97)

Hutto and Myin's REC thus seems to be consistent with a disjunctivist theory of enactive perception (e.g., Beaton, 2013), although any appeals to counterfactual interactions would need to be cashed out in contentless terms, for example in terms of attunement.

The limits of brain–body–environment equality. Normal perceptual experience consists of both mind-dependent as well as mind-independent aspects, i.e., both a subjective perspective and the objective world. However, Hutto and Myin do not account for the constitution of an organism's own perspective on the world. In their eagerness to clear basic minds from all remnants of internalism, Hutto and Myin fall into the opposite extreme, namely the eradication of any difference between the body and the environment.

> To suppose that what is constitutive of mentality must reside in organisms or their brains alone is to endorse a Senior Partner Principle, holding that [. . .] only brains bring mentality to the party. In the place of this, we promote the more even-handed Equal Partner Principle as the right way to understand basic mental activity. Accordingly, contributions of the brain are not prioritized over those of the environment. (p. 137)

However, this assumed absolute equality is not supported by the evidence of our embodiment or of our lived experience. Our living body (including the brain) is more

complex than our immediate environment, and from our first-person perspective we always perceive the world from the point of view of our bodies, and not the other way around. How shall we account for this complexity asymmetry and experiential perspectivalness if there is supposed to be nothing that distinguishes the living body from its environment? This unexplored tension may be one reason why Hutto and Myin do not apply their Equal Partner Principle consistently to all aspects of mentality. At least when it comes to basic minds they are committed to the "strongest reading of the Embodiment Thesis" which leads them to assert:

> If mentality is not at root content involving, there is no reason to suppose, even in principle, that it is possible to individuate and isolate some portion of organismic activity — a portion that falls short of an organism's total way of responding to some worldly offering — that can be identified with properly cognitive activity. (pp. 11–12)

And yet Hutto and Myin back down when it comes to explaining the phenomenal aspects of consciousness. For example, they allow that the "minimal supervenience basis for phenomenality might be narrow" (p. 8). More specifically, they encourage us to "take phenomenality to be nothing but forms of activities — *perhaps only neural* — that are associated with environment-involving interactions" (p. 169; emphasis added). Hutto and Myin thereby abstain from choosing between either aligning their REC approach with more conservative embodied approaches (e.g., Clark, 2009, 2012), or with other enactivisms that apply the Embodiment Thesis in a more consistent manner (e.g., Noë, 2009; Ward, 2012). This lack of commitment to the latter, more radical option is especially puzzling given that Hutto and Myin want to dissolve the hard problem of consciousness precisely by appealing to environment-involving interactions, such as feeling the softness of a sponge by squishing it (p. 177).

Hutto and Myin also remain conservative in allowing that there are basic minds without any phenomenality. They accept that the doings of basic minds exhibits a minimal intentional directedness, but insist that this is not yet sense-making and not yet necessarily accompanied by any phenomenality.

> The more modest claim, which we endorse, is that basic interest-driven ways of responding provide the right platform for understanding how mentality can be intentionally directed yet also wholly embodied and enactive. Certain organisms are not only set up so that they are intentionally directed at situations that can bear on their interests; *in some cases*, their ways of responding are also phenomenally charged. (p. 36; emphasis added)

But this proposal immediately raises a host of difficult questions. What defines a living system's interest? How does its interest-driven responding account for its intentional directedness? How is it possible for a basic mind to pursue interests without any kind of phenomenality, not even a basic sense of concern? A first step toward addressing these issues is to recognize that extensiveness does not necessarily entail equality.

Hutto and Myin proposes "there is no way to isolate properly mentality-constituting 'inner' organismic responses from 'outer' ones that allegedly stand over and against the former as mere causal contributions from the environment" (p. 6). This is in line with autopoietic enactivism, which has also started to argue that organismic activity is not confined within the body (Di Paolo, 2009a; Virgo, Egbert, and Froese, 2011). But autopoietic enactivism is striving for a middle way between the Senior Partner Principle and the Equal Partner Principle. The key idea here is that there is an essential asymmetry between the living body and the environment, because the living body's metabolic

self-construction (autopoiesis) has the active role in bringing about the existence of an individual with world-involving relationships in the first place (Barandiaran, Di Paolo, and Rohde, 2009; Di Paolo, 2005). This relational asymmetry is inherent in biological embodiment; it constitutes an organism's perspective and directs it toward relevant aspects of its environment (Varela, 1997).

Di Paolo (2009a) has pointed out that the circularity of the living, i.e., the fact that an organism's being is its own doing, is at odds with the functionalist aim of providing a substrate-independent account of its operation. The substrate-dependence of the living also constitutes their precariousness, which can be understood as their inevitable mortality and therefore as their original source of concern and care in the world (Weber and Varela, 2002). But mortality has no place in cognitivism, as can be seen from the quasi-religious theories some technology multimillionaires have about immortalizing themselves in a future computer. Enactivists, on the other hand, cannot let themselves be deluded by such functionalist fantasies: "Precariousness does not refer to a positive material property that could be captured functionally, but to the impermanence of any relevant positive property of the substrate" (Di Paolo, 2009a, p. 16). Enactivism speaks to the core of human existence: our time here is limited, but that is precisely what makes each moment meaningful.

Debating the scope and limits of basic minds. This finally brings us to Hutto and Myin's second misgiving about autopoietic enactivism. They are concerned about the facility with which properties of the human mind are generalized to basic minds.[3] For example, they hold that it is not plausible that the simplest living systems are capable of sense-making, which is one of the key tenets of autopoietic enactivism (Di Paolo et al., 2010). Hutto and Myin "prefer the more austere talk of informationally sensitive responses to natural signs" (p. 78). They reject the idea that basic minds are capable of sense-making partly because, for them, talk of creating, generating, or making meaning is closely associated with traditional theories of semantic content. However, even if autopoietic enactivism is cleared of this charge, which may have more to do with its careless language than a commitment to content, there still remains a deeper difference at play. Hutto and Myin disagree that basic minds have a capacity for interpretation (p. 36), a capacity that is presupposed by sense-making.

We can understand the origin of Hutto and Myin's conservatism regarding basic minds in terms of their theory of teleosemiotics, i.e., a combination of Dretske's teleofunctionalism and Millikan's teleosemantics freed from appeals to semantic content. They aim to use this theory to assert that "experiencing organisms are set up to be set off by certain worldly offerings — that they respond to such offerings in distinctive sensorimotor ways that exhibit a certain minimal kind of directedness and phenomenality" (p. 19). Hutto and Myin are also confident that REC can explain how it is that most animal doings consist in "motivated" and "skillful" dealings with the world (p. 50). They therefore suggest that "if REC has the right resources for explaining the wide class of such doings, then it has the potential to explain quite a lot of what matters to us when it comes to understanding mind and cognition" (p. 50). However, it is not clear how their teleosemiotics is supposed to live up to this formidable task.

Hutto and Myin partially go along with Dretske so as to characterize the activity of basic minds as informationally sensitive responding. This allows them to talk about the movements of behavior-based robots and the doings of insects in an essentially interchangeable way (pp. 41–43). This equivalence between robotic systems and living systems further helps to explain why they disagree that basic minds are capable of sense-

[3]See Wheeler (2011) for an extended discussion of related worries.

making: they disagree because they set the bar for basic minds much lower, even including basic mechanical systems. For autopoietic enactivism, talk about basic minds is only applicable to living beings: it does not matter how intelligent the behavior-based robot appears to be, nor does it matter, in direct contrast to a Millikanian theory of intentionality, whether its behavior has been selected by artificial evolution (Froese and Ziemke, 2009).

To be sure, an organism's natural history can provide useful guidance when we want to understand the processes that make the organism responsive to aspects of the environment. But we should not conflate such functional accounts with operational accounts (Varela, 1979). The fact that natural selection played a role in the past cannot explain how an organism's responsiveness is actually realized. Ultimately, only the organism's processes in the present are responsible for its responsiveness. Yet Hutto and Myin fail to offer any account of what constitutes individuality, agency, motivations, and skills, or how it is possible that something matters with respect to an organism's interests.

Autopoietic enactivism has the framework to explain the directed and interest-driven nature of basic minds, and it does so in an operational manner without appealing to semantic content. As mentioned before, this framework is based on autopoiesis as the source of an organism's mortality and its concerned doings (Weber and Varela, 2002). Beginning with Varela's (1992) biology of intentionality, and continuing with Di Paolo's (2005) biology of normativity, there is a tradition that addresses the missing pieces of REC head on (Thompson, 2007, 2011). Motivation and skills are also key themes of autopoietic enactivism (McGann, De Jaegher, and Di Paolo, 2013). Moreover, the foundations of autopoietic enactivism are being put to the practical test in the field of artificial life, for example in studies of the emergence of self-driven adaptive behavior (Froese, Virgo, and Ikegami, 2014) and normativity (Barandiaran and Egbert, 2014).

The extent of the challenge posed by autopoietic enactivism to the traditional foundations of cognitive science can be gauged by the practical difficulties faced by a biologically grounded enactive AI (Froese and Ziemke, 2009). For instance, it is not even clear if the circularity of autopoiesis is computable by a Turing machine, in principle (Letelier, Marín, and Mpodozis, 2003). Similarly, if it is accepted that autopoiesis is constitutive of life and mind, then its precariousness rules out functionalism (Di Paolo, 2009a). Conversely, if this precariousness is constitutive of a meaningful perspective, then functionalism rules out embodied sense-making. The implications are that the bar for basic minds should be set much higher than mere informationally sensitive responsiveness, and that a full rejection of representationalism must be accompanied by a full rejection of functionalism in order to avoid falling into nihilistic behaviorism.

Overcoming the cognitive gap of enactivism. Where does the story go from here? Hutto and Myin face a major challenge that is shared by all enactivism: how to bridge the "cognitive gap" which separates the activity of basic minds from the abstract cognition that is achieved by adult human minds (Froese and Di Paolo, 2009). As Hutto and Myin put it: "The ultimate task is to explain how basic minds make the development of contentful forms of cognition possible when the right supports, such as shared social practices, are in place" (p. 36). However, the viability of such an explanation should not be judged by its success of rehabilitating representationalism at the level of specifically human cognition (Hutto, 2013a). The general idea is to appeal to development in a culturally enriched environment: "The capacity to think using contentful representations is an example of a late-developing, scaffolded, and socially supported achievement. It originates from and exists, in part, in virtue of social practices

that make use of external public resources, such as pen, paper, signs and symbols" (p. 152). Importantly, this capacity to think using contentful representations is not explained in terms of acquiring a capacity to manipulate bona fide internal symbols: "rather, what is gained is an ability to perform operations that previously required the manipulation of external symbols but have now become possible in the absence of external symbols." In other words, scaffolded cognition "becomes (up to a certain degree) independent of context" (p. 152). It seems that explaining higher-level cognition in terms of context-independence is becoming a major theme of enactivism, and I look forward to seeing how Hutto and Myin will develop this aspect in future work.

