The Journal of Mind and Behavior

Vol. 37 No. 3 Summer 2016
Vol. 37 No. 4 Autumn 2016

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 multi-directional 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.

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. (Note: we typically do not publish empirical research.)

JMB offers an immediate 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 refereed journal, and all decisions will be made by peer reviewers, 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 post (one copy) or by email:

Raymond Chester Russ, Ph.D., Editor
The Journal of Mind and Behavior
Department of Psychology
University of Maine
5742 Little Hall
Orono, Maine 04469–5742

Tel. (207) 581-2057
Email: jmb@maine.edu

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: (207) 581-2057.

continued inside back cover

The Journal of Mind and Behavior

Editorial Board

CONTENTS OF BACK ISSUES

Volume 2, Number 3, Autumn 1981

Volume 2, Number 4, Winter 1981

Volume 3, Number 1, Winter 1982

Volume 3, Number 2, Spring 1982

Volume 3, Number 3, Summer 1982 (Special Issue)

On Ersatz Teleologists and the Temptations of Rationalism: Some Reactions to Some of the Reactions. *James T. Lamiell, Georgetown University.*
Are We All Clear On What a Mediational Model of Behavior Is? *Joseph F. Rychlak, Loyola University of Chicago.*

Volume 8, Number 3, Summer 1987

Emerging Views of Health: A Challenge to Rationalist Doctrines of Medical Thought. *William J. Lyddon, University of California, Santa Barbara.*
Information, Communication and Organisation: A Post-Structural Revision. *Robert Cooper, University of Lancaster, England.*
How Thoughts Affect the Body: A Metatheoretical Framework. I*rving Kirsch, University of Connecticut, and Michael E. Hyland, Plymouth Polytechnic.*
Consciousness and Commissurotomy: I. Spheres and Streams of Consciousness. *Thomas Natsoulas, University of California, Davis.*

Volume 8, Number 4, Autumn 1987 (Special Issue)

Inhibition in the Brain by Charles E. Ribak, University of California, Irvine (Editor).
Biochemistry of Glycinergic Neurons. *Edward C. Daly, Roudebush VA Medical Center.*
Immunocytochemical Characterization of Glycine and Glycine Receptors. *R.J. Wenthold, National Institutes of Health, and R.A. Altschuler, University of Michigan.*
Distribution of Inhibitory Amino Acid Neurons in the Cerebellum With Some Observations on the Spinal Cord: An Immunocytochemical Study With Antisera Against Fixed GABA, Glycine, Taurine, and β-Alanine. *Ole P. Ottersen and Jon Storm–Mathisen, University of Oslo.*
GABA-Peptide Neurons of the Primate Cerebral Cortex. *Edward G. Jones, University of California, Irvine.*
GABAergic Inhibition in the Neocortex. *K. Krnjevic, McGill University.*
Physiology of GABA Inhibition in the Hippocampus. *R.C. Malenka, R. Andrade, and R.A. Nicoll, University of California, San Francisco.*
Inhibitory Processes in the Thalamus. *M. Steriade and M. Deschenes, Université Laval.*
Neurotransmitter Modulation of Thalamic Neuronal Firing Pattern. *David A. McCormick and David A. Prince, Stanford University School of Medicine.*
What Do GABA Neurons Really Do? They Make Possible Variability Generation in Relation to Demand. *Eugene Roberts, Beckman Research Institute of the City of Hope.*
GABAergic Abnormalities Occur in Experimental Models of Focal and Genetic Epilepsy. *Charles E. Ribak, University of California, Irvine.*
Inhibition, Local Excitatory Interactions and Synchronization of Epileptiform Activity in Hippocampal Slices. *F. Edward Dudek, Tulane University School of Medicine, and Edward P. Christian, University of Maryland School of Medicine.*
Inhibition in Huntington's Disease. *M. Flint Beal, David W. Ellison, and Joseph B. Marcin, Massachusetts General Hospital.*

Volume 9, Number 1, Winter 1988

On Complementarity and Causal Isomorphism. *Douglas M. Snyder, Berkeley, California.*
Methodological Complementarity: With and Without Reductionism. *Michael E. Hyland, Plymouth Polytechnic, and Irving Kirsch, University of Connecticut.*
On Human Nature: A Look at the Subject from Karol Wojtyla's Work The Acting Person. *Paul G. Muscari, State University College of New York at Glens Falls.*
On the Radical Behaviorist Conception of Pain Experience. *Thomas Natsoulas, University of California, Davis.*
From Philology to Existential Psychology: The Significance of Nietzsche's Early Work. *Jerry L. Jennings, University of Pennsylvania.*

Volume 9, Number 2, Spring 1988

Are "Dialogic" Data Positive? *Salomon Rettig, Hunter College.*
Relativity, Complementarity, Indeterminacy, and Psychological Theory. *Mark Garrison, Kentucky State University.*

Contemporary Models of Consciousness: Part I. *Jean E. Bums, Consciousness Research, San Leandro, California.*
The Pluralistic Approach to the Nature of Feelings. *Thomas Natsoulas, University of California, Davis.*
Complementarity and the Relation Between Psychological and Neurophysiological Phenomena. *Douglas M. Snyder, Berkeley, California.*
The Moon Is Not There When I See It: A Response to Snyder. *Mark Garrison, Kentucky State University.*

Volume 11, Numbers 3 and 4, Summer and Autumn 1990 (Special Issue)

Challenging the Therapeutic State: Critical Perspectives on Psychiatry and the Mental Health System by David Cohen, Université de Montréal (Editor).
Introduction: The Medical Model as the Ideology of the Therapeutic State. *Ronald Leifer, Ithaca, New York.*
Toward the Obsolescence of the Schizophrenia Hypothesis. *Theodore R. Sarbin, University of California, Santa Cruz.*
Institutional Mental Health and Social Control: The Ravages of Epistemological Hubris. *Seth Farber, Network Against Coercive Psychiatry.*
Deinstitutionalization: Cycles of Despair. *Andrew Scull , University of California, San Diego.*
Twenty Years Since Women and Madness: Toward a Feminist Institute of Mental Health and Healing. *Phyllis Chesler, College of Staten Island, CUNY.*
The Ex-Patients' Movement: Where We've Been and Where We're Going. *Judi Chamberlin, Ruby Rogers Advocacy and Drop-In Center, Cambridge, Massachusetts.*
AIDS and the Psycho-Social Disciplines: The Social Control of "Dangerous" Behavior. *Mark S. Kaplan, University of Illinois at Urbana–Champaign.*
Therapeutic Professions and the Diffusion of Deficit. *Kenneth J. Gergen, Swarthmore College.*
The Futility of Psychotherapy. *George W. Albee, University of Vermont.*
The Name Game: Toward a Sociology of Diagnosis. *Phil Brown, Brown University* and *Harvard Medical School.*
Subjective Boundaries and Combinations in Psychiatric Diagnoses. *John Mirowsky, University of Illinois at Urbana–Champaign.*
Brain Damage, Dementia and Persistent Cognitive Dysfunction Associated with Neuroleptic Drugs: Evidence, Etiology, Implications. *Peter R. Breggin, Center for the Study of Psychiatry, Bethesda, Maryland.*
The Political Economy of Tardive Dyskinesia: Asymmetries in Power and Responsibility. *David Cohen, Université de Montréal, and Michael McCubbin, York University.*
Electroshock: Death, Brain Damage, Memory Loss, and Brainwashing. *Leonard Roy Frank, San Francisco, California.*
Behavior in a Vacuum: Social-Psychological Theories of Addiction that Deny the Social and Psychological Meanings of Behavior. *Stanton Peele, Mathernatica Policy Research, Princeton, New Jersey.*
The Conceptual Bind in Defining the Volitional Component of Alcoholism: Consequences for Public Policy and Scientific Research. *Richard E. Vatz, Towson State University, and Lee S. Weinberg, University of Pittsburgh.*
False Accusations of Sexual Abuse: Psychiatry's Latest Reign of Error. *Lee Coleman, Berkeley, California.*
Law and Psychiatry: The Problems That Will Not Go Away. *Thomas Szasz, State University of New York, Syracuse.*

Volume 12, Number 1, Winter 1991 (Special Issue)

Consciousness and Commissurotomy: III. Toward the Improvement of Alternative Conceptions. *Thomas Natsoulas, University of California, Davis.*
Perception Without Awareness and Electrodermal Responding: A Strong Test of Subliminal Psychodynamic Activation Effects. *Joseph M. Masling, State University of New York at Buffalo, Robert F. Bomstein, Gettysburg College, Frederick G. Poynton, State University of New York at Buffalo School of Medicine, Sheila Reed, University of Wyoming, and Edward S. Katkin, State University of New York at Stony Brook.*

On the Modeling of Emergent Interaction: Which Will it Be, The Laws of Thermodynamics, or Sperry's "Wheel" in the Subcircuitry? *Larry R. Vandervert, Spokane, Washington.*

Volume 13, Number 1, Winter 1992

Causal Knowledge: What Can Psychology Teach Philosophers? *Evan Fales and Edward A. Wasserman, The University of Iowa.*
Quantum Theory and Consciousness. *Ben Goertzel, University of Nevada, Las Vegas.*
Consciousness and Commissurotomy: IV. Three Hypothesized Dimensions of Deconnected Left-Hemispheric Consciousness. *Thomas Natsoulas, University of California, Davis.*
The Physiology of Desire. *Keith Butler, University of New Orleans.*
Constructivist Psychology: A Heuristic Framework. *William J. Lyddon and James T. McLaughlin, University of Southern Mississippi.*

Volume 13, Number 2, Spring 1992

Residual Asymmetrical Dualism: A Theory of Mind–Body Relations. *Arthur Efron, State University of New York at Buffalo.*
Toward a Model of Attention and Cognition, Using a Parallel Distributed Processing Approach. Part 3: Consequences and Implications of the Sweeping Model. *Gregory Christ, University of Ottawa.*
Being at Rest. *Douglas M. Snyder, Los Angeles, California.*
Neurophysiological Speculations on Zen Enlightenment. *Gerhard H. Fromm, University of Pittsburgh.*
Toward an Improved Understanding of Sigmund Freud's Conception of Consciousness. *Thomas Natsoulas, University of California, Davis.*

Volume 13, Number 3, Summer 1992

How to be a Scientifically Respectable "Property-Dualist." *Ran Lahav, Southern Methodist University, and Niall Shanks, East Tennessee State University.*
A Plea for the Poetic Metaphor. *Paul G. Muscari, State University of New York at Glens Falls.*
Quantum Mechanics and the Involvement of Mind in the Physical World: A Response to Garrison. *Douglas M. Snyder, Berkeley, California.*
Turnabout on Consciousness: A Mentalist View. *R.W. Sperry, California Institute of Technology.*
Intentionality, Consciousness, and Subjectivity. *Thomas Natsoulas, University of California, Davis.*

Volume 13, Number 4, Autumn 1992

Humanistic Psychology, Human Welfare and the Social Order. *Isaac Prilleltensky, Wilfrid Laurier University.*
On Private Events and Theoretical Terms. *Jay Moore, University of Wisconsin–Milwaukee.*
A Teleologist's Reactions to "On Private Events and Theoretical Terms." *Joseph F. Rychlak, Loyola University of Chicago.*
On Professor Rychlak's Concerns. *Jay Moore, University of Wisconsin–Milwaukee.*
Appendage Theory—Pro and Con. *Thomas Natsoulas, University of California, Davis.*
Freud on Dreams and Kosslyn on Mental Imagery. *Derek Drakoulis Nikolinakos, Temple University.*

Volume 14, Number 1, Winter 1993

Altered Sensory Environments, Altered States of Consciousness and Altered-State Cognition. *Joseph Glicksohn, Tel Aviv University* and *The Open University of Israel.*
CPU or Self-Reference: Discerning Between Cognitive Science and Quantum Functionalist Models of Mentation. *Kim McCarthy, University of Oregon* and *Columbia College Chicago, and Amit Goswami, University of Oregon.*
The Naturalists versus the Skeptics: The Debate Over a Scientific Understanding of Consciousness. *Valerie Gray Hardcastle, Virginia Polytechnic Institute and State University.*
Relativism in Gibson's Theory of Picture Perception. *David M. Boynton, University of Maine.*
A New Kind of Transference. *Lauren Lawrence, The New School for Social Research.*

Volume 14, Number 2, Spring 1993

Some Personal Reflections on the APA Centennial. *Seymour B. Sarason, Yale University.*
Consciousness$_4$: Varieties of Intrinsic Theory. *Thomas Natsoulas, University of California, Davis.*

Volume 15, Number 3, Summer 1994

The New Schizophrenia: Diagnosis and Dynamics of the Homeless Mentally Ill. *Alvin Pam, Albert Einstein College of Medicine.*

A Neural Network Approach to Obsessive–Compulsive Disorder. *Dan J. Stein and Eric Hollander, Mt. Sinai School of Medicine.*

On the Distinction Between the Object and Content of Consciousness. *Thomas Natsoulas, University of California, Davis.*

Quantum Physics and Consciousness, Creativity, Computers: A Commentary on Goswami's Quantum-Based Theory of Consciousness and Free Will. *Michael G. Dyer, University of California, Los Angeles.*

Volume 15, Number 4, Autumn 1994

The Depersonalization of Creativity. *Paul G. Muscari, State University College of New York at Glens Falls.*

The Unconscious: A Perspective from Sociohistorical Psychology. *Carl Ratner, Humboldt State University.*

How the Brain Gives Rise to Mathematics in Ontogeny and in Culture. *Larry R. Vandervert, American Nonlinear Systems.*

An Introduction to Reflective Seeing: Part II. *Thomas Natsoulas, University of California, Davis.*

The Structure of Awareness: Contemporary Applications of William James' Forgotten Concept of "The Fringe." *David Galin, University of California, San Francisco.*

Volume 16, Number 1, Winter 1995

Introduction to "Newton's Legacy for Psychology." *Brent D. Slife, Brigham Young University.*

Waiting for Newton. *Thomas H. Leahey, Virginia Commonwealth University.*

Ripples of Newtonian Mechanics: Science, Theology, and the Emergence of the Idea of Development. *Brian Vandenberg, University of Missouri–St. Louis.*

Psychology and Newtonian Methodology. *Piers Rawling, University of Missouri–St. Louis.*

Newtonian Time and Psychological Explanation. *Brent D. Slife, Brigham Young University.*

Temporality and Psychological Action at a Distance. *Richard N. Williams, Brigham Young University.*

Newton, Science, and Causation. *James E. Faulconer, Brigham Young University.*

Can Post-Newtonian Psychologists Find Happiness in a Pre-Paradigm Science? *Paul A. Roth, University of Missouri–St. Louis.*

Volume 16, Number 2, Spring 1995

Some Developmental Issues in Transpersonal Experience. *Harry T. Hunt, Brock University.*

Monistic Idealism May Provide Better Ontology for Cognitive Science: A Reply to Dyer. *Amit Goswami, University of Oregon, Eugene.*

On the Quantum Mechanical Wave Function as a Link Between Cognition and the Physical World: A Role for Psychology. *Douglas M. Snyder, Los Angeles, California.*

Consciousness and Commissurotomy: VI. Evidence for Normal Dual Consciousness? *Thomas Natsoulas, University of California, Davis.*

Volume 16, Number 3, Summer 1995

Consciousness, Thought, and Neurological Integrity. *Grant Gillett, University of Otago Medical School.*

Unsolvable Problems, Visual Imagery, and Explanatory Satisfaction. *Marc F. Krellenstein, New School for Social Research.*

Postmodernity and Consciousness Studies. *Stanley Krippner, Saybrook Institute, and Michael Winkler, University of Denver.*

A Radical Reversal in Cortical Information Flow as the Mechanism for Human Cognitive Abilities: The Frontal Feedback Model. *Raymond A. Noack, San Diego, California.*

Consciousness$_3$ and Gibson's Concept of Awareness. *Thomas Natsoulas, University of California, Davis.*

Volume 16, Number 4, Autumn 1995

The Internet and Research: Explanation and Resources. *David A. Allie, Phoenix Systems.*

Chaos and Related Things: A Tutorial. *Bruce J. West, University of North Texas.*
The Copenhagen Interpretation. *Henry Pierce Stapp, University of California, Berkeley.*
Quantum Mechanics, Chaos and the Conscious Brain. *Chris King, University of Auckland.*
Science of Consciousness and the Hard Problem. *Henry Pierce Stapp, University of California, Berkeley.*
Nonlinear Brain Systems With Nonlocal Degrees of Freedom. *Gordon G. Globus, University of California, Irvine* and *Catholic University of Brasilia.*
Magic Without Magic: Meaning of Quantum Brain Dynamics. *Mari Jibu, Okayama University Medical School* and *Notre Dame Seishin University, and Kunio Yasue, Notre Dame Seishin University.*
Quanta within the Copenhagen Interpretation as Two-Neuro-Algorithm Referents. *Larry Vandervert, American Nonlinear Systems.*
The Brain and Subjective Experience: Question of Multilevel Role of Resonance. *Paul D. MacLean, NIMH Neuroscience Center at St. Elizabeths.*
Nonlinear Dynamics and the Explanation of Mental and Behavioral Development. *Paul van Geert, University of Groningen.*
Nonlinear Neurodynamics of Intentionality. Walter J. Freeman, *University of California, Berkeley.*
Dynamics and Psychodynamics: Process Foundations of Psychology. *Hector C. Sabelli, Center for Creative Development, Linnea Carlson–Sabelli, Rush University, Minu Patel, University of Illinois at Chicago, and Arthur Sugerman, Center for Creative Development.*
Phase Transitions in Learning. *Günter Vetter, Michael Stadler, and John D. Haynes, University of Bremen.*

Volume 18, Number 4, Autumn 1997

A Neuromuscular Model of Mind with Clinical and Educational Applications. *F.J. McGuigan, Institute for Stress Management, United States International University.*
The Presence of Environmental Objects to Perceptual Consciousness: An Integrative, Ecological and Phenomenological Approach. *Thomas Natsoulas, University of California, Davis.*
Wholeness as the Body of Paradox. *Steven M. Rosen, College of Staten Island/CUNY.*
William James and the Challenge of Methodological Pluralism. *Stephen C. Yanchar, Brigham Young University.*
Ideas About a New Psychophysiology of Consciousness: The Syntergic Theory. *Jacobo Grinberg–Zylberbaum, National Autonomous University of Mexico* and *National Institute for the Study of Consciousness.*

Volume 19, Number 1, Winter 1998

The Case for Intrinsic Theory: III. Intrinsic Inner Awareness and the Problem of Straightforward Objectivation. *Thomas Natsoulas, University of California, Davis.*
Analysis of Adverse Behavioral Effects of Benzodiazepines with a Discussion on Drawing Scientific Conclusions from the FDA's Spontaneous Reporting System. *Peter R. Breggin, Center for the Study of Psychiatry and Psychology.*
Defining "Physicalism." *Robert M. Francescotti, San Diego State University.*
The Physics of Metaphysics: Personal Musings. *Aleksandra Kasuba, New York City, New York.*

Volume 19, Number 2, Spring 1998

States of Consciousness and Symbolic Cognition. *Joseph Glicksohn, Bar-Ilan University.*
The Easy and Hard Problems of Consciousness: A Cartesian Perspective. *Frederick B. Mills, Bowie State University.*
Tertiary Consciousness. *Thomas Natsoulas, University of California, Davis.*
The Foundation Walls that are Carried by the House: A Critique of the Poverty of Stimulus Thesis and a Wittgensteinian–Dennettian Alternative. *Wendy Lee–Lampshire, Bloomsburg University.*
Dynamic Interactionism: Elaborating a Psychology of Human Possibility and Constraint. *Jack Martin and Jeff Sugarman, Simon Fraser University.*
On Behaviorism, Theories, and Hypothetical Constructs. *Jay Moore, University of Wisconsin, Milwaukee.*

Volume 21, Numbers 1 and 2, Winter and Spring 2000 (Special Issue)

Brain, Knowledge, and Self-Regulation by Asghar Iran-Nejad, University of Alabama (Editor).

Foreword. *Asghar Iran-Nejad, University of Alabama.*

Introduction: The Current State of the Biofunctional Theory of Cognition. *Suzanne Hidi, University of Toronto.*

Bartlett's Schema Theory and Modern Accounts of Learning and Remembering. *Asghar Iran-Nejad and Adam Winsler, University of Alabama.*

Bartlett, Functionalism, and Modern Schema Theories. *William F. Brewer, University of Illinois at Urbana–Champaign.*

Sources of Internal Self-Regulation with a Focus on Language Learning. *Yasushi Kawai, Hokkaido University, Rebecca L. Oxford, Columbia University, and Asghar Iran-Nejad, University of Alabama.*

Response to "Sources of Internal Self-Regulation with a Focus on Language Learning." *Susan R. Schapiro, University at Buffalo, SUNY.*

Knowledge, Self-Regulation, and the Brain–Mind Cycle of Reflection. *Asghar Iran-Nejad, University of Alabama.*

Keep the Solution, Broaden the Problem: Commentary on "Knowledge, Self-Regulation, and the Brain–Mind Cycle of Reflection." *Richard S. Prawat, Michigan State University.*

The Biofunctional Theory of Knowledge and Ecologically Informed Educational Research. *George G. Hruby, University of Georgia.*

Rethinking the Origin of Morality and Moral Development. *Stacey Alldredge, Emmanuel College, and W. Pitt Derryberry, Michael Crowson, and Asghar Iran-Nejad, University of Alabama.*

Models of Moral Development. *Stephen J. Thoma, University of Alabama.*

A Nonlinear, GA-optimized, Fuzzy Logic System for the Evaluation of Multisource Biofunctional Intelligence. *Abdollah Homaifar, Vijayarangan Copalan, and Lynn Dismuke, North Carolina A&T State University and Asghar Iran-Nejad, University of Alabama.*

Commentary on "A Nonlinear, GA-optimized, Fuzzy Logic System for the Evaluation of Multisource Biofunctional Intelligence." *Gerry Dozier, Auburn University.*

The Nature of Distributed Learning and Remembering. *Asghar Iran-Nejad, University of Alabama and Abdollah Homaifar, North Carolina A&T State University.*

Commentary on "The Nature of Distributed Learning and Remembering." *Edward W. Tunstel, Jr., California Institute of Technology.*

The Brain Between Two Paradigms: Can Biofunctionalism Join Wisdom Intuitions to Analytic Science? *Eleanor Rosch, University of California, Berkeley.*

Knowledge Acquisition and Education. *Merlin C. Wittrock, University of California, Los Angeles.*

Issues in Self-Regulation Theory and Research. *Paul R. Pintrich, University of Michigan, Ann Arbor.*

Heeding Prawat and Hruby: Toward an Articulation Between Biofunctional and Post-modern Theories of Human Experience. *Jerry Rosiek and Asghar Iran-Nejad, University* of Alabama.

Volume 21, Number 3, Summer 2000 (Special Issue)

Toward a Unified Psychology: Incommensurability, Hermeneutics, and Morality by Stephen C. Yanchar and Brent D. Slife, Brigham Young University (Editors).

The Problematic of Fragmentation: A Hermeneutic Proposal. *Stephen C. Yanchar and Brent D. Slife, Brigham Young University.*

Progress, Unity, and Three Questions about Incommensurability. *Stephen C. Yanchar, Brigham Young University.*

Are Discourse Communities Incommensurable in a Fragmented Psychology? The Possibility of Disciplinary Coherence. *Brent D. Slife, Brigham Young University.*

On What Basis are Evaluations Possible in a Fragmented Psychology? An Alternative to Objectivism and Relativism. *Kristoffer B. Kristensen, Brent D. Slife, and Stephen C. Yanchar, Brigham Young University.*

Overcoming Fragmentation in Psychology: A Hermeneutic Approach. *Frank C. Richardson, University of Texas at Austin.*

Fragmentation, Hermeneutics, Scholarship, and Liberal Education in Psychology. *Jack Martin, Simon Fraser University.*

Putting It All Together: Toward a Hermeneutic Unity of Psychology. *Stephen C. Yanchar and Brent D. Slife, Brigham Young University.*

Volume 21, Number 4, Autumn 2000

Consciousness and Conscience. *Thomas Natsoulas, University of California, Davis.*

Experiences of Radical Personal Transformation in Mysticism, Religious Conversion, and Psychosis: A Review of the Varieties, Processes, and Consequences of the Numinous. *Harry T. Hunt, Brock University.*

Self-Organization in the Dreaming Brain. *Stanley Krippner, Saybrook Graduate School and Research Center, and Allan Combs, University of North Carolina at Asheville.*

Eliminativist Undercurrents in the New Wave Model of Psychoneural Reduction. *Cory Wright, University of Mississippi.*

Causation and Corresponding Correlations. *William V. Chambers, Experior Assessments.*

Volume 22, Number 1, Winter 2001

Epistemic Unification. Mitchell R. Haney, Missouri Western State College, and Herman *E. Stark, South Suburban College.*

Historical Origins of the Modern Mind/Body Split. *Richard E. Lind, Springfield, Missouri.*

The Case for Intrinsic Theory: V. Some Arguments from James's Varieties. *Thomas Natsoulas, University of California, Davis.*

Right Brain Damage, Body Image, and Language: A Psychoanalytic Perspective. *Catherine Morin, Institut National de la Santé et de la Recherche Médicale, Stéphane Thibierge, Université de Poitiers, and Michel Perrigot, Hôpital La Salpêtrière.*

A Spinozist Approach to the Conceptual Gap in Consciousness Studies. *Frederick B. Mills, Bowie State University.*

Volume 22, Number 2, Spring 2001

The Split-Brain Debate Revisited: On the Importance of Language and Self-Recognition for Right Hemispheric Consciousness. *Alain Morin, Ste-Foy, Québec, Canada.*

The Case for Intrinsic Theory: VI. Incompatibilities within the Stream of Consciousness. *Thomas Natsoulas, University of California, Davis.*

Apart from Genetics: What Makes Monozygotic Twins Similar? *George Mandler, University of California, San Diego* and *University College London.*

The Concept of Mental Illness: An Analysis of Four Pivotal Issues. *Robert L. Woolfolk, Princeton University.*

Is Crime in the Genes? A Critical Review of Twin and Adoption Studies of Criminality and Antisocial Behavior. *Jay Joseph, La Familia Counseling Service.*

Volume 22, Number 3, Summer 2001

On the Intrinsic Nature of States of Consciousness: Attempted Inroads from the First-Person Perspective. *Thomas Natsoulas, University of California, Davis.*

Structural Causation and Psychological Explanation. *Jeffrey Hershfield, Wichita State University.*

Conceiving Simple Experiences. *Michael V. Antony, University of Haifa.*

Free Will and Events in the Brain. *Grant R. Gillett, Bioethics Center, University of Otago.*

Can Dynamical Systems Explain Mental Causation? *Ralph D. Ellis, Clark Atlanta University.*

Volume 22, Number 4, Autumn 2001

Metaphor and Consciousness: The Path Less Taken. *Joseph Glicksohn, Bar-Ilan University.*

Complexity Theory, Quantum Mechanics and Radically Free Self Determination. *Mark Stephen Pestana, Grand Valley State University.*

The Affiliation of Methodology with Ontology in a Scientific Psychology. *Matthew P. Spackman and Richard N. Williams, Brigham Young University.*

The Process of Knowing: A Biocognitive Epistemology. *Mario E. Martinez, Institute of Biocognitive Psychology.*

The Concrete State: The Basic Components of James's Stream of Consciousness. *Thomas Natsoulas, University of California, Davis.*

The Concrete State Continued. *Thomas Natsoulas, University of California, Davis.*

Volume 23, Numbers 1 and 2, Winter and Spring 2002 (Special Issue)

Choice and Chance in the Formation of Society: Behavior and Cognition in Social Theory by Robert E. Lana, Temple University.

Volume 23, Number 3, Summer 2002

Volume 23, Number 4, Autumn 2002

Volume 24, Number 1, Winter 2003

Volume 24, Number 2, Spring 2003

Volume 24, Numbers 3 and 4, Summer and Autumn 2003

Volume 28, Number 1, Winter 2007

Volume 28, Number 2, Spring 2007

Volume 28, Numbers 3 and 4, Summer and Autumn 2007

Volume 29, Numbers 1 and 2, Winter and Spring 2008 (Special Issue)

An Evolutionary Perspective on Happiness and Mental Health. *Bjørn Grinde, Norwegian Institute of Public Health.*

Meditation on a Mousetrap: On Consciousness and Cognition, Evolution, and Time. *Stephen E. Robbins, Fidelity Information Services.*

Teleology and the Meaning of Life. *Osamu Kiritani, New England Institute for Cognitive Science and Evolutionary Studies.*

Consistent Nihilism. *Jasper Doomen, Leiden University.*

Book Review: *Islam and Science: The Intellectual Career of Nizām al-Dīn al-Nīsābūrī* by Robert G. Morrison. Reviewed by John Walbridge, Indiana University.

Volume 33, Numbers 3 and 4, Summer and Autumn 2012

Computers, Persons, and the Chinese Room. Part 2: The Man Who Understood. *Ricardo Restrepo, Instituto de Altos Estudios Nacionales.*

A Theory of Hemispheric Specialization Based on Cortical Columns. *Robert A. Moss, Bon Secours St. Francis Hospital, Ben P. Hunter, Dhara Shah, and T.L. Havens, Forest Institute of Professional Psychology.*

Dreaming: Physiological Sources, Biological Functions, Psychological Implications. *Matthew Merced, The George Washington University* and *Independent Practice, Washington D.C.*

Counterfactuals, Belief, and Inquiry by Thought Experiment. *Jonathan Leicester, The Royal Prince Alfred Hospital.*

Déjà Vu Explained? A Qualitative Perspective. *Luke Strongman, Open Polytechnic of New Zealand.*

The Equilibration of the Self and the Sense of Sublation: Spirituality in Thought, Music, and Meditation. *Ed Dale, Stockton Hall Psychiatric Hospital.*

Book Review: *The Spiritual Gift of Madness: The Failure of Psychiatry and the Rise of the Mad Pride Movement* by Seth Farber. Reviewed by Richard Gosden, Bingie, NSW Australia.