Conclusions

Hutto and Myin have provided a useful service for cognitive science, and especially for all varieties of enactivism, by deconstructing the foundations of representationalism. They do an admirable job of undermining all attempts of naturalizing representationalism that are based on information theory. They systematically demonstrate that information does not have content and that information processing does not explain semantics. Along the way they also clear up ambiguities in the enactive literature by highlighting the nature of extensive minds without content.

And yet despite Hutto and Myin's apparent radicalness, the dialogue with autopoietic enactivism has also revealed a number of conservative choices. It is interesting to note that Hutto and Myin refer to their approach as "Radical Enactive (or *Embodied*) Cognition," and it often makes sense to see their proposal more as a radicalization of embodied cognition, while retaining its functionalism. Elsewhere Hutto and Myin write:

> There is another possible move that must be avoided [by representationalism]. It is the extreme deflationary maneuver of holding that the representational story is only committed to organismic responses to covariant information. [. . .] to go this way would make representationalism indistinguishable from non-representationalism: the two positions would collapse into the same proposal and they would no longer be rivals. (Hutto and Myin, in press)

This is where Hutto and Myin's REC and autopoietic enactivism part ways. The latter rejects functionalism, and thus breaks with cognitivism altogether. To be sure, this move was not easy even for the other varieties of enactivism. Autopoiesis was largely absent from the first book on the enactive approach to cognitive science (Varela et al., 1991) and was only incorporated later on (Thompson, 2007). Similarly, the first account of sensorimotor enactivism was not specifically restricted to living beings (O'Regan and Noë, 2001), but Noë (2009) now recognizes that "the problem of mind is that of the problem of life" and, even more provocatively, that "the problem of consciousness, then, is none other than the problem of life. What we need to understand is how life emerges in the natural world" (p. 41). Hutto and Myin are clearly hesitant to become a part of this more radical enactive trend. To be sure, neither was it necessary for them to have committed themselves to this deep life-mind continuity in this book, since the book's strategic objective was largely a critique of representationalism. However, it remains to be seen whether their lingering functionalism will allow them to build up a replacement framework, which can successfully answer such questions as: What defines an individual? What defines agency? What defines meaning?

References

Barandiaran, X., Di Paolo, E.A., and Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior, 17,* 367–386.

Barandiaran, X., and Egbert, M.D. (2014). Norm-establishing and norm-following in autonomous agency. *Artificial Life, 20,* 5–28.

Beaton, M. (2013). Phenomenology and embodied action. *Constructivist Foundations, 8,* 298–313.

Buhrmann, T., Di Paolo, E.A., and Barandiaran, X. (2013). A dynamical systems account of sensorimotor contingencies. *Frontiers in Psychology, 4*. doi: 10.3389/fpsyg.2013.00285

Clark, A. (2009). Spreading the joy? Why the machinery of consciousness is (probably) still in the head. *Mind, 118,* 963–993.

Clark, A. (2012). Dreaming the whole cat: Generative models, predictive processing, and the enactivist conception of perceptual experience. *Mind, 121,* 753–771.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36,* 181–204.

Di Paolo, E.A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences, 4,* 429–452.

Di Paolo, E.A. (2009a). Extended life. *Topoi, 28,* 9–21.

Di Paolo, E.A. (2009b). Overcoming autopoiesis: An enactive detour on the way from life to society. In R. Magalhães and R. Sanchez (Eds.), *Autopoiesis in organization theory and practice* (pp. 43–68). Bingley: Emerald Group Publishing Limited.

Di Paolo, E.A., Rohde, M., and De Jaegher, H. (2010). Horizons for the enactive mind: Values, social interaction, and play. In J. Stewart, O. Gapenne, and E.A. Di Paolo (Eds.), *Enaction: Toward a new paradigm for cognitive science* (pp. 33–87). Cambridge, Massachusetts: MIT Press.

Dretske, F. (1988). *Explaining behavior: Reasons in a world of causes*. Cambridge, Massachusetts: MIT Press.

Dreyfus, H.L. (1992). *What computers still can't do: A critique of artificial reason.* Cambridge, Massachusetts: MIT Press.

Dreyfus, H.L. (2007). Why Heideggerian AI failed and how fixing it would require making it more Heideggerian. *Philosophical Psychology, 20,* 247–268.

Froese, T. (2010). From cybernetics to second-order cybernetics: A comparative analysis of their central ideas. *Constructivist Foundations, 5,* 75–85.

Froese, T. (2011). From second-order cybernetics to enactive cognitive science: Varela's turn from epistemology to phenomenology. *Systems Research and Behavioral Science, 28,* 631–645.

Froese, T. (2012). From adaptive behavior to human cognition: A review of enaction. *Adaptive Behavior, 20,* 209–221.

Froese, T., and Di Paolo, E.A. (2009). Sociality and the life-mind continuity thesis. *Phenomenology and the Cognitive Sciences, 8,* 439–463.

Froese, T., and Di Paolo, E.A. (2011). The enactive approach: Theoretical sketches from cell to society. *Pragmatics & Cognition, 19,* 1–36.

Froese, T., and Stewart, J. (2010). Life after Ashby: Ultrastability and the autopoietic foundations of biological individuality. *Cybernetics and Human Knowing, 17,* 7–50.

Froese, T., Virgo, N., and Ikegami, T. (2014). Motility at the origin of life: Its characterization and a model. *Artificial Life, 20,* 55–76.

Froese, T., and Ziemke, T. (2009). Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artificial Intelligence, 173,* 366–500.

Gallagher, S. (2008). Are minimal representations still representations? *International Journal of Philosophical Studies, 16,* 351–369.

Harvey, I. (2008). Misrepresentations. In S. Bullock, J. Noble, R.A. Watson, and M.A. Bedau (Eds.), *Artificial life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems* (pp. 227–233). Cambridge, Massachusetts: MIT Press.

Hutto, D.D. (2011). Philosophy of mind's new lease on life: Autopoietic enactivism meets teleosemiotics. *Journal of Consciousness Studies, 18,* 44–64.

Hutto, D.D. (2013a). Enactivism, from a Wittgensteinian point of view. *American Philosophical Quarterly, 50,* 281–302.

Hutto, D.D. (2013b). Exorcising action oriented represenations: Ridding cognitive science of its Nazgûl. *Adaptive Behavior, 21,* 142–150.
Hutto, D.D., and Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. Cambridge, Massachusetts: MIT Press.
Hutto, D.D., and Myin, E. (in press). Neural representations not needed — no more pleas, please. *Phenomenology and the Cognitive Sciences.* doi: 10.1007/s11097–013–9331–1
Letelier, J.C., Marín, G., and Mpodozis, J. (2003). Autopoietic and (M,R) systems. *Journal of Theoretical Biology, 222,* 261–272.
Maturana, H.R. (1974). Cognitive strategies. In H. von Foerster (Ed.), *Cybernetics of cybernetics* (pp. 457–469). Urbana: Biological Computer Laboratory.
Maturana, H.R., and Varela, F.J. (1987). *The tree of knowledge: The biological roots of human understanding.* Boston: Shambhala Publications.
McGann, M., De Jaegher, H., and Di Paolo, E.A. (2013). Enaction and psychology. *Review of General Psychology, 17,* 203–209.
Millikan, R.G. (2005). *Language: A biological model*. New York: Oxford University Press.
Myin, E., and Degenaar, J. (2014). Enactive vision. In L. Shapiro (Ed.), *The Routledge handbook of embodied cognition* (pp. 90–98). New York: Routledge.
Noë, A. (2004). *Action in perception*. Cambridge, Massachusetts: MIT Press.
Noë, A. (2009). *Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness*. New York: Hill and Wang.
O'Regan, J.K., and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24,* 939–1031.
Pascal, F., and O'Regan, K. (2008). Commentary on Mossio and Taraborelli: Is the enactive approach really sensorimotor? *Consciousness and Cognition, 17,* 1341–1342.
Proulx, J. (2008). Some differences between Maturana and Varela's theory of cognition and constructivism. *Complicity: An International Journal of Complexity and Education, 5,* 11–26.
Seth, A.K. (in press). A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synaesthesia. *Cognitive Neuroscience*. doi: 10.1080/17588928.2013.877880
Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind.* Cambridge, Massachusetts: Harvard University Press.
Thompson, E. (2011). Reply to commentaries. *Journal of Consciousnes Studies, 18,* 176–223.
Varela, F.J. (1979). *Principles of biological autonomy*. New York: Elsevier.
Varela, F.J. (1992). Autopoiesis and a biology of intentionality. In B. McMullin and N. Murphy (Eds.), *Autopoiesis and perception: A workshop with ESPRIT BRA 3352* (pp. 4–14). Dublin: Dublin City University.
Varela, F.J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition, 34,* 72–87.
Varela, F.J., Thompson, E., and Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, Massachusetts: MIT Press.
Virgo, N., Egbert, M.D., and Froese, T. (2011). The role of the spatial boundary in autopoiesis. In G. Kampis, I. Karsai, and E. Szathmáry (Eds.), *Advances in artificial life: Darwin meets von Neumann. 10th European Conference, ECAL 2009* (pp. 234–241). Berlin: Springer Verlag.
von Foerster, H. (1980). Epistemology of communication. In K. Woodward (Ed.), *The myths of information: Technology and postindustrial culture* (pp. 18–27). Madison: Coda Press.
von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. London: Falmer.
Ward, D. (2012). Enjoying the spread: Conscious externalism reconsidered. *Mind, 121,* 731–751.
Weber, A., and Varela, F.J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences, 1,* 97–125.
Wheeler, M. (2010). Minds, things, and materiality. In L. Malafouris and C. Renfrew (Eds.), *The cognitive life of things: Recasting the boundaries of the mind* (pp. 29–38). Cambridge: McDonald Institute for Archeological Research.
Wheeler, M. (2011). Mind in life or life in mind? Making sense of deep continuity. *Journal of Consciousness Studies, 18,* 148–168.

The Journal of Mind and Behavior
Winter and Spring 2014, Volume 35, Numbers 1 and 2
Pages 83–92
ISSN 0271–0137

What to Believe Now: Applying Epistemology to Contemporary Issues. David Coady. Oxford: Wiley–Blackwell, 2012, x + 202 pages, $96.95 hardcover, $31.99 paperback.

Reviewed by Andrew Alexandra, University of Melbourne

What to Believe Now: Applying Epistemology to Contemporary Issues is a lively, interesting, and stylishly written book. The author, David Coady, draws from an eclectic mix of epistemological theory to illuminate — albeit sometimes briefly — a range of currently controversial topics. These include the claims of "epistemic democrats" that democracy is better able than other political systems to "track the truth," and the debate about whether votes in democratic elections should be understood as statements, as preferences, or as resources; torture; and government surveillance and privacy. However, the heart of the book, and its most significant contributions, lies in its assessment of the epistemic credentials of a number of sources of popular beliefs, in particular the testimony of experts, rumours, conspiracy theories and the blogosphere. Accordingly, this review will focus on those assessments. Though there is a certain amount of overlap, the content of each is distinct enough to merit individual consideration.