Volume 34, Number 1, Winter 2013

The Use of the Classical Twin Method in the Social and Behavioral Sciences: The Fallacy Continues. *Jay Joseph, Private Practice, Oakland, California.*

Explaining Consciousness: A (Very) Different Approach to the "Hard Problem." *Paul F. Cunningham, Rivier University.*

Psychotherapy and the Brain: The Dimensional Systems Model and Clinical Biopsychology. *Robert A. Moss, Bon Secours St. Francis Hospital.*

The Flow of Time as a Perceptual Illusion. *Ronald P. Gruber, Stanford University Medical Center, and Richard A. Block, Montana State University.*

Book Review: *Clinical Psychology: An Introduction* by Alan Carr. Reviewed by Geoffrey L. Thorpe, University of Maine.

Volume 34, Number 2, Spring 2013

Deep Naturalism: Patterns in Art and Mind. *Liz Stillwaggon Swan, Mercyhurst University.*

Aristotle and Modern Cognitive Psychology and Neuroscience: An Analysis of Similarities and Differences. *James M. Stedman, University of Texas Health Science Center San Antonio.*

Mentalism as a Radical Behaviorist Views It — Part 1. *J. Moore, University of Wisconsin –Milwaukee.*

The Locus of Stimuli Meaning in the Influence of Attention on Movement: Meaning-Dependent Response Activation Model. *J.L. Vilchez, Universidad Internacional de La Rioja.*

Problematizing Tye's Intentionalism: The Content of Bodily Sensations, Emotions, and Moods. *Juan J. Colomina, The University of Texas at Austin.*

Book Review: *On Orbit and Beyond: Psychological Perspectives on Human Spaceflight* by Douglas A. Vakoch [Editor]. Reviewed by George Michael, Westfield State University.

Volume 34, Numbers 3 and 4, Summer and Autumn 2013

Mentalism as a Radical Behaviorist Views It — Part 2. *J. Moore, University of Wisconsin – Milwaukee.*

The Neurobiology of Transference. *Aslihan Sayin, Gazi University, Ankara, Turkey, and Mehmet Emin Ceylan, Üsküdar University, Istanbul, Turkey.*

Psi and the Problem of Consciousness. *George Williams, Federal Communications Commission–Washington DC.*

Critical Notice: *Butterfly in the Typewriter: The Tragic Life of John Kennedy Toole and the Remarkable Story of* A Confederacy of Dunces by Cory MacLauchlin. Reviewed by Leslie Marsh, University of British Columbia.

Critical Notice: *How Things Shape the Mind: A Theory of Material Engagement* by Lambros Malafouris. Reviewed by Duilio Garofoli, Eberhard Karls Universität Tübingen and Senckenberg Research Institute.

Volume 35, Numbers 1 and 2, Winter and Spring 2014

Feeling. *Jason Brown, New York City.*

ADHD as Emergent Institutional Exploitation. *Lincoln Stoller, Mind Strength Balance, Inc.*

Experimental Methods for Unraveling the Mind–Body Problem: The Phenomenal Judgment Approach. *Victor Yu. Argonov, Pacific Oceanological Institute of the Russian Academy of Sciences*

Critical Notice: *Radicalizing Enactivism: Basic Minds without Content* by Daniel D. Hutto and Erik Myin. Reviewed by Tom Froese, National Autonomous University of Mexico.

Critical Notice: *What to Believe Now: Applying Epistemology to Contemporary Issues* by David Coady. Reviewed by Andrew Alexandra, University of Melbourne.

Book Review: *Transference and Countertransference Today* by Robert Oelsner [Editor]. Reviewed by William Fried, Private Practice, New York City.

Book Review: *Schéma Corporel, Image du Corps, Image Spéculaire. Neurologie et Psychanalyse* [*Body Schema, Body Image, Specular Image. Neurology and Psychoanalysis*] by Catherine Morin. Reviewed by Dorothée Legrand, CNRS, Ecole Normale Supérieure.

Volume 35, Number 3, Summer 2014

Knowing How it Feels: On the Relevance of Epistemic Access for the Explanation of Phenomenal Consciousness. *Itay Shani, Kyung Hee University.*

Development of the Self in Society: French Postwar Thought on Body, Meaning, and Social Behavior. *Line Joranger, Telemark University College.*

Expressivism, Self-Knowledge, and Describing One's Experiences. *Tero Vaaja, University of Jyväskylä.*

"Feeling what Happens": Full Correspondence and the Placebo Effect. *André LeBlanc, John Abbott College* and *Concordia University.*

Book Review: *The Peripheral Mind: Philosophy of Mind and the Peripheral Nervous System* by István Aranyosi. Reviewed by Michael Madary, Universität Mainz.

Volume 35, Number 4, Autumn 2014

Conscious States of Dreaming. *Luke Strongman, Open Polytechnic of New Zealand.*

Higher-Order Thoughts and the Unity of Consciousness. *Lowell Friesen, Booth University College.*

Expansion of the Concept of Mental Disorder in the DSM–5. *Guy A. Boysen and Ashley Ebersole, McKendree University.*

Intentionality and the Aristotelian–Thomistic View of Concepts. *Thomas L. Spalding, University of Alberta, James M. Stedman, University of Texas Health Science Center San Antonio, Curtis Hancock, Rockhurst University, and Christina L. Gagné, University of Alberta.*

Book Review: *Conservatism and Pragmatism in Law, Politics, and Ethics* by Seth Vannatta. Reviewed by Luke Philip Plotica, Virginia Polytechnic Institute and State University.

Volume 36, Numbers 1 and 2, Winter and Spring 2015

A Radical Embodied Approach to Lower Palaeolithic Spear-making. *Duilio Garofoli, Eberhard Karls Universität Tübingen.*

Is That Me? Sense of Agency as a Function of Intra-psychic Conflict. *Travis A. Riddle, Columbia University, Howard J. Rosen, Columbia University, and Ezequiel Morsella, San Francisco State University* and *University of California, San Francisco.*

Forms of Momentum Across Time: Behavioral and Psychological. *Timothy L. Hubbard, Fort Worth, Texas.*

Singular Thought: The Division of Explanatory Labor. *Andrei Moldovan, University of Salamanca.*

The Journal of Mind and Behavior

Vol. 37 Nos. 3 and 4

Summer and Autumn 2016

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. Social psychology–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

Summer and Autumn 2016 Vol. 37 Nos. 3 and 4

CONTENTS

The Journal of Mind and Behavior
Summer and Autumn 2016, Volume 37, Numbers 3 and 4
Pages 183–198
ISSN 0271–0137

Non-Human Origins of Human Perception in the Pre-Pleistocene

Gregory C. Hoffmann and Michael S. Gordon

William Paterson University

In this essay we argue that the human perceptual and sensory mechanisms, which have been described as part of the emergence of our species during the Pleistocene, are part of a much earlier evolutionary trend. Evidence for the pre-human development of our perceptual systems is explored using the comparative literature of non-primates and non-mammals. Furthermore, we argue that evolutionary psychology theorists have tended to misconstrue the mechanisms of perception through an anthrocentric lens. Other lines of thought contend that much of hominid cognition and perception is evolutionarily unique to the point that a broad cognitive discontinuity exists between humans and other species. While the emergence of our species during the Pleistocene clearly has a significant influence on the human brain and mind, it is our contention that perception, and, arguably, the basis of most cognition, is related to much more longstanding environmental constraints as they impacted biological development. Comparative evidence from primates, other mammals, and non-mammalian species, in addition to an evaluation of evolutionary forces and history, are used in support of this argument. The human mind seems to be ancient in its architecture having been sculpted by longstanding and pre-human ecological constraints originating in perceptual mechanisms that significantly pre-date the Pleistocene.

Keywords: evolutionary psychology, perception, animal cognition

Evolutionary psychology has been used to promote hypotheses for several critical mechanisms in the development of perception–action links (e.g., Cosmides and Tooby, 2013; McBurney, Gaulin, Devineni, and Adams 1997; New, Cosmides, and Tooby, 2007; Tooby and Devore, 1987). Frequently these hypotheses emphasize

The authors gratefully acknowledge Drs. Mike Abrams of New York University and Lena Struwe of Rutgers University who helped inspire this investigation and Dr. Thomas Heinzen who provided his guidance and consultation as we embarked on this project. We thank two anonymous reviewers and Dr. Raymond Russ for their thoughtful commentary and guidance, and Drs. Robert Kurzban, Dan O'Brien, and William Mace for their early feedback on the manuscript. Finally, the authors acknowledge the support of the William Paterson University Summer Research Stipend awarded by the College of Humanities and Social Sciences. Correspondence concerning this article should be addressed to Michael S. Gordon, Ph.D, Department of Psychology, William Paterson University, 300 Pompton Road, Wayne, New Jersey 07470. Email: gordonm10@wpunj.edu

evolutionary forces that shaped the dawn of humanity during the Pleistocene epoch (approximately two million to 11,000 years ago). As humans became a species and entered the cognitive niche (Tooby and Devore, 1987), the environment and necessities of day-to-day life required a certain set of perceptual and cognitive abilities to enable survival and sustainability. Modern humans seem to have developed a cognitively unique genus with the advent of symbolic and linguistic systems (Penn, Holyoak, and Povinelli, 2008a, 2008b). Moreover, some perspectives have emerged from cognitive archeology and philosophical anthropology which suggest that the cognitive evolution in hominins is based on an emphasis in sociality (Sterelny, 2007), and subsequently the emergence of language and tool use in modern humans (e.g., Garofoli, 2015; Huffman, 1986; Ingold, 1996; Penn, Holyoak, and Povinelli, 2008a, 2008b).

As Penn and colleagues have argued, there is a discontinuity between human cognition and the rest of the animal kingdom (Penn, Holyoak, and Povinelli, 2008a, 2008b). While it is important to consider the evolutionary pressures during the Pleistocene that directly influenced human cognitive developments, one can speculate that substantial portions of the cognitive architecture, and our human "program," had evolved in the pre-Pleistocene (before humans emerged as a species). In particular, cognition would have been shaped in response to the perceptual systems and the information that those systems afford. Hence, it may be important to review some of the Pleistocene-based mechanisms that have been proposed and to consider what pre-human perceptual architecture and evolutionary forces may have shaped human cognition (e.g., Shaw and Kinsella–Shaw, 2012; Swenson and Turvey, 1991; Turvey and Carello, 2012). With respect to this review, we will revisit some of the specific challenges of using Pleistocene vs. pre-Pleistocene models that have been proposed (e.g., Heyes, 2012; Panksepp and Panksepp, 2000; Sterelny, 2007). Further analysis will investigate particular examples in perception–action cycles as well as the cognitive systems supporting these processes. These examples will be referenced with specific regard to how perception–action cycles may have formed from a thermodynamic perspective (Swenson and Turvey, 1991). Finally, some conclusions will be considered that reflect this extended view of adaptive mechanisms and how that might shape our understanding of human perception and behaviors.

In this essay we will specifically address the perceptual mechanisms posited by evolutionary psychologists to exist within the human cognitive representation system and we will examine evidence that is, perhaps, indicative of other species harboring such perceptual mechanisms. Each of these examples of perceptual mechanisms has been speculated as part of human evolutionary history and adaptation. However, using the tools of cross-species comparison and phylogenetic records, we believe that much of these "human" adaptations can be better understood as part of more general mammalian and pre-mammalian evolutionary trends. As we trace through this evidence we will clarify some of the problems associated with using an anthrocentric approach. Finally we will review some of the pressing issues and caveats of these theories.

Early Perceptual Mechanisms

The detection of motion for the use of finding food sources and avoiding danger has been a long-standing requirement for most species. There are a number of mechanisms evident in humans that would seem to demonstrate our link to pre-mammalian (and pre-Pleistocene) developments. Among these should be considered Reichardt detectors (Lu and Sperling, 1995; Van Santen and Sperling, 1985); collary discharge for motion perception (e.g., McCloskey, 1981; Stark and Bridgeman, 1983; Yasui and Young, 1975); and the use of tau for optic flow and collision detection (Beardsley, Sikoglu, Hecht, and Vaina, 2011; Lee, 1998; Loomis and Beall, 1998). The importance of these mechanisms is that each offers an effective and necessary solution for determining how our bodies move through space, and anticipating safe and effective behavioral responses.

Reichardt detectors, for example, were originally studied in flies (Reichardt and Poggio, 1979), and are evident in insects, reptiles, and humans (Van Santen and Sperling 1984, 1985) to detect motion across a receptive field. These low-level motion detectors use a relatively simple and mechanistic approach to the detection of moving objects. Receptive fields within the eye are compared using the coincidence of their stimulation at a common neuron. When a set of receptive fields representing two positions on the retina are stimulated, then, based on their relative alignment, spacing, and orientation, they may produce concurrent stimulation at the neuron — thus resulting an indication of motion. To stimulate these patches, motion must occur at a very particular trajectory relative to the position of the eye and failing to do so will potentially stimulate one, but not both receptive areas. Consequently, no indication of motion would occur for that detector. By employing a number of Reichardt detectors to represent the gamut of advantageous trajectories, these simple mechanisms can provide a complete range of sensitivities to different velocities and orientations of motion. Motion detection of this sort is important for safety and successful behavior. The cross-species evidence with insects and other non-mammalian species suggests an ancient development of this mechanism, but one that is now evident in humans.

An example of a more sophisticated form of motion detection is the sensitivity to optical rates of expansion in the form of tau. Tau has been suggested as a principle mechanism for diving Gannett's (birds) and mammals to determine self-direction and a range of time-to-arrival estimations within optic flow fields (e.g., Lee and Reddish, 1981; Tresilian, 1991; Warren, 1995). As described with Reichardt detectors, the sensitivity to optic flow has been critical for addressing the perceptual challenges of self-motion detection and collision detection; suggesting a common origin of this mechanism such that it is now shared across multiple mammalian and non-mammalian species. While tau is a direct function of the inverse square law in optics and acoustics, the processing mechanisms for those ecological constraints were necessarily more complicated than Reichardt detectors. With tau, sensitivity to expansion and local deformations of optical patterns became critical for tracking motion (Koenderink, 1985), and thus a more specific type of pattern sensitivity was required by those biological systems that made use of them.

The use of mechanisms to detect corollary discharge further contributes to the perception of optic flow, in as much as one must account for personal body movements of the eyes, head, and neck to appropriately account for personal versus global changes in the environment. Consequently, with an account of corollary discharge, and sensitivity to changes over time with local and global tau, the optic flow field is populated with a reliable indication of critical self-motion information. As an initial set of examples, it is clear that with motion perception many of the mechanisms employed within human perception represent long-standing evolutionary problems. The architecture to solve those problems was necessary prior to our species' emergence — hence the evidence of these mechanisms in the much earlier developed avian and insect species. While it may be easy to find examples of these kinds of mechanisms with relatively complex, but seemingly low-level problems like motion, we would further argue that higher-level cognitive–perceptual mechanisms also represent pre-human, and often pre-mammalian, adaptations.

As one considers the early development of motion sensitivity, these mechanisms may be seen to be related to more general principles of the environment, including physical ecological constraints (e.g., Swenson and Turvey, 1991). Swenson and Turvey have postulated a direct connection of perception–action cycles with thermodynamic principles. In this framework, entropy production and energy conversion from food sources are the primary selective forces for motion and motion perception. By extension, one could formulate that the early perceptual mechanisms for motion may have evolved in response to an organism's consumption and conversion of energy resources. These early motion-based mechanisms would seem to provide a critical means of entropy production across and between biological systems. The capture and digestion of one species serves the continued entropy of the ecology. Over evolutionary time and with progression toward complexity, organisms have increased in perceptual sophistication towards efficient entropy production. Simply put, the thermodynamic systems approach encapsulates evolution within a pre-Pleistocene (and pre-biological) framework predicated on fundamental, physical properties of the universe. Among content-specific traits that potentially indicate a higher-level of perceptual processing, one might include the incredible human proficiency for the visual recognition of faces. Our sensitivity to both configural and feature-level aspects of faces, and the recruitment of specialized neural processing in the fusiform gyrus, suggest the high cognitive demand and capacity with this function, and perhaps even a modular development for its detection (see Bruce and Young, 2012; for a review see also Puce, Allison, and McCarthy 1999). Our demonstrated and specific sensitivity to human faces, despite a normative insensitivity to facial differentiation of non-humans, is a clear indication of a human-specific perceptual demand.

Despite this human-specific function, there is also very clear evidence that the mechanisms for facial processing are not human-specific (e.g., Pascalis and Bachevalier, 1998; Tarr and Gauthier, 2000). Comparative research with *Macaca mulatta* monkeys shows their preferential attention to monkey over human faces (Pascalis and Bachevalier, 1998); and both chimpanzees and monkeys have been found to use neural mechanisms that are analogous to those employed by humans for face detection (e.g., Eifuku, De Souza, Tamura, Nishijo, and Ono, 2004; Parr, Hecht, Barks, Preuss, and Votaw, 2009; but see Perrett et al., 1988). Chickens have also been found to be responsive to face stimuli in recognizing con-specifics (Rosa–Salva, Farroni, Regolin, Vallortigara, and Johnson, 2011; Rosa–Salva, Regolin, and Vallortigara, 2010, 2012). Consequently, it seems that with facial recognition among conspecifics, our high-level sensitivity to face information is not specifically a human-based cognitive trait. Certainly face perception is critical, and the capacity to adapt somewhat unique processing to support our expertise with this ability is apparent in humans. With respect to this review, however, we merely highlight that our expertise for faces seems to reflect the evolutionary turn of a human-specific modification of an already effective and critical perceptual strategy that is well-evident across several species.

Cognitive Maps, Representation, and Wayfinding

Evidence of higher-level cognitive–perceptual functions have also been identified in spatial mapping and processes (Krasnow et al., 2011; McBurney, Gaulin, Devineni, and Adams., 1997; New, Krasnow, Truxaw, and Gaulin, 2007). It appears that females have a spatial gathering advantage over males, and that this capacity represents the use of cognitive mapping and allocentric environmental representations. Presumably our sexual dimorphism of cognitive mapping may have evolved in ancestral humans because of sex-based social roles in hunting and gathering. This would potentially reflect a human disconnect as an extension of the unique sociality of our species (e.g., Sterelny, 2007). The habitual involvement in gathering by females, and hunting by males, is hypothesized to have supported a sexual dimorphism in spatial processing: humans encode the locations of higher-quality resources more efficiently than lower-quality resources (New, Krasnow et al., 2007). This gathering navigation theory has also been used to predict the formation of a highly accurate and landmark-based environmental representation by women in order to encode the location of food in familiar environments (New, Cosmides, and Tooby, 2007). In contrast, males are theorized to harbor an advantage in encoding more general, and non-landmark specific, spatial awareness for hunting animal prey in unfamiliar environments (James and Kimura, 1997; Krasnow et al., 2011; New, Krasnow et al., 2007). While socio-cultural adaptations in humans would seem to have produced this sexual dimorphism, New, Krasnow et al. point out that sexual dimorphic foraging is not human-specific. With respect to other primates, there has been a recorded male bias in chimpanzees even though hunting provides only a small percentage of the chimpanzee food source (Stanford, Wallis, Matama, and Goodall, 1994). Consequently, the sexual dimorphism in spatial memory and representation is a product of human evolution, and one that continues to be evident in current spatial attention and memory tasks (Krasnow et al., 2011). However, evidence suggests that (a) there may have been less cultural/functional sexual dimorphism among humans than previously speculated (e.g., Fuentes, 2012) and (b) that, again, the neural dimorphism of spatial processing between the sexes may predate human evolution (Jacobs, Gaulin, Sherry, and Hoffman, 1990).

Of course, sexual dimorphism at a variety of levels is found across mammalian and reptilian sub-species and this neural dimorphism in spatial memory may reflect much more basic and phylogenetically early adaptation. Moreover, the marked differences in male and female hippocampus for spatial processing tends to emerge at four years of age (with girls outperforming boys on various spatial tasks). At age five boys slightly outperform girls; and at age 11, boys significantly outperform girls in spatial testing (Linn and Petersen, 1985). Taken together, these findings may be used to suggest that, whatever the level of dimorphism, it may become exaggerated with an individual's experience; it does not necessarily reflect a specific evolutionary mechanism; and does not appear to indicate an anthrocentric adaptation. Sexual dimorphism in spatial ability has been documented in a range of mammals. Meadow voles were tested in a variety of mazes and a sexual dimorphism in spatial abilities was revealed wherein males outperformed females (Gaulin and Fitzgerald, 1986, 1989; Kavaliers, Ossenkopp, Galea, and Kolb, 1998). Rats have been tested on a variety of mazes and have demonstrated a similar dimorphism favoring males in spatial ability (Cimadevilla et al., 1999; Dawson, 1972; Einon, 1980; Joseph, Hess, and Birecree, 1978; Seymoure, Dou, and Juraska, 1996). Primate testing with Rhesus monkeys has also shown a sexual dimorphism in spatial ability favoring males (Lacreuse, Herndon, Killiany, Rosene, and Moss, 1999). These findings suggest that sexual dimorphism in spatial cognition exists in other mammals and may be indicative of an earlier evolved sexually dimorphic spatial cognition phenotype in hominins.

These evolutionary branch-offs in mammalian spatial evolution are evidence that a common ancestor in mammalian evolution, and more importantly the human lineage, had a sexually dimorphic perceptual precursor to human spatial cognition. Hence, both the kind of dimorphism and the nature of our spatial cognition appear to be a part of a broader evolutionary trend for wayfinding. Males outperforming females across species is not always the case; there appears to be some ecological constraints at work as female cowbirds have been found to have more accurate spatial representation than males (Guigueno, Snow, MacDougall–Shackleton, and Sherry, 2014). Inasmuch as our cognitive mechanisms derive from our representation and navigation through space, one might further surmise that sex-based differences in spatial cognition represent pre-Pleistocene evolution.

Sex-based differences in perceptual processing occur in many animal clades, however, human spatial representation is unique in that we all possess the cognitive architecture to perceptually capture the environment through the use of geographical maps (Wang and Spelke, 2002). Although the cognitive architecture of humans seems to be among the most sophisticated of all animals, a diverse range of animals from mammals to insects have perceptual and cognitive capabilities that allow for complex navigational skills including those required for efficient foraging and nesting. Wang and Spelke (2002) have noted three distinctive systems that are present in a wide range of animal clades: (1) a path integration system that constantly updates location relative to the environment; (2) a place recognition system that uses template-matching of environmental landmarks; and (3) a re-orientation system that complements the path integration system when it has been disrupted. Of particular interest is that the studies of these key systems suggest that the primary means of efficient foraging and nesting are manifested in a large variety of animals and are hardly restricted to humans, despite a reliance on higher-order spatial processing and representation. Various animal clades including humans are known to rely on geometric patterns to determine their place in space and to categorize new visual stimuli (Rosa–Salva, Sovrano, and Vallortigara, 2014; Tommasi, Chiandetti, Pecchia, Sovrano, and Vallortigara, 2012). Hence, it may be plausible that a spatial gathering mechanism has an earlier phylogenetic origin and one that is non-specific to human evolution.

As noted above, the gathering navigation theory of spatial navigation is that animals preferentially encode the locations of higher-quality energy resources over lower-quality energy resources. Gathering navigation theory is a form of spatial navigation that is also consistent with the notion of efficient entropy production (e.g., Swenson and Turvey, 1991). As per theory and the Swenson and Turvey framework, gathering may be fundamentally guided and shaped by the environment to promote an organism's ability to find higher-quality energy resources. Individual success, as posited in gathering navigation theory, allows for thermodynamic efficiency in that it serves biological and physical environmental demands.

Shared Mechanisms for Cognitive Representations of the World

The seemingly high-level mechanisms that may be mistaken as having developed during the Pleistocene may be a part of a longer evolutionary trend. Cosmides and Tooby (2013) have argued that "Evolutionary psychologists emphasize hunter–gatherer life because it takes a long time for natural selection to build a computational adaptation of any complexity" (p. 203). To clarify one of these mechanisms in humans one can consider the category-specific attentional allotment (e.g., New, Cosmides et al., 2007). In fact, New, Cosmides et al. posited that "…the *human* attention system evolved to reliably detect certain category-specific selection criteria" (2007, p. 16598). It is interesting to consider this claim for it suggests what could be construed as a bias in the field. Namely, that higher-level perceptual knowledge evolved out of a period of *human*-specific changes, rather than an earlier mammalian or pre-mammalian mechanism for detection of this category-specific information. Even tacitly excepting from this statement that there is a basis in the biology with deeper phylogenetic origins, the function is described as functionally human. With respect to this example, category-specific visual attention is understood as a sensitization and capacity to separate animals (both humans and non-human) from a complex visual background. This capacity improves reaction times for detecting a category-specific change in animate objects (humans and non-human animals) within a visual landscape relative to inanimate objects. This category-specific attention to biological objects is thought to support improved performance encoding environments and potentially salient aspects therein. We reiterate that New and his colleagues concluded that this is a critical human visual-attention mechanism and one that supports our expertise using biological motion and categorization. Consequently, these suppositions are consistent with the animate monitoring hypothesis (New, Cosmides et al., 2007): whereby animate objects are consequential time-sensitive elements in a visual scene. In contrast, inanimate features of the environment (plants, rocks, etc.) are less time-sensitive but are vital categories to the human perceiver. While it may be that the human propensity to detect an animate object, such as a predator or prey, over an inanimate object was refined during the Pleistocene, one can speculate that category-specific visual attention to animate objects has deeper ancestral roots. It is this possibility that we can explore by examining parallel abilities in non-human species.

Sensitivity to detecting animate objects over inanimate objects in the visual environment was essential to the survival of humans (New, Cosmides et al. 2007). Moreover, it is crucial to the survival of any animal species that it can be potential prey, or that it has a requirement to catch prey, and thus to be sensitive to biological motion. Infants have a moderately well-developed visual system and can visually attend to objects, people, and events readily over inanimate stimuli in the visual environment; they are prepared to process sensory information about motion and integrate time and space from the very onset of their lives (Frankenhuis, Barrett, and Johnson, 2012). This suggests that sensitivity to biological motion may be an innate capacity in humans. One might further argue that there is a perceptual mechanism for sensitivity to biological motion in humans, and, consequently, that like many human psychological capacities, this adaptation emerged during the Pleistocene. However, it is plausible that sensitivity to biological motion has an earlier phylogenetic origin.

Evidence shows that a variety of non-human animal species have sensitivity to biological motion (self-propelled motion) over inanimate motion. This has been demonstrated in non-human primates, namely baboons (Parron, Deruelle, and Fagot, 2007). It is also striking to find that a similar innate ability has been found in newly hatched chicks (Mascalzoni, Regolin, and Vallortigara, 2010). That advantage and whatever common phylogeny we share with avians would seem to have allowed for, and likely supported, the development of biological motion sensitivity in the pre-Pleistocene.

With respect to biological sensitivity, chimpanzees (e.g., Vonk, Jett, and Mosteller, 2012) distinguish natural categories and form concepts of the environment. These categories would necessarily reflect the social–cognitive requirements of the chimpanzee community, as a means of expressing appetitive and aversive — potentially dangerous stimuli — for the social group (Sterelny, 2007). Hence, these higher level perceptual representations could potentially have contributed to the cognitive disconnect between humans and other mammals. However, non-mammals, such as pigeons, demonstrate a capacity for the categorization of animate/inanimate objects but do not form concepts regarding whether a categorized object is a predator or innocuous (see Fersen and Lea, 1990). Among hominids, it has been found that chimpanzees can discriminate natural categories in the same manner as humans using category-specific knowledge of animate objects (Vonk, Jett, Mosteller, and Galvan, 2013). In addition, the chimpanzees were found to be able to distinguish between animal and non-animal objects on a computer screen. This finding would seem to be indicative of a parallel ability in chimps to humans, and thus may represent that biological category distinction evolved at an earlier period in evolution. Vonk et al. (2013) have further noted that studies have shown a more general ability among primates for recognition of biological objects. Also, the phylogenetically distant mammal, the black bear, has been found to distinguish between natural categories and to form concepts of animal types (Vonk, Jett, and Mosteller, 2012). Hence, while there may be social constraints on category perception and knowledge formation (e.g., Sterelny, 2007), the capacity for this ability does not seem specific to hominids. In fact, as noted by the researchers, black bears were found to be able to learn category-specific knowledge in a shorter period of development than were great apes (Vonk et al., 2012, 2013).

Hence, it is plausible to theorize that the perceptual mechanism of category-specific visual attention to animate objects in humans is not a Pleistocene-based adaptation. The finding that we may share these fairly high-level and abstract perceptual and cognitive abilities for category-specific judgments with other mammals may indicate that humans, and perhaps even our closest relatives on the phylogenetic tree (e.g., ancestral hominids and apes, monkeys and pro-simians — most of which have not been studied for this adaptation), may share a pre-Pleistocene common ancestor with the perceptual and cognitive precursors to our current evolved perceptual phenotype. It may be important to repeat there is some tacit acceptance of that earlier evolutionary history, and yet a persistent bias to describe the traits as grounded within the Pleistocene. Given the evidence supporting an earlier evolutionary history, that bias to use the Pleistocene, and the dawn of humanity, in theorization of mechanisms is potentially misleading and inconsistent with the necessity for a consideration of more longstanding ecological constraints.

Social primates have been found to actively follow the visual gaze of conspecifics as they apprehend their environment to look for food and watch out for predators (Tomasello, Call, and Hare, 1998; Tomasello, Hare, and Fogleman, 2001). It is of adaptive significance for social primates to follow social cues that warn conspecifics of potential food or predators (McNelis and Boatright–Horowitz, 1998). Social cues, such as this example with visual gaze, may not be of the same significance to modern humans as they were to ancestral hunter–gatherers. No longer are most modern humans required to forage, hunt, or be vigilant for predators as would have been required of our Pleistocene ancestors. Nevertheless, humans have certain behaviors that other primates possess and we also share perceptual and cognitive mechanisms that are derived from shared ecological demands (Haun, Jordan, Vallortigara, and Clayton, 2010; Lauder, 1994).