Experts

Let me begin with expert testimony. On the face of it, it seems that we should give more weight to the testimony of experts than to that of laypeople. But what distinguishes the experts? According to Coady "being an expert is simply a matter of being well informed about a subject, that is, having a significantly greater store of accurate information about it than most people . . . " (p. 28). Generally, of course, we should believe experts when they make claims in their area of expertise. But what should we believe when experts disagree? At least in some cases of expert disagreement, there is a clear majority on one side of the dispute. In those cases, perhaps laypeople should accept the view of the majority. Coady considers the salient case of anthropogenic climate change, which exemplifies such a pattern; though there are some dissenting climate scientists, the large majority of climate scientists are convinced of anthropogenic climate change. Is that a reason for the layperson to do likewise? Alvin Goldman (Goldman, 2001, p. 99) and others think that in cases like these, numbers alone should not count (Elga, 2010, p. 177; Kelly, 2010, p. 148). The argument goes as follows. If Bob accepts a claim made

---

Correspondence concerning this review should be sent to Andrew Alexandra, School of Historical and Philosophical Studies, Old Quad, University of Melbourne, Victoria 3010, Australia. Email: a.alexandra@unimelb.edu.au

by Andrew because of Andrew's supposed expertise, Bob becomes what Goldman calls a "non-discriminating reflector" of Andrew. Suppose Bob then retells the claim to Charles, who has previously also heard it from Andrew. According to Goldman, Bob's testimony does not give Charles a further reason to believe the claim, in addition to the reason he already had after hearing the claim from Andrew. To think it does is, in effect, to count the same evidence twice. Now consider a situation where the views of the majority of experts are not formed independently of each other (in the most extreme case, there is only one expert who has formed her view independently, and all the rest of the majority are non-discriminating reflectors of that view) but a minority of experts each independently comes to a view dissenting from the majority view. Here, it may be that the minority view is the one which should be accorded more weight.

Coady argues, I think persuasively, that that conclusion is too swift. He does so by appeal to the notion of meta-expertise — the capacity accurately to identify who possesses expertise in an area (that is, who is more likely to possess true information in that area). If, say, the members of the majority possess such meta-expertise, then even if most of them are non-discriminating reflectors, the fact that they accept the claims of some (or even only one) over others, is a reason to give the claims they accept greater weight than the claims they reject. Since climate scientists obviously possess meta-expertise in their field, then, on this line of reasoning, the fact that most of them believe that anthropogenic climate change is occurring is powerful reason to do likewise, even if the beliefs of most of those scientists depend on claims of a few of their colleagues. And even where particular scientists are not themselves expert in the field of climate science, given their grasp of scientific method, they presumably possess meta-expertise in relation to the claims of those who do (or claim to) possess such expertise. Moreover, the structure of the institution of science, with its various formal markers of scientific expertise (such as academic qualifications, publications in peer-reviewed specialist academic journals etc.) means that even laypeople can possess meta-expertise (or perhaps by this stage meta-meta-expertise) in respect of who has scientific (meta-)expertise. So, contra Goldman et al., even if it were the case that most of the scientists who accept anthropogenic climate change are non-discriminating reflectors of the views of a small number of their peers, and most of the smaller number of scientists who reject it do so because they have considered the evidence on their own account, it is still reasonable for a layperson to accept the views of the majority.

In the course of his discussion of expertise, Coady also considers whether there are moral experts and, if so, whether moral philosophers are such experts. Coady denies that there are any moral experts, and so, *a fortiori*, that moral philosophers are moral experts. He claims not just that as a matter of fact that there are no moral experts, but that there *cannot* be, "because morality is too vast and amorphous a subject for anyone to be *significantly* better informed than *most* people about it . . ." (p. 54, emphasis in original). He draws an analogy with science. He thinks it too is a "vast and amorphous subject," and so also one where there cannot be experts. Certainly, both science and morality are vast, at least in the sense that there are huge numbers of phenomena which fall within their purview. It is less clear that they are amorphous: many theories of morality, for instance, hold that there is a small number of overarching moral principles which apply to particular, apparently diverse, events. In any case, Coady's rejection of the possibility of moral expertise on the grounds of the vast and amorphous nature of morality surely proves too much. Many subjects, including history, geography, the law etc. are vast and amorphous. Are we to say that there are no experts in history, or geography or . . . ?

If we reject the claim that there cannot be moral experts because of the vast and amorphous nature of morality, then it is at least possible that there are moral experts. And, indeed, if we do reject that claim, it follows from Coady's own definition of expertise that there *are* moral experts. Recall that for Coady "being an expert is simply a matter of being well informed about a subject, that is, having a significantly greater store of accurate information about it than most people . . . . " Now consider a morally sensitive historian, who has read widely about a range of cultures and periods. Since she knows many more facts about human behaviour than most of us, and is capable of making generally accurate moral evaluations of that behaviour, she also knows many more moral truths (assuming that there are such things, as Coady and common opinion both hold) than most of us. So, on Coady's definition, she is a moral expert.

What this shows is not, I take it, that there actually are moral experts, but rather that if we do not think that examples such as the sensitive historian force us to accept that there are, we need to reject, or perhaps broaden, Coady's definition of expertise. In any case, Coady's definition seems to me unduly narrow. Sometimes, as Coady has it, experts are just people who know more about some subject than most others. But sometimes they are people who can do things (including finding things out) that most of us can't, and sometimes they are people who can help people do things they can't otherwise do (and sometimes they are people who have all these abilities). Moreover, to call someone an expert is typically to accord them the status of an authority in their area of expertise, that is, someone whose opinion about what to do or believe provides a pre-emptive reason for the layperson to do or believe as the expert says, that is, a reason which supplants whatever reasons the layperson may otherwise have acted on. "The expert told me to" is *ceteris paribus* a sufficient justification for acting in accordance with the expert's direction (including coming to hold a belief).

It is the idea of the expert as an authority in this sense which we tend to resist in the case of morality. Some people clearly are better morally informed and more sensitive than others. Nevertheless, if one person is facing a morally difficult choice and takes a particular course of action on the advice of a second, acknowledged on all hands as morally wiser than the first, it is the actor, not her adviser, whom we hold morally responsible. That is, the opinions of the morally wise do not function as pre-emptive reasons for the less enlightened. This seems to be partly constitutive of the concept of moral autonomy. If so, moral autonomy is strikingly at odds with rational autonomy since, as noted above, the rational person recognises, and defers to, the views of experts and ought to do so. Probing the source and justification for this difference in any serious way is obviously beyond the scope of this discussion, but two possibilities suggest themselves. One is that, given the centrality of morality to human life, in the course of maturation all ordinarily intelligent people become morally competent deliberators. The morally wise can assist us in our deliberations in morally difficult situations by pointing to moral reasons, and their force, which it would otherwise be more difficult for us to see, but it is *our* recognition of those reasons which explains and justifies our action. A second possibility is that, given the importance of moral conscientiousness, we in effect apply strict liability rules to moral behaviour, as a way of trying to ensure that people do give due consideration to morally significant choices.

Expert reports are, of course, strong bases for belief. After his discussion of expertise, Coady looks at a number of sources of information which have a less elevated reputation, in particular, rumors, conspiracy theories, and the blogosphere. He argues that their reputation is unduly negative, and, that like more respectable sources of information such as the senses and testimony, they may in fact be valuable, even if not infallible, sources for the critical inquirer. Coady thinks each of these sources of information can

be defended in similar terms (see, for example, p. 139). I am not convinced by that claim, nor by the defence he does provide of rumors and the blogosphere, though his views about conspiracy theories seem to me to be powerfully persuasive. Let me look at each of these sources in turn, beginning with rumor.

Rumor

Rumor is obviously an important source of information, for good or bad. It has, however, received little attention from philosophers, though more from social scientists. As a glance at the relevant social science literature shows, it is a slippery concept, with a range of definitions on offer. Coady takes rumor to possess two defining features: it is communication which has passed through many informants, and it "has not been endorsed by an institution with official status" (p. 97).

Despite its unsavoury reputation, Coady aims to "defend rumor against those who argue (or simply assume) that rumors are always, or typically false" (p. 87). Coady argues, by contrast, that "many rumors are credible . . . and that in general the fact that a proposition is rumored to be true is evidence in favour of it being true" (p. 87). I find it hard to believe that anyone could seriously maintain a claim so obviously contrary to common experience as that rumors are *always* false, and Coady does not provide a reference for anyone who does so. On the other hand, rumors (like their close relative, gossip) certainly are widely seen as epistemically and even morally suspect.

Before examining Coady's attempt to rehabilitate rumor's standing, let me point to what I take to be problems with his definition of rumor. First, on that definition, one of the essential features of rumor is that its content is information which is not officially sanctioned. What counts as official here? According to Coady ". . . to describe a communication as 'official' is to say that it is endorsed by an institution with significant power (especially the power to influence what is widely believed) at the time and place in question" (p. 97). The fundamental problem with this description is that it detaches the concept of official communication from a connection with the occupation of offices, particularly state offices, by those who issue such communication. That means that, on the one hand, information provided by, say, a state-run newspaper will not count as official if it is widely seen as mendacious (as such newspapers often were in Communist states) and so lacks the power to influence what is widely believed. On the other hand, it means that information provided by an institution which is broadly trusted (as some religious organisations have been in repressive states) will count as official, even where that institution is illegal and operates underground. This surely gets things the wrong way around: whatever the people of the USSR thought of *Pravda*, for example, and however unlikely to believe its reports, those reports surely counted as official.

Furthermore, Coady's definition entails that many things will be counted as rumors which are not usually seen as such. Think, for example, of the kinds of stories about family members which people often tell. These often satisfy both conditions of Coady's definition: they have not been officially endorsed (and given the kinds of events they cover, such as first meetings of spouses, often not capable of being so endorsed) and they have passed through a number of informants. It jars my ear, at least, to call such stories rumors.

One important difference between broadly circulated unofficial claims which are not seen as rumors and those which are seems to be that rumors remain unverified. Official confirmation (sometimes) provides verification, but there are many other ways in which claims can be verified, most obviously by the rumored event actually occurring.

In any case, lack of verification is often taken as one of the defining features of rumor. It is noteworthy that on any approach which does take lack of verification (official or otherwise) as a defining feature of rumors, a rumor is not identified by its content, but rather by its place in a process of communicative transmission. An utterance becomes a rumor at some stage in that process — on Coady's account, after it has passed through a number of informants — and may cease to be a rumor, either by being discredited and broadly disbelieved, or by being verified.