Human Cognition in the Pleistocene

"Modification and perhaps specialization of the digestive tract and dentition to take advantage of these new food sources provides an explanation for the hominid radiation that took place two to three million years ago separating hominids into several different 'specialists'"(Tooby and DeVore, 1987, p. 212). This intermediate hominid adaptive specialization would have ended when one diverging branch advanced far enough into the cognitive niche that its general solution to local adaptive problems proved superior to the specialists.

Tooby and DeVore (1987) have argued that the Pleistocene era has been most critical to the development of human cognition. It is, in fact, undeniable that the later development of the neocortex has supported a dramatic change in the complexity and capacity for cognitive and meta-cognitive abilities (see Marino, 2000). It is also evident that this more recent adaptation of the brain during the Pleistocene is fundamental to modern human cognition. Hence, while it is important to accept and embrace the more recent changes in cognition during the Pleistocene, there are also some foundational reasons to speculate that perception and actions may have evolved as part of an evolutionary trend that significantly predates the Pleistocene. Whether we consider *Australopithecines* (4 million years), the *Ardipithecines* (5.5 million years), stretching back almost 13 million years to *Pierolapithecus*, the requirements of successful sensory processing and motor behaviors would seem to have remained relatively stable. Simply put, the physics for somatosensation, audition, and vision, and the biochemistry for taste and smell are stable and would have created a stable environmental context subserving the development of the senses. And none of these sensory challenges would be affected by an ultrasocial community context unique to modern human cultures (e.g., Sterelny, 2007). Moreover, the senses as they developed would have created, perhaps, the primary root of the cognitive system. Hence, to a certain extent, one might effectively argue that while human cognition is a recent development, the approach to thought and the information on which we cogitate are rooted in very primitive and pre-Pleistocene perceptual functions; and even the thermodynamics of the universe.

The Problem with Evolution as Part of the Human Pleistocene

Above we reviewed several systems that seem to predate human evolutionary development. Similary, Panksepp and Panksepp (2000) argued in their now well-cited paper, *The Seven Sins of Evolutionary Psychology*, that the human emotional and motivational systems in the brain are very ancient, and are indeed more ancient than the Pleistocene. As addressed in this essay and elsewhere, theorists in evolutionary psychology often propose that various mental faculties and abilities are fundamentally human in origin and adaptation. This seems to contradict a breadth of research in comparative neuroscience — hence the argument for an explicit set of guidelines to understand evolutionary forces with the "Seven Deadly Sins." To paraphrase this argument, evolutionary psychology theorists have ignored neurobiological evidence which indicates that higher-level cognitive functions, and many conscious states of the human mind, are rooted in very primitive "hard-wired" affective and motivation systems found in all mammals (and many other phyla). In effect the proposed domain specificity of the human mind may be a result of the domain generality of the neocortex interacting with innate subcortical systems that evolved, and significantly predate, the Pleistocene. Certainly many of these principles apply directly to sensory mechanisms and their evolution. The problem, then becomes, how does one describe adaptations as human-based, or as human-evolution? Simply put, in structuring the problem in this manner we diminish the more longstanding and profound influences of the environment.

When it comes to perception, the fundamental evolutionary challenges have not dramatically altered with the advancements in human cognitive complexity. Essentially we maintain the same underlying needs to identify dangers from safety; food from toxin; comfort from pain. How we go about solving these problems has changed dramatically with cultural complexity and social change, but many of the fundamental perceptual problems have been addressed long before and with great success before any of these human-based progressions occurred. If one were to argue that human perceptual adaptations are, in fact, human-based, then one would necessarily argue that human-based perception is solving for a different environment, and set of challenges, that existed before our species. We contend that the environment does not require a species-specific mechanism, even if our particular niche may vary with phenotype. Perceptual mechanisms are simply and directly required for interaction with that environment rather than a human-specific challenge. To address, interact, and understand our environment for successful perception we necessarily had to have effective perceptual mechanisms to support the emergence of our species. Fundamentally, the development of our species is subject to universal thermodynamic constraints and more specific, but enduring, dangers and comforts of own our niche. Surely, there has been significant adaptation that is specific to the human brain and mind with respect to increasing cognitive complexity, but the crucial adaptations from which this subsequent evolution has taken place is much older. It is on these grounds that we contend that the human perceptual systems and mechanisms are adaptations that are not fundamentally human specific.

Moreover, if we were to suppose that human biological architecture and perceptual mechanisms emerged in the Pleistocene and thereafter, then we might predict unique mechanisms in the human visual, auditory, somatosensory systems relative to our closest genetic lineage via speciation. These mechanisms may well have varied as we incorporated social and technological advances that promoted a unique niche for humans relative to other species (e.g., Huffman, 1986; Ingold, 1996; Sterelny, 2007). Specifically, chimpanzees, or at least monkeys and other close mammals from which we have diverged in the pre-Pleistocene, should present numerous examples of unique speciated perceptual mechanisms or neural processing. While some differences do exist, the architecture, neural plasticity, and major perceptual obstacles (e.g., of finding food, sex, etc.), remain surprisingly consistent across our species. Again, the limited divergence suggests that the major evolutionary forces acted upon and instigated the development of common mechanisms, shared by humans, in the pre-Pleistocene. The examples put forward earlier in this essay indicate that other species may possess certain perceptual mechanisms also found in humans and show that perhaps the many perception–action links thought to exist in humans as having evolved in the Pleistocene may have actually evolved further in the past.

In the extant perceptual and cognitive literature, the human niche has received considerable attention. We have argued that the multiplicitious niche of our forebears that has structured pre-human species of various climates, size and space constraints, and hard-won survival mechanisms, are still a major influence in our genetic code and the foundations of modern perception/cognition. In focusing on adaptations only in our most recent epoch of evolution, we ignore the potentially critical shaping events that have biased our most recent adaptations. To paraphrase a now well-established idea, our brains are structured as a heterogeneity of mechanisms, processes, and parts (e.g., Cosmides and Tooby, 2013). This mélange of pieces presents a seemingly uniform cognitive state, while actually drawing on thousands of tiny biological hacks of nature's code. Limiting the scope of our thoughts to these most recent developmental eras constrains both our theories and imaginations to just the latest adaptations in a struggle for biological success that has endured for millions of years.

References

Beardsley, S. A., Sikoglu, E. M., Hecht, H., and Vaina, L. M. (2011). Global flow impacts time–to–passage judgments based on local motion cues. *Vision Research*, *51*, 1880–1887.

Bruce, V., and Young, A. W. (2012). *Face perception*. New York: Psychology Press.

Cimadevilla, J. M., González–Pardo, H., López, L., Dıaz, F., Cueto, E. G., García–Moreno, L. M., and Arias, J. L. (1999). Sex–related differences in spatial learning during the early postnatal development of the rat. *Behavioural Processes*, *46*, 159–171.

Cosmides, L., and Tooby, J. (2013). Evolutionary psychology: New perspectives on cognition and motivation. *Annual Review of Psychology*, *64*, 201–229.

Dawson, J.L.M. (1972). Effects of sex hormones on cognitive style in rats and men. *Behavior Genetics, 2*, 21–42.

Eifuku, S., De Souza, W. C., Tamura, R., Nishijo, H., and Ono, T. (2004). Neuronal correlates of face identification in the monkey anterior temporal cortical areas. *Journal of Neurophysiology, 91*(1), 358–371.

Einon, D. (1980). Spatial memory and response strategies in rats: Age, sex and rearing differences in performance. *The Quarterly Journal of Experimental Psychology*, *32*(3), 473–489.

Fersen, L. V., and Lea, S. E. (1990). Category discrimination by pigeons using five polymorphous features. *Journal of the Experimental Analysis of Behavior*, *54*(2), 69–84.

Frankenhuis, W.E., Barrett, H.C., and Johnson, S.P. (2012). Developmental origins of biological motion perception. In K. Johnson and M. Shiffrar (Eds.) *People watching: Social, perceptual, and neurophysiological studies of body perception* (pp. 123–140). New York: Oxford University Press.

Fuentes, A. (2012). *Race, monogamy, and other lies they told you: Busting myths about human nature*. Berkeley: University of California Press.

Garofoli, D. (2015). Cognitive archaeology without behavioral modernity: An eliminativist attempt. *Quaternary International*, *405*, 125–135.

Gaulin, S. J., and Fitzgerald, R. W. (1986). Sex differences in spatial ability: An evolutionary hypothesis and test. *American Naturalist*, 127, 74–88.

Gaulin, S. J., and Fitzgerald, R. W. (1989). Sexual selection for spatial–learning ability. *Animal Behaviour*, *37*, 322–331.

Guigueno, M. F., Snow, D. A., MacDougall–Shackleton, S. A., and Sherry, D. F. (2014). Female cowbirds have more accurate spatial memory than males. *Biology Letters*, *10*, 20140026.

Haun, D., Jordan, F. M., Vallortigara, G., and Clayton, N. S. (2010). Origins of spatial, temporal and numerical cognition: Insights from comparative psychology. *Trends in Cognitive Sciences*, *14*(12), 552–560.

Heyes, C. (2012). New thinking: The evolution of human cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2091–2096.

Huffman, T. N. (1986). Cognitive studies of the Iron Age in southern Africa. *World Archaeology*, *18*(1), 84–95.

Ingold, T. (1996). Situating action V: The history and evolution of bodily skills. *Ecological Psychology*, *8*(2), 171–182.

Jacobs, L. F., Gaulin, S. J., Sherry, D. F., and Hoffman, G. E. (1990). Evolution of spatial cognition: Sex-specific patterns of spatial behavior predict hippocampal size. *Proceedings of the National Academy of Sciences*, *87*(16), 6349–6352.

James, T. W., and Kimura, D. (1997). Sex differences in remembering the locations of objects in an array: Location–shifts versus location–exchanges. *Evolution and Human Behavior*, *18*(3), 155–163.

Joseph, R., Hess, S., and Birecree, E. (1978). Effects of hormone manipulations and exploration on sex differences in maze learning. *Behavioral Biology*, *24*(3), 364–377.

Kavaliers, M., Ossenkopp, K. P., Galea, L. A., and Kolb, B. (1998). Sex differences in spatial learning and prefrontal and parietal cortical dendritic morphology in the meadow vole, *Microtus pennsylvanicus*. *Brain Research*, *810*(1), 41–47.

Koenderink, J. J. (1985). Space, form and optical deformations. In D.J. Ingle, M. Jeannerod, and D.N. Lee (Eds.), *Brain mechanisms and spatial vision* (pp. 31–58). The Hague: Springer Netherlands.

Krasnow, M. M., Truxaw, D., Gaulin, S. J., New, J., Ozono, H., Uono, S., Ueno, T., and Minemoto, K. (2011). Cognitive adaptations for gathering-related navigation in humans. *Evolution and Human Behavior*, *32*(1), 1–12.

Lacreuse, A., Herndon, J. G., Killiany, R. J., Rosene, D. L., and Moss, M. B. (1999). Spatial cognition in rhesus monkeys: Male superiority declines with age. *Hormones and Behavior*, *36*(1), 70–76.
Lauder, G. V. (1994). Homology, form, and function. In B. K Hall (Ed.), *Homology: The hierarchical basis of comparative biology* (pp. 151–196). San Diego: Academic Press.
Lee, D. N. (1998). Guiding movement by coupling taus. *Ecological Psychology*, *10*(3–4), 221–250.
Lee, D. N., and Reddish, P. E. (1981). Plummeting gannets: A paradigm of ecological optics. *Nature*, *293*(5830), 293–294.
Linn, M. C., and Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, *56*(6), 1479–1498.
Loomis, J. M., and Beall, A. C. (1998). Visually controlled locomotion: Its dependence on optic flow, three-dimensional space perception, and cognition. *Ecological Psychology*, *10*(3–4), 271–285.
Lu, Z. L., and Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, *35*(19), 2697–2722.
Marino, L. (2000). Convergence of complex cognitive abilities in cetaceans and primates. *Brain, Behavior, and Evolution, 59*(1–2), 21–32.
Mascalzoni, E., Regolin, L., and Vallortigara, G. (2010). Innate sensitivity for self-propelled causal agency in newly hatched chicks. *Proceedings of the National Academy of Sciences*, *107*(9), 4483–4485.
McBurney, D. H., Gaulin, S. J., Devineni, T., and Adams, C. (1997). Superior spatial memory of women: Stronger evidence for the gathering hypothesis. *Evolution and Human Behavior*, *18*(3), 165–174.
McCloskey, D. I. (1981). Corollary discharges: Motor commands and perception. In J.M. Brookhart and V.B. Mountcastle (Eds.), *The handbook of physiology* (pp. 1415–1447). Bethesda: American Physiological Society.
McNelis, N. L., and Boatright–Horowitz, S. L. (1998). Social monitoring in a primate group: The relationship between visual attention and hierarchical ranks. *Animal Cognition*, *1*(1), 65–69.
New, J., Cosmides, L., and Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences*, *104*(42), 16598–16603.
New, J., Krasnow, M. M., Truxaw, D., and Gaulin, S. J. (2007). Spatial adaptations for plant foraging: Women excel and calories count. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1626), 2679–2684.
Panksepp, J., and Panksepp, J. B. (2000). The seven sins of evolutionary psychology. *Evolution and Cognition*, *6*(2), 108–131.
Parr, L. A., Hecht, E., Barks, S. K., Preuss, T. M., and Votaw, J. R. (2009). Face processing in the chimpanzee brain. *Current Biology*, *19*(1), 50–53.
Parron, C., Deruelle, C., and Fagot, J. (2007). Processing of biological motion point-light displays by baboons (Papio papio). *Journal of Experimental Psychology: Animal Behavior Processes*, *33*(4), 381–391.
Pascalis, O., and Bachevalier, J. (1998). Face recognition in primates: A cross-species study. *Behavioural Processes*, *43*(1), 87–96.
Penn, D. C., Holyoak, K. J., and Povinelli, D. J. (2008a). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31*(02), 109–130.
Penn, D. C., Holyoak, K. J., and Povinelli, D. J. (2008b). Darwin's triumph: Explaining the uniqueness of the human mind without a *deus ex machina*. *Behavioral and Brain Sciences*, *31*(02), 153–178.
Perrett, D. I., Mistlin, A. J., Chitty, A. J., Smith, P. A. J., Potter, D. D., Broennimann, R., and Harries, M. (1988). Specialized face processing and hemispheric asymmetry in man and monkey: Evidence from single unit and reaction time studies. *Behavioural Brain Research*, *29*(3), 245–258.
Puce, A., Allison, T., and McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top–down processing on face-specific potentials. *Cerebral Cortex*, *9*(5), 445–458.
Reichardt, W., and Poggio, T. (1979). Figure–ground discrimination by relative movement in the visual system of the fly. *Biological Cybernetics*, *35*(2), 81–100.
Rosa–Salva, O., Farroni, T., Regolin, L., Vallortigara, G., and Johnson, M. H. (2011). The evolution of social orienting: Evidence from chicks *(Gallus gallus)* and human newborns. *PLoS One*, *6*(4), e18802.
Rosa–Salva, O., Regolin, L., and Vallortigara, G. (2010). Faces are special for newly hatched chicks: Evidence for inborn domain-specific mechanisms underlying spontaneous preferences for face-like stimuli. *Developmental Science*, *13*(4), 565–577.

Rosa–Salva, O., Regolin, L., and Vallortigara, G. (2012). Inversion of contrast polarity abolishes spontaneous preferences for face-like stimuli in newborn chicks. *Behavioral Brain Research*, 228(1), 133–143.

Rosa–Salva, O., Sovrano, V. A., and Vallortigara, G. (2014). What can fish brains tell us about visual perception?. *Frontiers in Neural Circuits*, *8*, 119–119.

Seymoure, P., Dou, H., and Juraska, J. M. (1996). Sex differences in radial maze performance: Influence of rearing environment and room cues. *Psychobiology*, *24*(1), 33–37.

Shaw, R. E., and Kinsella–Shaw, J. (2012). Hints of intelligence from first principles. *Ecological Psychology*, *24*(1), 60–93.

Stanford, C. B., Wallis, J., Matama, H., and Goodall, J. (1994). Patterns of predation by chimpanzees on red colobus monkeys in Gombe National Park, 1982–1991. *American Journal of Physical Anthropology*, *94*(2), 213–228.

Stark, L., and Bridgeman, B. (1983). Role of corollary discharge in space constancy. *Perception & Psychophysics*, *34*(4), 371–380.

Stephenson–Jones, M., Samuelsson, E., Ericsson, J., Robertson, B., and Grillner, S. (2011). Evolutionary conservation of the basal ganglia as a common vertebrate mechanism for action selection. *Current Biology*, *21*(13), 1081–1091.

Sterelny, K. (2007). Social intelligence, human intelligence and niche construction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 719–730.

Swenson, R., and Turvey, M. T. (1991). Thermodynamic reasons for perception–action cycles. *Ecological Psychology*, *3*(4), 317–348.

Tarr, M. J., and Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, *3*, 764–770.

Tommasi, L., Chiandetti, C., Pecchia, T., Sovrano, V. A., and Vallortigara, G. (2012). From natural geometry to spatial cognition. *Neuroscience & Biobehavioral Reviews*, *36*(2), 799–824.

Tomasello, M., Call, J., and Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, *55*(4), 1063–1069.

Tomasello, M., Hare, B., and Fogleman, T. (2001). The ontogeny of gaze following in chimpanzees, *Pan troglodytes* and rhesus macaques, *Macaca mulatta. Animal Behaviour*, *61*(2), 335–343.

Tooby, J., and DeVore, I. (1987). The reconstruction of hominid behavioral evolution through strategic modeling. In W.G. Kinzey (Ed.), *The evolution of human behavior: Primate models* (pp. 187–237). Albany: SUNY Press.

Tresilian, J. R. (1991). Empirical and theoretical issues in the perception of time to contact. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(3), 865–876.

Turvey, M. T., and Carello, C. (2012). On intelligence from first principles: Guidelines for inquiry into the hypothesis of physical intelligence (PI). *Ecological Psychology*, *24*(1), 3–32.

Van Santen, J. P., and Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, *1*(5), 451–473.

Van Santen, J. P., and Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, *2*(2), 300–320.

Vonk, J., Jett, S. E., and Mosteller, K. W. (2012). Concept formation in American black bears, *Ursus Americanus*. *Animal Behaviour*, *84*(4), 953–964.

Vonk, J., Jett, S. E., Mosteller, K. W., and Galvan, M. (2013). Natural category discrimination in chimpanzees (*Pan troglodytes*) at three levels of abstraction. *Learning and Behavior*, *41*(3), 271–284.

Wang, R. F., and Spelke, E. S. (2002). Human spatial representation: Insights from animals. *Trends in Cognitive Sciences*, *6*(9), 376–382.

Warren, W. H. (1995). Self–motion: Visual perception and visual control. In W. Epstein and S. J. Rogers, (Eds), *Perception of space and motion* (second edition, pp. 263–325). San Diego: Academic Press.

Yasui, S., and Young, L. R. (1975). Perceived visual motion as effective stimulus to pursuit eye movement system. *Science*, *190*(4217), 906–908.

The Journal of Mind and Behavior
Summer and Autumn 2016, Volume 37, Numbers 3 and 4
Pages 199–218
ISSN 0271–0137

Scientific Realism, Psychological Realism, and Aristotelian–Thomistic Realism

James M. Stedman
University of Texas Health Science Center San Antonio

Matthew Kostelecky
St. Joseph's College at the University of Alberta

Thomas L. Spalding and Christina Gagné
University of Alberta

In this paper, we examine the attractiveness of scientific realism as a philosophical underpinning providing a realist interpretation of psychology. We begin by discussing how psychology arrived at scientific realism as a kind of default position, and discuss some of the advantages of scientific realism relative to non-realist philosophical approaches to psychology. We then raise several potential problems with the naïve adoption of scientific realism for psychology. We argue that these problems show that scientific realism cannot provide a coherent and comprehensive realist underpinning for psychology, and that scientific realism, if taken seriously, has some quite pernicious effects on the field. In particular, scientific realism would divide all of psychology into the scientific and the non-scientific. However, because scientific realism has no clear criteria for what counts as scientific, this distinction, in practice, tends to collapse into a naïve materialist reductionism. We then describe Aristotelian–Thomistic (A–T) realism, and show how it might be adopted to provide a more coherent and comprehensive philosophical underpinning for psychology. We show that the A–T approach avoids the problems that we identified with scientific realism as a philosophical underpinning for psychology. Importantly, unlike scientific realism, the A–T approach maintains a clear realist orientation while providing clear principles for understanding the extent to which humans have epistemological access to reality by matching appropriate methods of inquiry for various subjects of rational inquiry, rather than elevating the scientific method to the status of a principle. Thus, we argue that the A–T approach could provide a solidly realist philosophical underpinning to the whole field of psychology that does not suffer from the defects common to the naïve acceptance of scientific realism.

Keywords: scientific realism, Aristotelian–Thomistic philosophy, metatheory

---

Correspondence concerning this article should be addressed to James M. Stedman, Ph.D., Department of Psychiatry, University of Texas Health Science Center, 7703 Floyd Curl Drive, San Antonio, Texas 78229. Email: Stedman@uthscsa.edu

When pressed to explain their "philosophy of science" for psychology, most psychologists would present some psychological version of scientific realism, holding that the psychological states and their interactions investigated by psychology actually exist in the animal and human subjects studied and can be described and verified through the usual scientific methods. For example, Paul Meehl (1993), who later in his career characterized himself as a scientific realist, said this regarding the general intelligence factor, "g":

> What sort of existential status . . . do we — and ought we — impute to factors? For a scientific realist, a factor is presumably a physical entity possessing a quantitative property. The physical entity exists in the person; hence in the brain. The general intelligence factor g is "in" the CNS. (p. 5)

The scientific realist claim made by Meehl is more than just a philosophy of science claim; it is also an ontological claim about the nature of the reality of "g" and perhaps an epistemological claim about how that reality comes to be known. Hence, the claim of scientific realism in psychology is about the ontology of psychological realities and how these realities are known.

In this paper, we begin by discussing the potential strengths of scientific realism, and what makes it attractive to psychologists. We then point out a set of issues that we believe make scientific realism less attractive as an underpinning for psychology than, perhaps, most psychologists realize. Finally, we point out that there is another form of realism, namely Aristotelian–Thomistic realism, that could underpin psychology and that has some advantages over scientific realism.

Scientific Realism

Any discussion of the strengths of scientific realism must begin with a review of its philosophical predecessors, operationism and logical positivism, two distinct but intertwined underpinnings of psychology dating from the 1920s and, in fact, is still with us today. In 1927, Bridgeman published *The Logic of Modern Physics* in which he proposed an operational analysis of all the concepts in physics (space, time, velocity, mass, etc.) with the aim of *eliminating* all abstractions from the field. He did this because he was convinced that "metaphysical abstractions" had led to serious errors in physics. As a correction, he proposed that all concepts of physics are constituted by and defined by a set of operations; and he asserted the following: "the concept is synonymous with the corresponding set of operations" (1927, p. 5). Criticism forced Bridgeman later to acknowledge the need for including some theoretical constructs.

Logical positivists initially endorsed many of Bridgeman's ideas but, by the late 1930s, they rejected his brand of operationism as an oversimplification.

Ultimately, they came to see that no set of operations ever exhausts the meaning of a scientific concept. Hence, though operationism and logical positivism shared similar ideas, the two have distinct histories.

Although physics never embraced operationism, psychology did. In fact, somewhat ironically, it was Herbert Feigel, a member of Vienna Circle and a logical positivist, who influenced Harvard psychologist E. G. Boring and his student S. S. Stevens to take up Bridgeman's operational attitude. During the same time frame, Edward Tolman was influenced by Moritz Schlick, founder of the Vienna Circle, and Tolman became an eager advocate of operationism. However, Tolman turned Bridgeman's operationism on its head. Whereas Bridgeman sought to eliminate "metaphysical concepts," Tolman eventually introduced the notion of the "intervening variable," referring to psychological theoretical constructs he was attempting to operationalize. As stated by Green (1992) in his excellent review of operationism: "Where Bridgeman sought to rid science of metaphysical concepts, Tolman sought to legitimize them by attaching them to related physical operations" (p. 296). Many influential theory-building psychologists, with the notable exception of B. F. Skinner, took up Tolman's version of operationism, although, as clearly demonstrated in the 1945 *Psychological Review* symposium on operationism (Langfeld, 1945), there was little consensus on the specifics of the program. In fact, Rogers (1989) has argued there were a number of operationisms in vogue during the 1930s and onwards.

By the late 1940s, operationism was coming under heavy attack within psychology. MacCorquodale and Meehl (1948) and Cronbach and Meehl (1955) presented refined statements in an attempt to answer criticisms; however, serious critiques continued by Koch (1959) and others. By the 1970s, both operationism and logical positivism were largely discredited within philosophy itself (see, e.g., Fotion, 1995). Nevertheless, Tolman's view of operationism refused to die and elements of operationism continue even to this day (see Green, 1992, for an extensive and convincing treatment of this topic). One reason why Tolman's version of operationism has been so long-lived is that it is, in many respects, simply a reductionist approach — psychological entities are legitimated only by being tied to the purely physical. We will return to this point when we discuss the ways in which psychologists have adopted scientific realism.

As noted above, many serious psychological theorists, including many former operationalists, such as Meehl, began to advocate for scientific realism as a replacement for operationism. Others, who followed postmodernist trends asserting that "foundationalism" was dead, presented philosophies based on antirealist positions, such as social constructionism (Gergen, 1994) and hermeneutic theory (Messer, Sass, and Woolfolk, 1988), though these were clearly minority positions among psychologists. We will focus on scientific realism, and it should be noted again that scientific realism is a philosophy of science, an ontology, and an epistemology.

Scientific Realism: The General Theory

Chakravartty (2011) describes scientific realism as follows: "Scientific realism is a positive epistemic attitude toward the content of our best theories and models, recommending belief in both the observable and unobservable aspects of the world described by the sciences" (p. 1). This description implies variations in positions taken by scientific realists; and, indeed, that is the case. Most claim epistemic truth or approximate truth of scientific theories or, at least, aspects of those theories. Most make the same claim for both observable and unobservable elements of scientific theories. In the end, all variations hold that the best scientific theories produce true, or increasingly true, descriptions of the world.

Chakravartty (2011) identifies three features common to all scientific realism positions: (1) a metaphysical commitment to a "mind-independent" existence of the world; (2) semantic realism regarding scientific theories and findings; and (3) an epistemological commitment to scientific inquiry leading to knowledge of the world as it exists outside the mind. The metaphysical commitment to the "mind-independent" existence of the world is central to scientific realism, and, indeed, any form of realism. The realist metaphysical commitment was opposed by various figures in modernity, so much so that some questioned whether one could be certain about the existence of the world outside the mind. By contrast, scientific realism maintains that atoms, molecules, universes, and biological entities all exist as "things" and interact. The psychological extension of scientific realism is that psychology's theoretical entities can and should be considered in exactly the same way as the other scientific "things."

Semantic realism relates to the truth claims made by scientific theories and their findings. This is perhaps the heart of the scientific realist position, and is described by Chakravartty (2011) thusly: "Claims about scientific entities, processes, properties, and relations, whether they be observable or unobservable, should be construed literally as having truth values, whether true or false" (p. 3). Unpacking this statement from the perspective of a psychologist adapting scientific realism, we find that "scientific entities" includes all objects of scientific inquiry from the subatomic to political attitudes to any "thing" that can be investigated via the scientific method of knowledge generation. "Properties" are, of course, all of those qualities or traits belonging to "things," from size and shapes to motivations, feeling states, and beliefs, again all subject to study according to the scientific method. "Processes" and "relations" can perhaps be combined in the sense that these terms describe the interactions of "things," be it chemical binding of molecules or regularities found in mob violence.

"Observable" generally refers to the "things" of science that are directly perceivable by the senses and measureable in some form. "Unobservable" refers to many of the "things" of science which are theoretical and not directly observable but are postulated as causes: electrons or features of black holes, and from the psychologists' perspective, anger, beliefs, etc. Finally, the claim of scientific realism is that scientific propositions, laws, and even hypotheses are subject to a "true or false" judgment. If true, or approximately true, scientific realism claims them as true of the mind-independent world.

Finally, scientific realism holds that science produces true knowledge of the world. This is the epistemological claim, and is closely tied to semantic realism. It is countered by antirealist claims of all stripes which attack some or all of the three features noted above.

So, how do scientific realists support these claims? Putnam (1975) holds for the "miracle" argument which is basically that the principles of scientific realism are the only ones that explain the vast accomplishments of science; without scientific realism the history of science can only be explained as a miracle (see Lyons, 2003, Frost–Arnold, 2010, for more recent presentations). The corroboration argument (Hacking, 1983) is brought forth to support scientific realism's claim regarding unobservable entities. This claim is that, if a theoretical entity can be detected by two or more different instruments of measurement, this constitutes a basis for defending realism. Explanationists (Kitcher, 1993; Psillos, 1999), again in support of the reality of unobservables, assert that, when our best theories require unobservables and their interactions to predict, that fact supports the reality of those unobservables. Entity realism (Cartwright, 1983; Giere, 1988) makes a similar argument about unobservables, specifically, that, when the interactions of unobservables can be manipulated and produce the same outcomes over a number of trials, this fact supports their existence as real.

It should be noted that antirealists have produced counters to all of the supporting arguments stated above and, in fact, have provided counters to all of the principles of scientific realism (Gergen, 1994; Laudan, 1981). Indeed, these arguments in many respects actually require a mind-independent world in order to provide any real support for scientific realism, and thus appear to argue, at least partially, in a circle. Nevertheless, the vast majority of scientists across all disciplines continue to believe they produce knowledge regarding a mind-independent world. Psychologists are no different from physicists in that regard. Now that we have examined scientific realism as a general theory, it is time to turn attention to scientific realism as psychologists see it.