Whatever the problems with his definition of rumor, the plausibility of Coady's defence of the epistemic credentials of rumor does not rest on that definition. He points out that both those who recount a rumor (whom he calls rumor mongers) and those who hear it — who must of course in turn recount it to others if it is to survive — generally have some interest in its truth. Indeed, this seems just to be a specific case of a more general socio-linguistic phenomenon: speakers do not usually want to appear to be exceptionally credulous, badly informed or ill-willed; and usually we subject speakers' claims to some degree of filtering (however implicit) before adding them to our stock of beliefs. Assuming that each listener/speaker is less likely to decide to pass on a rumor which she has heard if she judges it to be implausible than if she judges it to be plausible, the epistemic community jointly acts as a filter for rumors, winnowing implausible ones, and spreading plausible ones. On the basis of these considerations, Coady (p. 103) maintains that "two closely related and widespread views about rumor . . . are unfounded" viz. "that there is a presumption against believing rumors" and that "belief in a particular rumor becomes less warranted the further it spreads." On the contrary, according to Coady, "in general the fact that a proposition is rumored to be true is evidence in favor of its being true . . ." and ". . . except in special circumstances, our warrant for believing a rumor will actually increase as the rumor spreads" (p. 103). The special circumstances which Coady has in mind are the operations of "selection pressures other than judgments of plausibility"(p. 103). A rumor may flourish "even though it is highly implausible, because it satisfies some deeply felt psychological or social need" (p. 103).

I am unconvinced by the claim that, in general, the further a rumor has spread the greater our warrant for believing it. Coady's argument to this effect seems to me to rest on an equivocation on the meaning of "plausibility." What might be called subjective plausibility is a hearer's judgment that a claim is believable; objective plausibility indicates that there is evidence available to a hearer which provides the hearer with good reason to believe a claim. Obviously, and notoriously, subjective and objective plausibility do not always track each other — many people (for example, climate sceptics) have beliefs which are not well supported by evidence available to them. Members of a group might find a false claim plausible in the light of other false beliefs they share. Moreover, contra the implication of Coady's claim, there is no necessary conflict between, on the one hand, plausibility, and on the other, selection pressures such as "deeply felt psychological and social needs." As the work of social psychologists on cognitive biases such as confirmation bias (interpreting new data in a way which confirms one's preconceptions) and the bandwagon effect (believing something because others in one's group do) has shown, those needs often influence judgments of plausibility. Hence, in a group which shares false beliefs which reflect deep-seated attitudes, rumors which are consistent with or reinforce those beliefs are likely to be found plausible and flourish, even if false (think for example of the prevalence of blood libel among anti-Semitic populations in Europe in the Middle Ages). That such rumors are widely disseminated in that group provides no warrant for believing them.

Notwithstanding, Coady's claims about rumors hold in certain settings: research from the 1940s on, investigating workplace rumors, for instance, has found both that

such rumors tend to be accurate, and that they become more accurate through their life (DiFonzo and Bordia, 2008). There are a number of distinctive features about rumors and rumor-mongering in such settings. Firstly, the topic of the rumor is (often) of material importance to the group within which it circulates. Hence, hearers have an incentive to ensure that they only accept true (not just plausible) rumors, while rumor-mongers have an incentive to ensure that the rumors they repeat are true since there is likely to be a cost to gaining a reputation for unreliability. Secondly, since there are likely to be a number of people in possession of information relevant to the truth of the rumor, a rumor is unlikely to survive unless it is (largely) true. Thirdly, in the kind of small and stable population typically found in workplaces, it is possible to know who is an (un)reliable source of information, either about matters relevant to the truth of a particular rumor, or more generally.

However, none of these things has to be true about rumors and rumor-mongering in general, and in many cases none of them will be true. I conclude that both the generalisations about rumors which Coady objects to (that there is a presumption against believing rumors and that belief in a particular rumor becomes less warranted the further it spreads) and their contraries, which he supports (that there is a presumption in favour of believing rumors, and that belief in a rumor becomes more warranted the further it has spread), are false, or at least misleading. We should assess the plausibility of rumors on a case-by-case basis, in the same way as we treat testimony in general, by considering the reliability of our source, and fit with other well-established beliefs.

Conspiracy Theories

Like rumors and rumor mongering, conspiracy theories, and theorising, have a bad name. Indeed, while it would be a rare person who could claim in good faith that he never listened to or spread rumors, the denial that one is a conspiracy theorist now seems to be obligatory for anyone who wants to be taken as minimally rational. Coady's discussion, drawing on his previous work (Coady, 2007), shows how peculiar that attitude is.

Again, as is the case with rumor, there is a variety of extant definitions of conspiracy, at least as that term figures in the phrase "conspiracy theory." These definitions all agree in understanding conspiracies as involving groups of people jointly and secretly planning to bring about some state of affairs. They may also see conspiracies as involving deception (in order to ensure secrecy, say) and being morally suspect in their goals or methods (which in turn helps to explain why conspiracies are kept secret).

The modern scorn for conspiracy theories seems to originate with Karl Popper's discussion in *The Open Society* of what he calls the conspiracy theory of society, which he characterizes as the belief that "whatever happens in society — especially happenings such as war, unemployment, poverty, shortages, which people as a rule dislike — is the result of direct design by some powerful individuals and groups" (1972, p. 123).

Anyone who did believe in the conspiracy theory of society in this sense clearly would be deeply irrational. Popper (p. 123) himself claimed that the conspiracy theory of society was "very widespread." That seems doubtful, to say the least, and Popper provides no evidence that it is true. In any case, when people accuse others of being conspiracy theorists, they do not usually seem to be accusing them of holding the conspiracy theory of society, at least in its unvarnished Popperian sense.

It would be absurd to believe that everything that happens in society (or even just the seriously bad things) is the product of conspiracy; it would be equally absurd to believe that nothing that happens is. Conspiracies patently exist, so on any plausible

definition of conspiracy, in itself it cannot be irrational to believe that there are conspiracies. Moreover, since conspiracies are by their nature secret, it hardly seems irrational to believe that some things that have happened are the result of undiscovered conspiracies.

What, then, is supposed to distinguish the irrational conspiracy theorist from the rational person who nevertheless accepts that at least sometimes events are the results of conspiracies? Coady considers various salient possibilities. Perhaps what identifies the conspiracy theorist is a mistaken belief about the frequency of conspiracy? While it is difficult (and given the secrecy which is an essential element of conspiracies, perhaps impossible), to know how often conspiracies occur, they clearly are a common enough social phenomenon. Much criminal activity involves conspiracies, and even in workplaces and family life they are hardly unknown. Perhaps, then, the conspiracy theorist is someone who has false beliefs, not about the frequency of conspiracies, but rather about their significance? But clearly there have been highly significant — indeed world historical — conspiracies, such as those which brought Lenin and Hitler to power, and, more recently, the conspiracy which led to the destruction of the World Trade Center in New York. Other possibilities that Coady considers are that conspiracy theorists are distinguished by their (irrational) belief(s) that conspiracies are frequently successful, or that Western governments, especially the United States, engage in conspiracies often, successfully or significantly. In fact, far from being irrational, those beliefs are true, as the evidence that Coady produces shows.

Of course, particular beliefs that some people hold about conspiracies are false and even irrational, and some people are prone to attribute events to conspiracies with little or no evidence, and even to take the presentation of countervailing evidence as demonstration of the existence of the conspiracy. But these are just particular instances of broad epistemic failings; they no more show that it is generally wrong to believe in conspiracies than children's belief in the Easter Bunny show that children shouldn't believe anything their parents tell them.

The Blogosphere

The twentieth century was the heyday of what Coady calls conventional media — large private or public organisations, staffed by professional journalists, seeking out and transmitting information, and providing commentary, about such matters as international affairs, domestic politics, high-profile court cases, sporting events and so on. Much of what most of us believed about such things ultimately came to us through the work of such media organisations, which became enormously profitable and influential by virtue of their control of the mechanisms for gathering and disseminating news. Recent developments in information technology, in particular the growth of the internet, have undercut that control by allowing many private individuals around the globe to transmit and receive information. Many people and groups now run internet sites — blogs — to provide access to information which in the past would have been (largely) confined to mainstream media outlets, and to allow and invite comment from the readership, which can then be published on the blog. The numbers and importance of blogs as sites of information and discussion have grown to the point where it is meaningful to speak of a distinct blogosphere.

Coady draws a distinction between the blogosphere and conventional media. Journalism in the blogosphere supposedly differs from that which occurs in conventional media in that it is carried out by non-professionals, who "are not part of any large, formally structured, institution" (p. 162). The principal form of research of these blogger journalists "consists in examination of documentary evidence, rather than interviews (and other forms of contact) with people in power" (p. 162).

Since the blogosphere has become such an important source of beliefs it is certainly worth examining its status as a source of belief, and comparing it to that of conventional media. Coady's discussion is still one of the few philosophical considerations of these matters. A problem inherent in any such discussion is the rapid pace of the developments in information technology as well as the use of such technology. Although Coady does acknowledge some interaction between the blogosphere and conventional media (p. 158), the distinction between them has become more blurred even in the relatively short time since the book was written, with popular blogs becoming commercially valuable items attracting the attention of media conglomerates, newspapers becoming increasingly interactive, and writers more and more moving between the blogosphere and older media outlets. Such changes are exemplified in the recent career of Glenn Greenwald, described as a "prominent blogger" on page 148 of the book. Greenwald has in fact spent the past few years working for the *Guardian* newspaper, leaving in October 2013 to help run a new independent, for-profit, on-line news site, bankrolled to the tune of $250,000,000 by an internet entrepreneur.

Coady aims to show that the blogosphere is a valuable source of information, indeed superior to the conventional media. He considers Alvin Goldman's (2008) claim that it is the conventional media which are superior, because they have filters in place to help guarantee the quality of published material, unlike the blogosphere. The filters Goldman has in mind include the use of fact checkers to vet reports for accuracy, the requirement that more than a single source is used for a story, and limitations on the use of anonymous sources. The extent to which the conventional media actually make use of such filters is, of course, an empirical matter, and Coady convincingly shows that in fact the record of the conventional media in utilizing such filters is spotty at best. Moreover, Coady points to common practices of conventional media which tend to undermine their validity. Foremost of these is the misguided emphasis on so-called balance. Balance, as it has come to be understood by the conventional media, involves giving "equal time to opposing positions," rather than making any judgment about the epistemic status of those positions (again, reporting of the climate change debate provides an apposite example), or taking an independent stance about the substance of those positions. In the case of political reporting, this has led to the media becoming virtually a mouthpiece for the political establishment, with the kinds of results seen in the retailing of reports of the existence of Saddam's weapons of mass destruction in the prelude to the second Gulf War.