Weaknesses of Scientific Realism as an Underpinning for Psychological Realism

Although above we have argued that most psychologists assume some form of scientific realism as a valid underpinning for the knowledge claims that they make (as do most scientists), and furthermore that scientific realism is a better approach than most of the others that have recently held sway (see also, Stedman, Sweetman, and Hancock, 2008), here we wish to point out several weaknesses of scientific realism precisely as a realist underpinning for psychology. To begin, we argue that most psychologists are, naively, psychological realists. In short, most psychologists adopt a realist perspective that exactly parallels scientific realism, in that they believe (a) that a world outside the mind exists; (b) that the theoretical claims of psychology are properly claimed to have truth values; and (c) that psychological inquiry can and does lead to truth or at least to increasing truth over time. Given this set of beliefs, it is easy to see the attraction of scientific realism for such psychologists — it is a way of understanding psychology as being contiguous with the rest of science, such that arguments for scientific realism become, de facto, arguments for psychological realism, as well. Thus, psychologists are free to go about their business, assuming that the question of the ontological, semantic, and epistemic status of their theoretical concerns has been (at least) largely settled.

However, we see several weaknesses in scientific realism that are particularly damaging to its ability to serve psychology's need for a realist underpinning, the first three of which relate to the final of the three general characteristics outlined above, the epistemological claim, and the final one of which relates to the nature of the arguments supporting scientific realism. First, because scientific realism makes an epistemological claim specifically for scientific inquiry, and is, at best, agnostic about the value of any non-scientific inquiry, scientific realism can only underpin a science. This limitation of the scope is critically important to scientific realism, because without this (and the related semantic claim that scientific theoretical entities — but not necessarily other theoretical entities — are to be treated as having truth values), scientific realism would simply be unqualified realism. Yet, historically one of the great difficulties for psychology as a field has been to determine whether and to what extent psychology is, in fact, a science. This argument continues to the present day (see, e.g., Kraus, 2013). Indeed, scientific realists can and do disagree about what counts as science, and hence, what kinds of claims can be warranted. Given the critical importance of the science vs. non-science distinction in scientific realism, one might expect that scientific realism has a clear way of making the distinction. Indeed, it seems clear that scientific realism must assume that there is some way of determining science from non-science, even though no such criteria are included in the descriptions of scientific realism given above. It is critical to note that the difference between science and non-science itself is not subject to scientific inquiry, and therefore, by scientific realism, whatever conclusion one might reach about whether something is or is not a science is of questionable warrant as a truth claim! Clearly, even if it is true, scientific realism cannot fully warrant psychological realism without auxiliary assumptions about how to differentiate science from non-science, and the assumption or demonstration that psychology in fact falls on the science side of that differentiation. Thus, adopting scientific realism as the underpinning for psychological truth claims sets up the possibility of a division between the scientific and non-scientific within psychology, as well. Is all of psychology a genuine science? Is none? Is only some? What are we to make of the truth claims of the "non-scientific" part? To put it bluntly, then, there is a serious question as to what extent scientific realism can actually underpin psychology as a whole. Scientific realism provides no resources for answering the question about what parts of psychology have warranted truth claims, and so scientific realism re-introduces or exacerbates the science vs. non-science divisions within psychology. Thus, psychologists' assumption that scientific realism is obviously convertible to psychological realism is problematic.

Second, often the epistemological claim of scientific realism is made even stronger, such that any truth claim not advanced as a result of scientific inquiry is deemed to be, not only potentially unwarranted, but simply nonsense (Williams, 2015, pp. 5–17 provides some history of this way of thinking, along with some recent examples; note also the parallel to earlier positivist and operationist claims). Thus, the stakes of the possible division of science vs. non-science within psychology are raised dramatically, as all non-science psychology (should any exist) is presumptively nonsense. This idea is problematic for many reasons, not just because of its effects on psychologists. For one thing, if taken seriously, it would mean that scientific realism itself is nonsense, as scientific realism is established and supported (as noted above) via various philosophical arguments, not via scientific inquiry. Furthermore, the idea that non-scientific claims should be treated as nonsense is rather clearly not something established via the scientific method! Note further that neither mathematical nor logical truths are established via the scientific method, so if taken seriously this scientific realism claim would also rule out mathematics and logic. Still, this rather obvious problem has not kept some psychologists and others from claiming that any non-scientific claim is presumptively nonsense (again, see Williams, 2015).

Third, despite the fact that scientific realism (as extended to psychological realism) might be taken to underpin psychological theoretical entities, psychologists (and others) very often resort to reductionist positions to make the transition from psychological entities to physical entities. It would be trivially easy to pull hundreds of quotations from the literature (indeed, see the discussion of "g" with which we open this paper, or the discussion of Tolman's operationist-inspired attempt to legitimize psychological constructs by linking them to physical operations), but as one example, consider the following, taken from Tuomela (1977, p. 39, original emphasis and punctuation): "It should be emphasized that thoughts thus understood are *actualities*, some goings on in the person's 'mind' (in the first place, though they ultimately will presumably turn out to be propositional brain processes)." It is important to notice that even here, in a book dedicated to a functionalist and realist description of human action and the theories thereof, the author finds it necessary to put the word *mind* in quotation marks, and to claim immediately that thoughts, despite being actualities, are not actualities in their own right, but instead are brain processes. In addition, of course, there is the difficulty about what exactly could count as a "propositional brain process" in the first place, and how some state or series of states of purely physical objects could be “propositional” in themselves, in any meaningful sense. However, our main point here is that even in a book dedicated to what is a psychological realism approach to human action, the author feels compelled to distinguish what is actually scientific from what is not, and that this distinction, in effect, comes down to what is purely physical. Indeed, the author wants to say that thoughts are actualities — real things — but cannot even get to the end of the sentence without conceding that they are not! Thus, scientific realism here does *not* in fact function to underpin psychological realism, except in the special case of a purely reductive, physicalist approach to psychology — in essence, scientific realism underpins psychological realism by simply removing the psychological.

One might reasonably ask why one should jump to reductionism, if scientific realism supports psychological realism — why, if scientific realism allows us to treat our theoretical constructs as real and as having truth values, do we feel the need to remove the “psychological” from those constructs? The obvious attraction of reduction is that the resulting level of inquiry is, naively speaking, clearly scientific. But, there are well-known and serious philosophical problems with attempting to understand psychological entities (such as thoughts or social processes) in purely reductionist ways (see, e.g., Madden, 2013, Ch. 5 for a very readable overview), so if scientific realism can only underpin psychology to the extent that psychology is reducible to a “real science,” then scientific realism provides no actual realist underpinning for psychology as a whole.

Of course, this strong wish to link to something that unquestionably counts as “science” is a rather obvious consequence of conditioning the acceptance of a realist orientation on the scientific mode of inquiry. Given that scientific realism provides no way to show whether something is or is not a science, the only real option for psychologists is to try to link psychological entities to something that will be accepted as scientific, even in the absence of any argument. Hence, psychologists often assume that the psychological entities entirely reduce to something physical, and then presume that scientific realism allows them to take a realist position about anything physical. Seen in this light, psychological realism is simply a purely reductionist and materialist metaphysical and epistemological philosophy of psychology.

Finally, as we noted above, most of the arguments specifically for scientific realism (the miracle argument, the corroborationist argument, the explanationist argument) are primarily epistemological arguments in favor of the reality of unobservable scientific entities, but they all, at a deep level, rely on the fundamental realist position that a mind-independent reality exists. Thus, for example, if we remove that basic realist assumption then the miracle argument loses all force — the consistent success of science is not a miracle if, in fact, it reflects only a mind-constructed, non-independent reality. And, obviously, if the corroboration of a measurement by different techniques only reflects the consistency of a mind-created reality, it tells us nothing about the mind-independent reality of that measurement. In short, arguments for scientific realism mostly assume realism, and focus on establishing that science provides an epistemological framework that allows one to access reality. Of course, other arguments for realism can also support scientific realism, so long as they provide support to the basic realist assumption without undermining the epistemological arguments for scientific realism, and similarly, they could provide support to psychological realism as so long as they neither undermine scientific realism or the adoption of scientific realism as a way to get psychological realism. Unfortunately, as we argue above, some of scientific realism's own claims make the extension to psychological realism questionable. Hence, a realist position that is more deeply grounded in metaphysical argument, as opposed to a primarily epistemological argument, and that could then be clearly extended to psychological entities, would be helpful.

In sum, we argue that scientific realism, despite the advantages that we discussed above relative to anti-realist accounts of science, turns out to be seriously flawed as a way of underpinning psychology, and that these weaknesses derive rather directly from scientific realism's epistemological and semantic claims that condition the reality of theoretical entities on their position within a scientific mode of inquiry. In the next section, we describe a different way of underpinning psychological realism, by adopting an Aristotelian–Thomistic form of realism.

Aristotelian–Thomistic Realism

In this section, we describe Aristotelian–Thomistic (A–T) realism as a set of general metaphysical principles based on a philosophical worldview, and show how this approach (particularly in the work of Thomas Aquinas) can be developed into a form of psychological realism. There are a few critical differences from traditional scientific realism. First, although A–T realism amply meets all three of the criteria put forward by Chakravartty (2011) as definitive of scientific realism, and thus has no problem with scientific realism's claims for the domain of physical science, its claims have a greater scope than do scientific realism's. More specifically, A–T realism does not limit our ability to determine truth to the scientific method (though the scientific method is certainly an appropriate method for finding truth in some areas of inquiry). A–T realism maintains limits on our ability to determine truth as appropriate to the particular topic of investigation. In this regard, it is important to recognize that the scientific method is not "the gold standard," but rather the method that is best adapted to the investigation of the physical. For example, mathematical truths do not need to be established by the scientific method, but by the method of proof. However, mathematical truths (and logical truths) are, in A–T realism, even more certain than any truth of the physical world, no matter how well established via the scientific method (if indeed, any mathematical or logical truth could ever be said to be established via the scientific method). Second, unlike scientific realism, A–T realism is a basic or fundamental outlook regarding the world, which can be developed in a manifold of directions, and allows different sorts of inquiry to cohere with each other. Thus, everything from what today we would call the hard sciences, to the psychological sciences, to the other social sciences, to the structure of scientific inquiry itself, to a study of logical inferences, and many things besides are not kept in isolation from each other. Rather, they are integrated into a holistic picture, in which each topic area has its own proper kinds of truth claims and methodologies. Third, any contemporary form of A–T realism is, at base, a philosophical approach to reality that is acutely aware of the problems of skepticism and induction, and the various forms of antirealism, and proceeds — with full knowledge of these issues — to describe the world as it is.

Those are some of the salient differences between A–T realism and scientific realism, but just what is A–T realism? At its most basic level, it is a comprehensive approach to reality as reality is and as it is manifested to humans by rational inquiry. The sort of "rational inquiry" that is undertaken is going to be different depending on the sort of subject matter that is being inquired into, such that some forms of inquiry are more certain or stable than others, and such that different methods of rational inquiry are more or less appropriate to the different sorts of subject matter. There is a type of hierarchy to different kinds of human knowing from more to less certain, with mathematics being the most certain. The physical sciences are less certain (they are not apodictically true or true a priori), but the physical sciences are still sufficiently grounded to believe that the propositions of the physical sciences are, roughly, true (if open to revision). Even granting such differences in subject areas, we can still distill several common metaphysical principles that range across this comprehensive worldview.

What follows is an introduction to A–T realism by way of presenting several important metaphysical presuppositions or principles.[1] To begin, beings exist and knowledge is of beings. That is to say, there really are things in the world and our knowledge really is of those things. This principle shows the fundamentally realist position of the A–T worldview. Second, our knowledge of things is conveyed to us, fairly reliably, via the senses. A–T realism, then, is grounded on empirical presuppositions. Third, the principle of non-contradiction (i.e., a thing cannot both be and not be at the same time in the same respect) holds across all human inquiry, and its logical corollary, the law of non-contradiction, holds for how we reason about things. Thus, our reasoning follows a pattern that is itself established in reality. Fourth (and this is related to the previous principle), there is a basic mirroring of reality in human cognition. This is not to say that the human cognition is infallible or completely reproduces external reality, but that

[1]This listing is compiled from across the Thomist and Aristotelian corpus. There is no place in either body of writings where either Thomas or Aristotle lay out their metaphysical principles in a wholly systematic way.

cognition operates in a basically reliable way, as the processes of nature do too. The human mind, a part of nature, picks up on the regularities found in nature and internalizes them through physiological and psychological processes. Fifth, and this is perhaps the most distinctive and controversial aspect for a contemporary reader, the whole of reality is looked at through the lenses of actuality and potentiality. To clarify, a given object of inquiry — a tree, say — is analyzed by its current act and the ways in which it could potentially act. This will quickly become an account of things in the watchwords of Aristotelian metaphysics: "form" and "matter," the former of which is a sort of correlate to actuality, while the latter is related to potentiality. The human being, again, a part of nature, is also submitted to this sort of analysis and looked at through the lenses of actuality and potentiality, as is human cognition itself.

To be sure, we can (and various historians of philosophy have done this in different ways) reconstruct an A–T philosophical anthropology, epistemology, or psychology, but those are disciplines which post-date both Aristotle and Aquinas. Instead, we believe that it is more instructive to see what is going on with respect to an account of the human being (philosophical anthropology), a justification or description of knowledge (epistemology), or an account of psychological processes such as human cognition (psychology) as embedded into the larger metaphysical framework of the A–T view. We should note that although we focus in this discussion on human cognition, but it is important to recognize that A–T realism can underpin the whole range of modern psychology, whether human or non-human animals, whether one is primarily interested in cognition or emotion, and so on.

Aristotle's and Thomas's descriptions of human psychological processes are accounts of the metaphysical relationship between the human mind and the things that humans know. The result is a metaphysical account of cognition, precisely because it aims to describe the relation between the knower and the object of scientific inquiry. This is a realist perspective *par excellance*, because both relata are presumed to be real and an account of the relation between the relata is what is sought. Clearly A–T realism is comprehensive in its scope, and, we think, superior to a naïve version of psychological realism, which uncritically accepts scientific realism. What, then, would a psychological realism formed within the backdrop A–T realism look like? We turn to that in our next section.

Aristotelian–Thomistic Psychological Realism

Aristotelean–Thomistic realism accounts for change in the world according to causes and effects, for, according to the whole Aristotelian tradition, knowledge is understood to be knowledge of a thing's cause(s). Now, A–T psychological realism is part of this larger approach to reality but specified down to human psychological processes, wherein those processes are understood to be one kind of change among others (again, we concentrate here on cognitive processes but the A–T analysis applies to other psychological processes, such as the emotions, as well). As such, A–T psychological realism shares the same basic methodological approaches as A–T realism and accounts for cognition as one kind of change among others. While the biologist looks to understand the causal mechanisms that make an organism be the sort of organism it is, the psychologist seeks to uncover the causal mechanisms that allow for a human to cognize an organism (or any external thing). In both cases, understanding the causal story of the thing investigated is key, even if the causal mechanisms differ somewhat between those that govern objects of the external world and those that govern how the mind comes to know an external thing.