By its nature, the blogosphere massively extends the offerings that are available in the marketplace of ideas, while simultaneously allowing much greater scrutiny and criticism of those ideas than was possible with the kind of one-way flow of information which characterized conventional media in the mechanical age. If one accepts a Millian approach to the epistemic benefits of free speech, it would seem to follow that the rise of the blogosphere is leading to more true beliefs being held, as Coady maintains. That said, there are features of the blogosphere as it actually exists, which are less epistemically desirable. The picture of the blogosphere which emerges from Coady's discussion is one which is largely inhabited by epistemically responsible agents, open to rational discussion and rebuttal. One does not have to travel far into the blogosphere to discover agents who are not so open, to put it mildly. As noted above, according to Coady, bloggers base their claims in examination of documentary evidence. While this may be true of Wikileaks and the like, it would require an implausibly liberal understanding of the idea of documentary evidence for it to hold true of the more polemical and demented sites, of which there are many. Further, the blogosphere has allowed for a much greater fragmentation of what might be called communities of

believers — groups who share fundamental commitments and beliefs — than was the case in the heyday of the conventional media. That fragmentation allows members of those communities to insulate themselves from challenges to their cherished beliefs. While the rise of the blogosphere may have made much more information available, in itself it does not necessarily improve people's capacity to engage critically with that information and draw well-founded conclusions from it. Again, I think Coady presents a somewhat over-idealised picture of the blogosphere to support his claims.

Despite the reservations I have expressed above about certain details of Coady's arguments, this is a good book. It deals with under-explored issues of real intellectual and practical importance in an imaginative and stimulating way. It is also unusually readable — pithy, accessible and often witty. Its virtues would, I think, make it an excellent text for upper-level courses in applied philosophy.

References

Coady, D. (2007). Are conspiracy theorists irrational? *Episteme: A Journal of Social Epistemology, 4,* 193–204.

DiFonzo, N., and Bordia, P. (2008). Rumor in organizational context. In D.A. Hantula (Ed.), *Advances in social and organizational psychology: A tribute to Ralph Rosnow* (pp. 249–274). Abingdon: Taylor and Francis.

Elga, A. (2010). How to disagree about how to disagree. In R. Feldman and T.A. Warfield (Eds.), *Disagreement* (pp. 175–186). Oxford: Oxford University Press.

Goldman, A.I. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological Research, 63,* 85–110.

Goldman, A.I. (2008). The social epistemology of blogging. In J. Van den Hoven and J. Weckert (Eds.), *Information technology and moral philosophy* (pp. 111–122). Cambridge: Cambridge University Press.

Kelly, T. (2010). Peer disagreement and higher-order evidence. In R. Feldman and T.A. Warfield (Eds.), *Disagreement* (pp. 11–74). Oxford: Oxford University Press.

Popper, K. (1972). *The open society and its enemies, volume 2: The high tide of prophecy: Hegel, Marx and the aftermath* (fourth edition). London: Routledge and Kegan Paul.

The Journal of Mind and Behavior
Winter and Spring 2014, Volume 35, Numbers 1 and 2
Pages 93–96
ISSN 0271–0137

Transference and Countertransference Today. Robert Oelsner (Editor). London and New York: Routledge, 2013, 361 pages, $54.95 paper.

Reviewed by William Fried, Private Practice, New York City

Of considerable salience among the contents of this book is the publication in English, for the first time, of Racker's "Observations on Countertransference as a Technical Instrument" [pp. 18–29].[1] It is the first and keynote paper. The 16 papers that follow it are to a greater or lesser degree, responses to and elaborations on Racker's ideas. Either because of difficulties with the translation or conceptual flaws, however, the paper falls short of other works by Racker, principally those collected in his book, *Transference and Countertransference,* that has been available in English since 1968.

Consider the statement, "If the analyst can use his negative transference reactions in favor of the treatment, he is usually able to overcome them" [Racker, p. 19]. Surely, the obverse is true: if the analyst can overcome his negative transference reactions, he can use them in favor of the treatment. If the translation is accurate, and Racker actually said the former, then his concept is flawed; if he meant the latter, then the translation is flawed.

Another illustration of a problem inherent in the translation is the following passage:

> Thus, the analyst's feelings of guilt — which on the one hand were inappropriate to the aggression committed, but on the other were appropriate to Bertha's depression — became a guide to the aggression of Bertha's moral superego against the ego where she had introjected the seductive and frustrating primary objects (the "primary persecutors"). [Racker, p. 22]

The passage is unnecessarily difficult to decipher because it agglomerates too many ideas into an inordinately long sentence. Greater clarity would be achieved by the use of shorter sentences.

The book begins with a forward by Claudio Laks Eizirik, followed immediately by Robert Oelsner's introduction. Together, the two might constitute an excellent review of the volume as a whole, with appreciative references to the contents of most or all of

---

Correspondence concerning this review should be addressed to William Fried, Ph.D., FIPA, 29 West 64th Street, New York, New York 10023. Email: billfried9@gmail.com

[1]Quotations from and references to the papers in this book will be bracketed to differentiate them from quotations and references from other sources. The latter will be enclosed in parentheses.

the papers included. It was therefore alarming to me that my own impressions of the works diverged from theirs in quite significant ways, as witness the examples above, for mere openers. The conclusion I drew from this discrepancy was that both Eizirik and Oelsner are fluent in Spanish, and therefore less limited in their comprehension of these materials than a reader restricted to English syntax and vocabulary.

One idea that is central to most of the papers in this book is that transference consists of unconscious fantasies about crucial objects (people), that the patient projects into the analyst; that is, he believes that the attributes of these objects belong to the analyst and behaves as though this were so. Another idea shared by most of the authors of these papers is that countertransference is the analyst's experience of and response to having these fantasies attributed to him. A third focus of many of the papers is the phenomenon of enactment, the tendency or inclination for the analyst to engage in overt behaviors that express the qualities of the patient's objects. When the patient exhibits behavior under the influence of the object or self representations to which he attributes unconscious fantasies, this too is defined as an enactment, the term having replaced "acting out," the word that was used for this purpose in the past.

Some of the papers raise and discuss the important issue of the analyst's transference and/or countertransference to psychoanalysis itself; to the specific version(s) of it she has been taught and acquired in the course of her education, training, personal analysis and experience; and to particular technical procedures or interventions. Lemma, for example, points out that there are institutes and supervisors that promote the value of the transference interpretation as though it, alone, is the element that distinguishes psychoanalysis from psychotherapy and is therefore the definitive intervention in psychoanalysis, beside which all others are regarded as subordinate [Lemma, pp. 129–130].

It is noteworthy that, though some of the authors hint at the influence on the analyst's approach of early childhood and pre-training experiences, none mentions the impact of the implicit theory of personality that the analyst develops from earliest infancy, the strength of which is often undetected throughout training, an oversight that may result in its later emergence in the form of obstructive countertransference reactions.

Another relevant phenomenon that elicits no criticism in these pages is the analyst's naive inclination to use questions as interventions, as illustrated in several of the case descriptions. It is as though the necessary alertness to unconscious meaning and defensive operations is turned off when the analyst finds herself asking a question in response to some unclarity or obscurity in the analysand's discourse. That is, rather than regard it as a possible resistance, an unconsciously determined omission or obfuscation, the analyst will ask for clarification by means of a straightforward question. This is an excellent example of the most common form of enactment; it is a verbal one that may occur several times in the course of a single session, and become integral to a defensive fabric across many sessions, with no awareness of its existence or function either by the analyst or, for that matter, her supervisor. If analysts are encouraged, as they are by all of the authors in the current volume, to examine their own in-session behavior for evidence of countertransference, how is it that the interrogative inclination eludes this process?

A pervasive difficulty with the thematic trope of this book is its language. To begin with, much confusion and needless exposition might have been averted had Racker and his successors not chosen to name the process by which the analyst becomes aware of the effects of being treated as the object(s) of the patient's unconscious fantasies, *countertransference*. The term had already been coined for a similar but crucially divergent

entity, the analyst's residue of unanalyzed, poorly analyzed, or incompletely analyzed unconscious conflict, defense, fantasy, affects, etc. that interfered with her ability to analyze. This residue would affect her work with each analysand in a different way, depending on the degree to which it might match aspects of the analysand's personality structure. The processes are sufficiently distinct to have required different names. The separation thus achieved would have been salutary.

The problem to which I have just referred is one side of something even more troubling. The use of an already saturated term to name a new discovery is paralleled by the invention of neologisms to name processes for which adequate nomenclature already exists. The authors of this book do this in abundance, contributing to further obscurity and difficulties of understanding. Here, for example, is a short list of such ad hoc words or phrases:

> "The countertransference position" [Faimberg, p. 51]; "The analyst's negativization of himself," "calculated vacillation of neutrality," "symbolizing historicization," "rechannel the cure" [Fainstein, pp. 77, 78];"misconceptions" [Weiss, p. 122]; "the mindful object," "the objectivization of subjectivity" [Lemma, pp. 127, 145]; "tragic ethics,""spectral non-structural models, non linear and non-hermeneutic,""ethical aesthetics" [Chuster, pp. 317, 220, 225]; "mishmashed," "mestizo," "terrorism of suffering" [Borgogno and Vigna–Taglianti, pp. 294, 296].

The specialized use to which the authors put many of these terms renders them less comprehensible to readers familiar with their generally accepted connotations. This tends to be the case even in the rare instances where the author attempts to define the novel sense and context for which the term has been recruited. Even if I were ignorant of more recent psychoanalytic coinage, I would remain concerned that a significant number of my colleagues might approach the task under the same handicap and I am reasonably certain that the editor and publishers would not wish deliberately to limit the size of the book's prospective readership.

My principal criticism of this volume as-a-whole is that each author tends to take the concepts of transference and countertransference in the direction that the author finds congenial to his or her own practice, experience and understanding of these terms. Each selects and quotes ideas from the works of certain earlier or contemporary writers, and ignores those of others that may have equal relevance. This selection is rendered possible by the absence of criteria for the validity of any particular set of ideas or its superiority over others for explanatory or therapeutic purposes. Indeed, the extant outcome research tends to confirm that no one psychodynamic approach is superior to others (Shedler, 2010).

In psychoanalysis, the concepts of theorists whose prominence tends to be a complex function of their writing, teaching, theorizing, and charisma, are likely to gain currency. But once they enter the circulatory system of the profession, they undergo changes and transformations that are significant. This volume illustrates the process well. Examples are Haydee Faimberg's [pp. 53–54] redefinition of Racker's concept of complementary identification; Racker's (1968) and Paula Heimann's (1950) assertion, contrary to Freud, of ". . . countertransference as not being synonymous with the analyst's neurosis on the basis that countertransference is triggered by the patient" [Faimberg, p. 55]; and Fainstein's [pp. 68–87] focus on the reformulation of the clinical implementation of countertransference based on Lacan's contributions.

Having registered my reservations about *Transference and Countertransference Today* in the foregoing paragraphs, it remains for me to commend the book for the ability of

most of its authors to provoke thinking about the uses of analysts' inner experiences as these experiences evolve in the clinical metier. These texts, despite the stylistic and semantic obstacles to which I earlier referred, are intellectually and psychologically challenging and will reward the reader willing to struggle with the prose. The discussions that stimulated me most were in Lemma's [pp. 127–149] paper on the uses and misuses of transference interpretations; Berman–Oelsner's [pp.177–195] contemporary Kleinian view of psychoanalysis with children; Faimberg's [pp. 49–67)] sensitive, though often occluded clinical observations; and Robert Oelsner's [pp. 236–255] more lucidly explicated case material. The book also renders an inadvertent service to the cause of improved psychoanalytic writing and dialogue by demonstrating the need for greater attention to a shared and versatile language in which to give voice to the slowly developing convergence of ideas from heretofore disparate analytic perspectives.

References

Heimann, P. (1950). On counter-transference. *International Journal of Psychoanalysis, 31,* 81–84.

Racker, H. (1968). *Transference and countertransference. The international psycho-analytical library* (Volume 73, pp. 1–196). London: The Hogarth Press and the Institute of Psychoanalysis.

Shedler, J. (2010). The efficacy of psychodynamic psychotherapy. *American Psychologist, 65,* 98–109.

The Journal of Mind and Behavior
Winter and Spring 2014, Volume 35, Numbers 1 and 2
Pages 97–106
ISSN 0271–0137

Schéma Corporel, Image du Corps, Image Spéculaire. Neurologie et Psychanalyse [Body Schema, Body Image, Specular Image. Neurology and Psychoanalysis]. Catherine Morin. Toulouse: Éditions érès, 2013, 214 pages, 13 euros.

Reviewed by Dorothée Legrand, CNRS, Ecole Normale Supérieure

---

1. Position of the Reader

In her book, *Schéma Corporel, Image du Corps, Image Spéculaire. Neurologie et Psychanalyse* [*Body Schema, Body Image, Specular Image. Neurology and Psychoanalysis*], Catherine Morin aims at understanding the "subjective consequences of strokes"[1] [« conséquences subjectives des accidents vasculaires cérébraux »] (p. 11) by relying on patients' reports, and by interpreting them from a perspective at the interface of neurology and psychoanalysis. Throughout the book, Morin gives a brief description of different concepts she relies on, concepts about which there is no consensus, neither in neurology nor in psychoanalysis, nor, even less, between these two disciplines; she quickly criticizes different positions, alternative to her own, positions from cognitive sciences, psychology, or neuropsychoanalysis, the latter discipline being younger than the other two but no less prolific on the topics at stake. Her rapid treatment of these topics appears as a way to avoid getting stuck in the maze of historical and/or contemporary debates on what is an object, what is a subject, what is a delusion, and, a question that is not the least weighted, what is a body, a body image, a body schema. But is this rapidity superficiality or efficiency? Both maybe, but here we will leave this question unanswered, to follow the path pursued by the author herself. Thus, we won't point to other definitions of the aforementioned notions, other definitions to which an objector may still object, and so on. A more interesting question to start from is one that Morin herself raises: given this theoretical and clinical setting, "What have we learned? That is to say: What did the patients teach us?" [« qu'avons-nous appris? C'est-à-dire: que nous ont appris les patients? »] (p. 189), and, in particular, what have we learned about the subjective consequences of brain injuries?

To enter into this question, we cannot but consider the way the author places herself in the position to learn from patients; this involves considering the way patients place themselves or are placed in a position to relate the subjective effects of their stroke

---

Correspondance concerning this review should be adressed to Dorothée Legrand, Ecole Normale Supérieure, Archives Husserl, 45 rue l'Ulm, 75005, Paris, France. Email: dorothee.legrand@ens.fr

[1]All translations are mine. The original French text is indicated in brackets after each quote.

in order to teach interlocutors who hold two perspectives at once: on the one hand, the patient's interlocutors aim at learning something about strokes, and on the other hand, they are "concerned to help him [the patient], involved in his rehabilitation" [« soucieux de l'aider, impliqués dans sa réadaptation »] (p. 71).

2. Position of the Author

From the outset, Morin defines her own position as "that of a physician–researcher, a doctor who seeks to understand what the pathology is from the point of view of the patient. I could specify: from the point of view of *a patient* when the interview was included in an investigation, from the point of view of *her patient* when the interview participated in the dialogue between a patient and his doctor" [« celle d'un médecin-chercheur, un médecin qui cherche à comprendre ce qu'est la maladie du point de vue du patient. Je pourrais préciser: du point de vue *d'un patient* lorsque l'entretien était inclus dans une recherche, du point de vue de *son patient* lorsque l'entretien participait au dialogue entre un malade et son médecin »] (p. 12). Physician–researcher? We should not underestimate the importance of the en-dash between the position taken by a physician oriented by psychoanalysis and that of a researcher who pursues an epistemic aim. Indeed, by this en-dash, Morin assumes, and enacts in her practice as in this book, that the same person can be both at once such a physician and such a researcher vis-à-vis another person, a patient. But we cannot ignore or overlook that the aims of these two practices are opposite to each other. While the clinician oriented by psychoanalysis, as Morin is, aims to listen to the patients singularly, without filtering in any way what reaches her ears, the researcher, on the contrary, aims at an epistemic benefit which, if only implicitly, immediately orients her listening on how the patient's speech can be integrated or not to her hypothesis, as indeterminate as it may be. What is at stake here is the whole question of the "neutrality" of the listener oriented by psychoanalysis. This neutrality guarantees that an *unconditional* hospitality can be given to the words given by the patient singularly; yet this neutrality is undermined by the epistemic objectives pursued by the clinician if she adopts *both at once* an analytical and an epistemic position. As a physician–researcher, Morin assumes that she does not bracket her position as a physician–analyst, when she undergoes her investigations, nor does she bracket her position as a researcher pursuing an epistemic aim, when she practices as a clinician. On the one hand, this meshing holds the promise of an epistemically rich medicine and a clinically rich research. On the other hand, this epistemologico–clinical meshing also contains the risk to elicit, orient, or enclose the patient's speech into the orthopedics, the normativity of a framework motivated by epistemic benefit, knowledge, and learning.

The point here is *not* to suggest that research and clinical work should operate separately; on the contrary, since any research that takes the patient's speech as "empirical data" has necessarily an effect on the subjective position of the patient relative to what he tells or does not tell to the researcher; in other words, any investigation involving patients should be conceived of in a clinical setting. But if one must assume the clinical significance of any research based on the patient's speech, it should also be noted that the epistemic objectives which animate an investigation are fundamentally incompatible with what animates the clinical encounter between a practitioner oriented by psychoanalysis and an incessantly singular patient. Because it is informed by her clinical practice, the investigation performed by Morin is rich, relevant, and operative, but the ethics of a clinical practice oriented by psychoanalysis would want that the patient's speech is received, given hospitality, listened to, regardless of, for example, the repre-

sentativeness of this particular patient compared with the group of brain-injured patients to whom he is assimilated in an epistemic perspective.

In the case of Catherine Morin, it is not uninteresting to note that it is "the patient's or caregivers' words [. . .] which made [her] quit the physiology of motor disability and orient [her] investigations towards the subjective consequences of strokes" [« des propos de patients ou de soignants [. . .] qui [l]'ont fait quitter la physiologie du handicap moteur et orienter [ses] recherches vers les conséquences subjectives des accidents vasculaires cérébraux »] (p. 11). It is suggested that the "therapeutic postulates" [« postulats thérapeutiques »] (p. 131) in a department of "neurological rehabilitation" [« rééducation neurologique »] (p. 11) would put some medical objectivity in tension with the subjectivity of brain-injured patients. However, this very tension would also be present between the psychoanalytical approach, on the one hand, and the epistemic aims of an investigation, on the other hand, at least in the sense that the latter aims at integrating "the patient's state" [« l'état du patient »] to what is known or knowable. The author is confronted with an orthopedic aim, therefore, not only in rehabilitation, but also in the epistemic approach.

This tension, between the position of "physician–researcher" and the subjectivity of the patient, can notably be found when Morin describes (quickly) the methodology of her qualitative research. She notably explains that the patients passed "semi-structured interviews conducted at [her] request, that is to say, and this was always explicitly stated, at the request of a doctor who seeks to understand what the pathology is from the standpoint of the patient" [« entretiens semi-directifs réalisés à [sa] demande, c'est-à-dire, et ceci était toujours formulé explicitement, à la demande d'un médecin qui cherche à comprendre ce qu'est la maladie du point de vue du patient »] (p. 70). Don't we hear here a tension between the demand emanating from the doctor who seeks to understand the pathology, and the point of view of the patient addressing himself to a clinician? The question of the impact of this demand on the patient and his words, the question of the impact of the clinical relationship on the words collected with an epistemic aim, the question of the impact of the epistemic context on the clinical encounter, none of these questions is asked; all should be. It is the *address* of the words, drawings, gestures of the patient that is at stake here, an address to another person that is essential to consider if the clinician oriented by psychoanalysis wants to put into practice the Lacanian idea according to which the clinician's attention is to be focused on what the patient says insofar as he says it *to the listener*, i.e., focused on the patient's act of saying, insofar as it is addressed to the clinician.

3. Position of the Patient

The position Catherine Morin gives to herself relative to the patients never ceases to intrigue the reader. Let us go back to what she says herself about it, to emphasize its correlate, i.e., the position she thereby gives to the patients. Morin, as we have seen, defines her own position as "that of a physician–researcher, a doctor who seeks to understand what the pathology is from the point of view of the patient. I could specify: from the point of view of *a patient* when the interview was included in an investigation, from the point of view of *her patient* when the interview participated in the dialogue between a patient and his doctor" [« celle d'un médecin–chercheur, un médecin qui cherche à comprendre ce qu'est la maladie du point de vue du patient. Je pourrais préciser : du point de vue *d'un patient* lorsque l'entretien était inclus dans une recherche, du point de vue de *son patient* lorsque l'entretien participait au dialogue entre un malade et son médecin »] (p. 12). "*A patient*," emphasized by the author by

opposition or in complementarity with "*her patient*," is heard here as "a patient among others," a patient "protected" by an anonymity which also deprives him of his subjective singularity, a patient who is integrated into the group of "the patients" whose brain is injured on the right or the left, a patient whose speech is analyzed systematically in order to be subjected to statistical tests, a patient, therefore, who is integrated within the epistemic framework of this physician–researcher. Moreover, this same person is also "*her patient*," emphasized by the author by opposition or in complementarity with "*a patient*." This possessive pronoun, and the fact that it is emphasized by the author, does not seem to indicate that the doctor takes the patient as her territory for exploration; rather, here, the doctor seems to consider as essential the fact that the patient speaks to her, addresses his speech to her ears: it is *her* patient and not the patient of any doctor, because the act of listening of this doctor is not substitutable to that of any doctor. In other words, for this patient, this doctor is *his* doctor.

Here, we see how a practice that would assume a hierarchical relationship between doctor and patient does not necessarily suffer from all the pitfalls which it is accused of, and, in the first place, it does not exclude but may rather allow respecting the patient's speech. This is assumed as such by Morin for whom it is "essential, not only to interrogate patients in a non-suggestive way and to leave room for their spontaneous discourse before questioning them about their deficits, but also to avoid systematically proposing interpretations drawn from normal psychology before characterizing the patients' discourse" [« essentiel, non seulement d'interroger les patients de façon non suggestive et de laisser la place à leur discours spontané avant de les questionner sur leurs déficits, mais aussi de ne pas proposer systématiquement des interprétations tirées de la psychologie normale avant d'avoir caractérisé le discours des patients »] (p. 175). It should be noted, however, that the "spontaneous discourse" of the patient, discourse whose spontaneity would be preserved thanks to the discretion of the clinician who puts her own questions aside, is actually and can only be addressed to an other. Therefore, the clinician's caution vis-à-vis any suggestion, and even her silence, does not imply that the patient delivers a speech that would be free of any influence of the clinician, this "influence" being the very structure of speech as it is addressed to an other.

4. Psychoanalysis

All of Morin's enterprise is motivated by psychoanalysis: the point for her is indeed to "describe [the] neurological disorders of self-representation in psychoanalytic terms" [« décrire [les] troubles neurologiques de la représentation de soi en termes psychanalytiques »] (p. 14). Among these psychoanalytic terms: object. Needless to say, the term "object" is not specifically analytical; moreover, it is not defined unequivocally within psychoanalysis itself, it is even a topic of division of this field into different fratricidal trends. Thus, we can only be surprised, and in fact hindered by the lack of definition of this term, even though Morin places it at the center of her conceptualization of the "right hemisphere syndrome" [« syndrome hémisphérique droit »] — I shall get back to this below.

Morin also seems to casually assume a point that generates significant tensions within psychoanalysis, and between the latter and some of its critics: symbolism, in which a patient's body parts, words, gestures, or drawings is taken as a metaphor of some general meaning. For example, the "symbolic, specificity, of the left side as the 'bad' side" [« spécificité symbolique du côté gauche comme 'mauvais' côté »] (p. 44); the hand as part of the "phallic signifiers" [« des signifiants phalliques »] (p. 64); the

eye and mouth as "displaced representations of the female sex" [« représentations déplacées du sexe féminin »] (p. 66), the lack of figuration of the mouth "as due to the sudden, traumatic introduction of impairments and disabilities insulating the subject from the social bond" [« comme liée à l'instauration brutale, traumatisante, de déficiences et d'incapacités isolant le sujet du lien social »] (p. 79), etc. This practice is striking with one of the patients who Morin presents in greater detail, Mr. E., categorized as displaying a right hemisphere syndrome, and who loves fishing. Morin interprets as follows: "torrent fishing is quite specifically a masculine activity, and identifying the fishing rod as a phallic representation is hardly risky" [« La pêche en torrent est une activité assez spécifiquement masculine, et repérer dans la canne à pêche une représentation phallique n'est guère risqué »] (p. 116). In contrast to Morin, I find it "risky" to tack a phallic representation onto any object that would be a bit long — a toothbrush, a spaghetti? Not only does this involve forgetting that the phallus is the signifier of lack, but also this runs the risk to use psychoanalysis as a sort of key of dreams that would tack significations onto the patient's manifestations, thereby veiling the singularity of his physical, mental, emotional, cognitive states. As Morin herself emphasizes, "more interesting is to relate the space between this instrument and the body and words of Mr. E." [« plus intéressant est de mettre en rapport l'espace entre cet instrument et le corps et les mots de monsieur E. »] (p. 116). Still, about Mr. E., we learn from Morin that "mouth and beak can be considered as sexual symbols" [« bouche et bec peuvent être considérés comme des symboles sexuels »] (p. 119) and that, therefore, by applying these general symbols to Mr. E. in particular, we could interpret their absence in his drawings as symbolizing "the 'erasure' of sexual concerns which this patient reports" [« l'"effacement' des préoccupations sexuelles dont fait état ce patient »] (p. 119). But which place does this "erasure of sexual concerns" take in the life of Mr. E. in particular? This is what we cannot respond to, on the basis of the absence of figuration of mouth and beak in the drawings of Mr. E., if we read such absence only through general symbolism. As Morin underscores herself, it seems more relevant to note that the raptor which Mr. E. draws is not only without any beak; it also presents spurs [*ergots*] which the patient explicitly associates to the sessions of *ergo*therapy he goes through since his brain injury. What is said here — through this raptor — about the position Mr. E. takes relative to the process of rehabilitation that the pathology imposes to him? Again, the use of symbolism does not seem to answer this question which is crucial clinically.

These criticisms being placed, let us suspend them here, to rather reveal the specifically psychoanalytic dimension of Morin's approach — a psychoanalytic dimension which is thus not tied to the "psychoanalytic terms" since these are not properly defined, nor to the use of symbolism, since the latter is not strictly psychoanalytic. Morin positions psychoanalysis in a place which is particularly favorable for its practice, i.e., between neurology and psychology, and outside the field of cognitivism. Morin properly stresses the difficulty there is to hold this position in a department of rehabilitation, at the hospital, "where two dangers threaten the therapist: 'psychologizing' everything as if the brain injury had no psychic [physical?] organic effects, 'cerebralizing' everything as if the patient were not entitled to or escaped common suffering" [« où deux écueils menacent le thérapeute : tout 'psychologiser' comme si la lésion cérébrale n'avait pas d'effets psychiques [physiques ?] organiques, tout 'cérébraliser' comme si le patient n'avait pas droit ou échappait à la souffrance commune »] (p. 17); "these two pitfalls are two faces of the same coin. They threaten us as soon as we seek to 'know how patients function' " [« ces deux écueils sont l'avers et le revers d'une même médaille. Ils nous menacent dès que nous cherchons à 'savoir comment fonctionnent les patients' »] (p. 190).

The difficulty of keeping this work "between neuroscience and psychoanalysis" [« entre neurologie et psychanalyse »] (p. 12) and of keeping psychoanalysis between neurology and psychology, is indubitable. But, as we said, Morin, in this book, seems to position psychoanalysis in a place which is particularly favorable for its practice; how so? Psychoanalysis — and this is what signed its birth certificate — never ceased to characterize physical symptoms linked to neurosis, notably by relating symptoms to functional disorders, and distinguishing them from organic lesions. For example, psychoanalysis meant to distinguish hysterical conversions from epilepsy. The analytically oriented psychosomatic approach blurred this distinction, by its attempt at determining which psychic structure would account for the emergence of eczema, ulcers, asthma, or other events described as psychosomatic: somatic disorders of psychic origin. But in cases of brain injury, there is no doubt about the organic etiology, which immediately prevents any psychologizing or psychosomatizing temptation which would interpret as psychic what is physiological. In this context, the field is left open for another question: Which sense or which role does the patient give to his troubles? Morin's question is not: What is the psychic structure which may explain mental disorders (of self-representation) which these patients suffer from? Her question is rather: Given their psychic structure, their past, their projects, etc., in which way do the patients live their injuries? And, as Morin points out, "it is only by listening to the patient talk about what happens to him that we can appreciate the particular position he holds as a subject relative to his pathology. But it is also only by listening to him that we will characterize his pathology itself" [« c'est seulement en écoutant le patient parler de ce qui lui arrive qu'on peut apprécier sa position particulière de sujet par rapport à sa pathologie. Mais c'est aussi seulement en l'écoutant qu'on va caractériser sa pathologie même »] (p. 190). Thus, in this approach, we do not only learn the patient's subjective position, which is essential for any clinical encounter, we also learn about the pathology itself, which is thereby characterized as a subjective disturbance — whose etiology is unambiguously cerebral. We are thus invited to a practice that inverts the psychosomatic approach: while psychosomatic means to account for the organic etiology of mental disorders, Morin's approach, "between" neurology and psychoanalysis, is interested in "psychic symptoms of neurological origin" [« symptômes psychiques d'origine neurologique »] (p. 44).

This work is thus particularly favorable to the practice of psychoanalysis, and for yet another reason. As the body is irreducible to the representation, knowledge, and mastery one has of it (pp. 41, 44, 47), disorders of body image and body schema are themselves irreducible to these cognitive categories: indescribable, incomprehensible, and inexplicable in these terms (p. 169). Thus, it is a non-cognitive clinical practice and theoretical conceptualization — here psychoanalysis — that is the most legitimate to account for such disorders. The subject is captured by the shape of his body, an object of the other's desire over which he can have neither knowledge nor mastery (pp. 47–48), and the clinician must be able to avoid capturing this body and its disorders into a mastery and knowledge of which she would hold the secret — diagnostic categories and brain mapping, for example. The singular reactions of each patient cannot be reduced "to stereotypes independent from his psychological structure shaped by his personal history" [« à des stéréotypes indépendants de sa structure psychologique façonnée par son histoire personnelle »] (Afterword, p. 196). What is at stake is how an injury and the disorders that it triggers will be inscribed into the "continuity of the psychic life of the patients" [« la continuité de la vie psychique des patients »] (Afterword, p. 194), inscription which the patient suggests when he speaks to *his* doctor, if the latter lends herself to listening to him singularly: "the existence of body image disorders cannot erase the psychic structuring of the subjects who are affected by them. In front

of each patient taken individually, whether or not he has body schema disorders, it is only by listening to what he says about himself and his body that we may adjust the dialogue with him" [« l'existence de trouble de l'image du corps ne saurait effacer la structuration psychique des sujets qui en sont frappés. Devant chaque patient pris individuellement, qu'il ait ou non des troubles du schéma corporel, c'est seulement en écoutant ce qu'il dit de lui-même et de son corps qu'on pourra ajuster le dialogue avec lui »] (p. 73).

Morin thus develops her work in a place which is both privileged and particularly difficult for the practice of psychoanalysis, out of the cognitive field, "between" cerebralizing and psychologizing. Now, we are dislodged from this place when Morin conducts a "multivariate descriptive analysis" [« analyse descriptive multivariée »] which aims at "identifying similarities and differences in a set of objects," [« déceler les ressemblances et les dissemblances dans un ensemble d'objets »] in this case, in a group of brain-injured subjects (p. 71). It is on the basis of such quantitative analysis that Morin can affirm that "self-portraits of patients with body schema disorders clearly stand out from those of patients without body schema disorders" [« les autoportraits des patients présentant des troubles du schéma corporel se démarquent clairement de ceux de patients sans troubles du schéma corporel »] (p. 73). Here, there is no place for *the patient* facing her singularly, for "a patient" [« *un patient* »] who she encounters as a researcher, for "*her patient*" [« *son patient* »] who she meets as a doctor: only "the patients" [« les patients »] remain. While she explains how "the patients" use personal pronouns according to whether they belong to the groups of right or left brain lesions, while she points out how "the patients" blend themselves in a group and blur their singularity (pp. 93–94) by using the "generic you" [« vous générique »] or the "collective we" [« nous collectif »] (p. 89), Morin herself uses the generic pronoun "they," thereby undifferentiating the singular subjects who addressed themselves to her (p. 86).

In doing so, Morin does not only stray away from the act of analytic listening whose singularity excludes comparativity; she also departs from the analytical conception of the speaking subject and his speech. In a psychoanalytic framework, indeed, we cannot stick to the patient's speech as if a subject who says "I" necessarily positioned himself subjectively and a subject who says "they" necessarily faded away subjectively; the reverse may as well be the case. The "subject of enunciation" [« sujet de l'énonciation »] (p. 92) does not manifest himself in his speech in the number of occurrences of the word "I." We should not overlook the fact that the subject of the act of speech may be absent from the "I" he says, or manifests himself in saying "we," just like we should not neglect the impossible coincidence, the systematic difference between the "I" said and the one who says it. Such negligence would involve taking what is said literally, rather than as a rebus where the subject reveals himself while veiling himself; it would suspend the hypothesis of the unconscious for the sake of quantitative analysis.

5. Who is my Hand?

We can now return to the first question that animates this entire book: from this particular place taken by psychoanalysis, a place which invites us to "navigate in a minefield of confusions and mis-sense" [« à parcourir un champ miné de confusions et de faux-sens »] (p. 91), "what have we learned? That is to say: what did the patients teach us?" [« qu'avons-nous appris ? C'est-à-dire : que nous ont appris les patients ? »] (p. 189), and in particular what have we learned about the "subjective consequences of strokes" [« conséquences subjectives des accidents vasculaires cérébraux »] (p. 11)?

Relevantly, Morin distinguishes "brain-injured patients without right hemisphere syndrome who, in cases of sensory disorders, report: 'It is *as if* my hand were not mine'" [« les patients cérébrolésés sans syndrome hémisphérique droit qui, en cas de troubles sensitifs, disent: « 'C'est *comme si* ma main ne m'appartenait pas' »]; "Right Hemisphere Syndrome patients who *claim* that their hand is that of someone else" [« patients SHD qui *affirment* que leur main est celle de quelqu'un d'autre »]; and "psychotic patients who are *convinced* that the control of their body is removed by an unpleasant being" [« patients psychotiques qui ont la *conviction* que la maîtrise de leur corps leur est retirée par un être antipathique »] (p. 159).

Here we see how grouping Right Hemisphere Syndrome patients with psychotic patients would neglect the specificity of psychotic delirium that breaks into the mental states and concrete life of the patient, and whose strength of conviction cannot be doubted by the patient, even when the latter shows awareness of his delusion as such. As opposed to the persecution experienced by the delirious patient, Right Hemisphere Syndrome patients rather seem to find a form of "consolation" with their hand, which presents to them their object of choice: as a mother who regards her hand as if it were the daughter she never had (pp. 146, 184).

In addition, we also see, throughout Morin's work, how grouping Right Hemisphere Syndrome with asomatognosia would neglect the specificity of the way in which Right Hemisphere Syndrome patients live their body. What troubles the Right Hemisphere Syndrome patient, and Morin makes it clear, is not a deficit of knowledge (*a-gnosis*) of his own body (*soma*). This characterization in terms of deficit suffers from two errors: a conception of normality as involving some knowledge of one's own body, somatognosia; a conception of pathology as a deficit vis-à-vis what characterizes normality, asomatognosia. If one questions these conceptions, Right Hemisphere Syndrome becomes more readable. So let us return to these two presuppositions which, although problematic, are nonetheless active throughout the medical approach.

Everything happens as if the researchers and doctors, who are interested in the lived body and its perturbations, predominantly think of body image as a more or less faithful reproduction of the body as it is objectively describable, i.e., as an object whose shape, location, weight, etc. can be measured by an impartial observer. The fidelity of the (mental) representation vis-à-vis the represented (a sort of mental equivalent of pictorial mimesis) would fall within normality; infidelity within pathology. But what is a body image?

One's body image — at least in the field of investigation that drives Morin's work — is not founded on, nor founds a form of knowledge of one's own body; the body image, mental or reflected by the mirror, is a construction that supports the subject's "misrecognition" [« méconnaissance »] (pp. 46–47) of his body. Misrecognition is here two-fold. First, there is a misrecognition of the "real neurological immaturity" [« l'immaturité neurologique réelle »] (p. 47) and of the sensory-motor incoordination with which it is correlated, i.e., a misrecognition of the "real" body that is disunited and uncontrolled, a misrecognition of the bodily fragmentation thanks to the veil of a unifying image of the body as "one." Second, there is a misrecognition of what the lived body owes to the relation of the subject to an other; an other who points at that body as his object of desire. There is thus a misrecognition of the fact that one's body image is incomplete, in the sense that what gives it its form is precisely unimaginable, unrepresentable. Indeed, it is the unimaginable desire of the other which gives its form to one's body image, as the other gazes at the body of the child facing the mirror and designates it: you are this image for me.

Your body image is thus a "knot" [« nouage »] (p. 46) between (1) a real body which is unknown and even unknowable, (2) another subject whose desire does not have

any image but who gives a name to the image of your body as what you are, and (3) a visible form of the body recognized as yours, a shape that wraps around the desire of the other as a lack which it veils without filling it. This knot, Morin says, is "normally unapparent" [« normalement inapparent »] (pp. 46, 53) and it is in this sense that the identification with the mirror image is not a recognition of a certain reality of one's own body in the mirror: what is at stake is neither cognition nor re-cognition, but identification — you are that.

What happens then when the knot unravels? The image is torn apart. It no longer gives the orthopedic shape of the entire body. Instead, its tear reveals the elements that took shape from the fact of being tied together: the real body and the desire of the other that I cannot see. It is the unimaginable — that which can have no image — that tearing the image reveals.

5.1. Patients with Disturbances of the Body Schema

The real body is what would be imposed to the patient, in the case of a left hemispheric lesion, without body schema disorders. The body whose image was pointed at by the desire of the other, this body is no more. Following the stroke, the sudden tearing of the desired, ideal, unifying image may reveal the body as raw material, inert, uncontrollable. But this is unimaginable. This body must be covered with white, since it is no longer covered with its own ideal image. It is thus that "the words and self-portraits of patients without disorders of the body schema have in common a certain silence on the paralyzed body" [« les paroles et les autoportraits des patients sans troubles du schéma corporel ont en commun un certain silence sur le corps paralysé »]; these patients "notice their paralysis only when they want to make a move and fail" [« ne constatent leur paralysie que lorsqu'ils veulent faire un mouvement et y échouent »] (p. 97), "despite a visible and asymmetric disability, [they] maintain a stable, erected, symmetrical image of their body, and react to the loss they have just been subjected to, following a classical process of mourning: it is little by little that they will unveil the normal neurotic misrecognition of the body and its pathological alterations, and that they will recognize the actual loss they have suffered" [« malgré un handicap visible et asymétrique, [ils] maintiennent une image érigée, stable et symétrique de leur corps, et réagissent à la perte qu'ils viennent de subir selon un processus de deuil bien classique: c'est petit à petit qu'ils vont lever la méconnaissance névrotique normale du corps et de ses altérations pathologiques et qu'ils vont reconnaitre la perte réelle qu'ils ont subie »] (pp. 111, 176). The mourning of the functionality of my body and of the ideal image I had of it is a gradual process that shows the loss as such, that localizes this loss in my life, and that reveals a body that works only partly, not like before. Through this process, another image of the damaged body can be built, to hide the horror of the sudden loss, and the patient can then find a body that he will inhabit with functionality, projects, desire (p. 165).

5.2. Patients without Disturbances of the Body Schema

There is another unimaginable dimension that the tearing of the image unravels: the desire of the other, desire that can never be given an image as such but that can be incarnated in a body part which has lost its image and functionality. While the left hemisphere lesion would leave intact the imaginary process which can then ignore the handicap, the right hemispheric lesion would affect the imaginary process itself, not just the ideal image that had been built. The image, then, cannot be rebuilt, and

cannot cover the unimaginable desire of the other; the object of this desire appears instead in the real body: incarnation.

While the patient who suffered a left hemispheric lesion covers his disability with white, the patient who suffered a right hemispheric lesion represents the physical alteration of his body crudely (p. 108); it would even be "this seemingly direct access to the seriousness of their situation" [« cet accès apparemment direct à la gravité de leur situation »] that would preclude these patients from integrating this knowledge; they are then anosognosic (p. 111). This can only be understood from the idea that the patient is not without knowing that he is hemiplegic, although he denies it consciously (p. 121). The misrecognition of hemiplegia would "only" be "apparent" [« apparente »] (p. 173), and the question that anosognosia asks us does not only involve determining what the non brain-damaged subject knows of the body, and what the brain-injured subject ignores of it; on the contrary, knowledge of the body in the normally neurotic subject is a misrecognition and the relation of the Right Hemisphere Syndrome subject to his body removes such misrecognition: this body, this paralyzed hand, becomes unimaginable, it is no longer part of the image of the patient's body, but incarnates an object of desire (pp. 124, 162, 164, 174).

6. *Conclusion*

Here, the ambiguity of the term "object" is instructive: either the patient's hemiplegic limb is experienced by him as a real object, an inert material thing or a thing which has a life of its own; or this limb is experienced by the patient as an object of desire, an object shaped by the desire of the other, an object whose status depends on the gaze of the other (p. 49). While a left hemispheric lesion would maintain an *image of an altered body*, with which the patient must compose, a right hemispheric lesion would rather provoke an *alteration of body imagery*. One way or another, therefore, while the image of the body knots matter and desire into a form to which the subject identifies himself, tearing up the image reveals the body as object: matter which desire does not innervate anymore, or incarnation of the desire of the other. That is the whole relationship of the body with desire, of one's body image with the desire of the other, that brain injuries shake in a way that can only be revealed by an approach which avoids both cerebralizing and psychologizing, an approach out of the cognitive field, such as a psychoanalytic practice that Morin articulates to neurology in order to listen to patients teaching her what is for them the subjective effects of their brain lesion.

A NOTE ON OUR BOOK REVIEW POLICY

We will accept book reviews for publication each issue. Authors wishing to submit book reviews are urged to write with the above interdisciplinary framework firmly in mind. All books *solicited* from publishers will be sent to selected individuals for review. JMB also accepts unsolicited reviews. Reviews should be absent of all titles except the name of the work reviewed, author of work reviewed, place of publication, publisher, date of latest publication, number of pages, and cost. Any individual wishing to submit a review should contact our Book Review Editor for further information: Steven E. Connelly, Ph.D., Department of English, Indiana State University, Terre Haute, Indiana 47809. Email: sconnelly@isugw.indstate.edu

JMB is abstracted or indexed in: *Cultures, Langues, Textes: La Revue de Sommaires; Current Contents (Social and Behavioral Sciences); EMBASE/Excerpta Medica; International Bibliography of Book Reviews; International Bibliography of Periodical Literature; Linguistics and Language Behavior Abstracts; Physics Abstracts; Psychiatric Rehabilitation Journal; PsychINFO/Psychological Abstracts; Research Alert; Social Science Citation Index; Social Work Abstracts; Sociological Abstracts; The Philosopher's Index.* The Journal of Mind and Behavior website is located at www.umaine.edu/jmb/.

The Journal of Mind and Behavior
Winter and Spring 2014 Vol. 35 Nos. 1 and 2

CONTENTS