On this account, the task of the psychologist is a bit more complicated than that of the scientist who studies the external world (be it a biologist or physicist or whomever), because the causal nexus the psychologist investigates is itself more complicated. That cognitive causal nexus includes the external object, the medium through which the information is relayed, and, especially, the active and passive aspects of the mind that are necessary for human cognition to occur. What immediately follows is a brief exposition of the causal story of how cognition occurs, according to Aristotelean–Thomistic psychological realism.[2]

Let us begin with an ordinary object of inquiry, say, an American Elm tree. How does cognition of this external thing occur on the A–T account? In the first instance, as a form of realism, Aristotelean–Thomistic psychological realism assumes the elm exists independently of any mind inquiring into it.[3] Moreover, the larger A–T worldview dictates that the elm is currently in act in various ways, its "actuality," and has various possible states of actuality that aren't being currently actualized, i.e., its "potentiality." The actuality of the tree, or "form" is conveyed

---

[2]It should be noted, briefly, that while we are calling this an Aristotelian–Thomist psychological realism, what follows is more a description of how Thomas Aquinas, in particular, specified the Aristotelian texts that he had before him. Many years had passed between Aristotle's time and that of Aquinas, and Aquinas was the beneficiary of a robust commentary tradition that had been trying to make sense of Aristotle for more than fifteen hundred years. Aquinas makes use of a more technical vocabulary than what one finds in Aristotle himself, due to this commentary tradition.

[3]The contemporary A–T realist knows that she may be deceived (she has read her Descartes et al.), but thinks that it is simply more rational to get on with the project of inquiring into the object of inquiry, rather than submitting everything to critical scrutiny. The skeptical currents that have had such a profound impact on modern philosophy (and thereby contemporary psychology) are often attributed to Descartes and Hume. In fact, these skeptical ideas have much deeper roots than either Descartes or Hume. Indeed, "Skepticism" had a long and noble academic pedigree even during Antiquity. The medievals, in particular, knew of various kinds of skepticism (often in a more radical form than what one finds in their modern forms) and usually decide that it is simply more rational to describe the processes of human cognition, rather than submit everything to critical doubt. In this sense A–T realism is clearly of a piece with how most contemporary scientists get on with their job of describing reality according to empirical standards rather than seeking to justify the project of knowing external things.

to and impressed upon the proper sense organ(s), either through a medium (like air or water) or immediately through contact, as in touch or taste. Upon initial reception, the form is communicated from the sense organs to the "interior senses," which is how the mind initially unifies disparate sense data and allows for the internal reception and retention of the form. At this point, there is an "intention" in the mind, which is to say that there is a cognition of the tree, but only the tree as a particular sensible object. That intention can be recalled when the object is not present and, as one encounters the elm more and more, one experiences it exhibiting different states of actuality, for example, with leaves in the spring and summer and bare in the winter. These different states of actuality can be recalled by the mind and contrasted to each other.

At this point in the cognitive story, the form of the external thing has been impressed upon what today we would call the brain of the cognizer, but only at the particular level.[4] That is certainly a kind of cognition, but it does not attain the universal status that is required for full-fledged understanding, for which the cognizer needs more than the particular intention of the cognized object. Indeed, for understanding one needs to cognize something about the thing that ranges over a multitude of similar things. For instance, if one had never encountered a tree but then suddenly did, one does not really understand what the tree is by simply sensing and retaining/recalling a sense image of that tree in the mind. That is to say, one does not really understand what the tree is by way of a particular cognition, even though one does know some things about it: at times it has leafy things on it, it has such and such colors, different textures in different parts, and so on. But, until one is able to classify it as a plant, tree, elm, and most specifically, as an American Elm, one does not really understand it. What is needed is a "concept," that is to say, some sort of universal understanding of the tree that accounts for how individual trees fit into a larger classificatory scheme.

The A–T tradition marks an important difference between humans and non-human animals, specifically in the capacity to form full-fledged concepts and then reason about or from those concepts. To account for a concept, the A–T psychologist will turn from the passive aspects of cognition (in which humans and higher animals are quite similar) to something "active" in the human mind itself. Aquinas calls this active component to human cognition the "agent intellect" whereas Aristotle calls it the "active intellect." In spite of this terminological difference both figures are insisting that there must be something

[4]The term "brain" here is anachronistic. The term used by the Aristotelian tradition is typically "soul." Today, "soul" has the connotation of something ghostly or immaterial, whereas for the Aristotelian tradition anything that is alive and material has a soul, for the soul is simply a way of demarcating material things that are alive and material things that aren't. Aquinas does say, however, that there is a particular organ located in "the middle part of the head" that is responsible for this kind of cognition, which is why we've used the word "brain" above.

active in the human mind itself to account for human understanding.[5] The primary active cause in the sensorial process, i.e., the sensed object, is not seen as sufficient to cause understanding, which grasps something universal about the particular sensed object.

The causal cognitive story of human understanding picks up at the point where the sensorial process stops. For higher animals and humans alike, the external object is sensed by means of the actuality or form of the thing being conveyed to the sense organs and from there to the appropriate organs for internal reception and retention of the form. Due to these physiological and psychological processes, the agent intellect has this particular intention before it and is able to see what is universal about that particular intention. As per our previous example, the human intellect is able to form a concept of the American Elm by which it understands why this particular tree is in the same class of things as other American Elm trees.

Another example that might make the issue more evident is that of triangles.[6] Sensible cognitions of triangular things occur all the time for humans and non-human animals alike, but that kind of cognition does not recognize the properties that make a triangle a triangle. So, at the sensorial level, we can cognize the triangular thing in front of us, determine whether to pursue, avoid, manipulate, or ignore it, but, until we have a concept of the triangle, we do not understand the features that make it a triangle, e.g., that it is a three-sided polygon. Coming to understand the triangularity of the triangle completely escapes the sensitive powers, because the external object (i.e., the triangular thing) only causes a particular intention of the thing. When the agent intellect sees what is universal in the particular it "abstracts" the universal from the particular that is presented to it via the sensorial process and forms a concept of the thing. Note here that the universal is not just abstracted in the sense of averaged — an averaged sensory triangle is no more triangular than a single sensible triangle is, and it is just as much a particular. It still has a particular (average) size, color, area, distribution of angles, etc. The key point is that the intellect abstracts from all of those particulars.

Now, this whole process, which results in a concept, is both empirically grounded and realist. It is realist insofar as the mind-independent existence of the object of inquiry is taken for granted and insofar as the term of the inquiry is in accounting for the mind's relation to other things. In terms of being empirically grounded,

---

[5]For Aristotle's use of the term see De anima III, chapters 4 and 5. For Aquinas's see, for example, Summa Theologica Ia, qq. 84–85 — though the term is found throughout his writings. The difference of vocabulary is due to the robust commentary tradition that Aquinas has inherited and which has specified Aristotle's vocabulary considerably. We employ Aquinas's terminology.

[6]Mathematical objects are, according to the A–T worldview, a bit easier to comprehend and use as examples than objects like trees, which itself works as a caution from seeking the same sort of certitude in other avenues of inquiry as one finds in mathematics, as, for example, in our knowledge of the physical world.

there is an Aristotelian dictum that the whole A–T tradition takes as basic: "There is no understanding without a sense image."[7] As such, a concept must always be traceable back to something sensed and conveyed to the intellect via the story we have rehearsed here.

As noted at the beginning of this section, A–T realism accounts for change in the world according to causes and effects, and knowledge is understood to be a knowledge of a thing's cause(s). Cognition is itself one kind of change, among others, for clearly there is a change of some sort in someone who does not cognize something but who then does cognize it. Moreover, there are different sorts of causes and effects referred to in accounting for cognition on the A–T model. At the sensorial level, the primary active cause is the external object which is impressed upon the sensitive cognitive processes and effects a sensible image in the cognizer. This sensible image then stands as the content upon which the intellect operates and from which it abstracts, which is necessary for the production of a concept of the thing.

Advantages of A–T Psychological Realism

In our description of the A–T account of cognitive processes, we found it useful to employ two different examples: one of an American Elm tree and one of triangles. The former is an ordinary object of scientific inquiry, whereas the latter is an object of mathematical inquiry. The latter is easier to comprehend and use as an example, because of the clarity and certainty that come along with mathematical objects, over and against things like trees, which we experience as always in some sort of motion. Aristotle and Aquinas were aware of the difference and used it as a warning to not expect the same kind of certainty in different modes of inquiry.[8] Thus, we can have clear and certain knowledge about many objects of mathematical inquiry, whereas certainty is not as achievable in the physical sciences, and is arguably even less achievable in the social sciences or other areas of rational inquiry. To expect the same level of certainty across different domains is to court error or skepticism or reductionism, and, as such, is not a wise path to follow.

In the first part of the paper we presented a series of complaints about the deleterious effects of an uncritical acceptance of scientific realism by psychologists, the first of which was that scientific realism provides no way to differentiate science from non-science, and thus no way to account for the scientific status of psychology. Aristotelean–Thomistic realism allows for an account of just this issue. In the first instance, psychology is simply not as certain as mathematics, but then neither are physics or biology. Secondly, as noted above, the task of the psychologist

---

[7] Aristotle, De anima, III, 7.

[8] See Aristotle's Ethics, Book I, chapter 3 (1094b12-15) and Aquinas's commentary upon the same text.

is more complicated than that of the biologist and the physicist, because to have knowledge in those domains is just to have knowledge of the causal story of the object of inquiry. In the case of the scientific study of an American Elm tree, the arborist understands how the tree comes about, what the tree's development should look like, how it will reproduce, and so forth. For psychologists, by contrast, the causal story includes aspects of those issues, but also includes an account of how mind and behavior relate to the things in the real world. Of course, the basic realism of the A–T approach includes not only that of psychologists focused on cognitive theory, or on explicitly observable behavioral responses, but also on psychologists interested in emotions, sensation, social and cultural relations. So, the A–T realism includes not only the cognitive account of, say, the concept of an elm, but also an emotional or esthetic response to the elm, or a memory evoked by the elm, or perhaps a specific cultural meaning associated with elms. One should expect the endeavor, by virtue of its complexity, to be more difficult and, yes, less certain. That does not mean that psychology is not a science, but that it is a difficult science that requires careful examination of assumptions, informed by empirical inquiry. Aristotelean–Thomistic realism, then, helps place psychology as a science and as the kind of science it is, while also allowing that different areas of psychology could themselves be more or less amenable to the scientific method, be more or less certain, and so on, but still be perfectly valid subjects for rational enquiry, and for the discovery of, or at least approximation to, truth about real psychological entities and processes.

A second complaint we raised was the extent to which scientific realism often does not admit the truth of any claim that is not advanced by scientific inquiry or, more seriously, that it dismisses any such claim as nonsense. This is, of course, self-referentially incoherent, since scientific realism is itself not the product of scientific inquiry. By contrast, A–T realism is quite assertive about its (reasonable) presuppositions: there are real things in the world and we can investigate them through rational processes. Different methods will be useful for different sorts of things and the scientific method, as practiced today, is entirely consonant with this approach. The scientific method, as incorporated by A–T realism, does not ground anything except the empirical inquiry and experimentation for which it is appropriate. Similarly, mathematical methods are appropriate for mathematical enquiry, logical methods for logical enquiry, and both produce results that are more certain than those of the scientific method (even when the scientific method is entirely appropriately applied to physical sciences). By the same token, philosophical methods are appropriate to philosophical enquiry, but may be more or less certain than the scientific method, depending on the specific area of philosophic enquiry. Again, the certainty or lack thereof is due to the nature of the subject of enquiry, and in all cases the appropriate method is also determined by the nature of the subject of enquiry. The scientific method is not, on the A–T model, converted unwittingly into a metaphysical standpoint. On this view, then, psychological methods (which are often very closely related to the methods of physical sciences, but also often not identical, and in some cases are quite different) are appropriate for psychological enquiry.

Third, we noted a persistent temptation to a very naïve reductionism, largely driven by the desire for psychology to be treated, without dispute, as science. Applying scientific realism to psychology uncritically results in a need to reduce psychology to the purely physical, in order that scientific realism can underpin realism for psychology. Because there is no other way for scientific realism to determine science from non-science, one must assume that the subject of the enquiry is purely physical — very few doubt that science, whatever else it is or does, applies to the purely physical. Thus, in the attempt to provide a realist underpinning for psychology, we remove the psychological as a valid area of enquiry in its own right. Psychology is only valid in as much as it is a "way of talking about" entities that are actually purely physical. In effect, we have to destroy psychology to save it, or at least to save a realist underpinning for it. This is simply not an issue for A–T realism. In the A–T approach, semantic and epistemological realism does not depend on an area of enquiry being amenable to the scientific method, in the way that scientific realism suggests. Instead, A–T realism is based on the whole, coherent A–T approach, from metaphysics through to the physical sciences and on to psychology, social sciences, and other areas of enquiry.

Questions About A–T Realism and Psychology

In this final section, we address several issues which arise from our previous exposition, and which are likely to stand as objections to adopting A–T realism as a way of providing a realist underpinning for psychology. First, we have not provided much detail about why the A–T realist avoids the problem of skepticism. The A–T approach, as we noted, assumes that external things are real and that our sensation of those things is basically reliable. Obviously, Aristotle pre-dates the modern versions of skepticism, and his cognitive theory did not present itself as opposed to such approaches, though this hardly means that his thought or, especially, later appropriations of Aristotelian thought is naïvely realist. Skepticism was an active school of thought in late antiquity (St. Augustine himself famously became a Skeptic for a period) and Aquinas clearly knows about the skeptical challenge to knowledge, but is simply more interested in a descriptive account of knowledge than being agnostic on the issue and thinks that the Aristotelian, descriptive approach has more to commend it. A contemporary advocate of the A–T approach will be very familiar with the skeptical currents that have so profoundly shaped modern philosophy, and with full awareness of these issues, will decide that it is simply more rational to get on with the project of providing an account of reality, than it would be to submit every piece of knowledge or the project of human knowledge as such to critical scrutiny. In this, the A–T realist shares much in common with the contemporary scientist (or scientific realist theorist, for that matter) who will not want to get bogged down in skeptical questions, but wants to describe things. The A–T realist is a realist because of a prior decision that a description of reality as it comes to us via the senses is more advantageous than the laborious and the never-quite-satisfactory task of justifying our knowledge of reality. To be clear, as we argued above, scientific realism shares this prior judgment — that is why, as we noted, the main arguments for scientific realism actually assume realism, and then go on to try to prove that the scientific method is a good way to learn something about the real.

Second, while we have emphasized that the A–T approach is thoroughly realist in orientation, and simultaneously that it limits the semantic and epistemic truth claims by (a) the nature of the subject being investigated and (b) the limitations of the methods attached to those subjects, we have said relatively little about the deeper limits on the human ability to identify the truth. Earlier, we clarified, "This is not to say that human cognition is infallible or completely reproduces external reality, but that cognition operates in a basically reliable way, as the processes of nature do too." It is now time to unpack this statement. Scientific realism limits the human ability to "get to truth" by limiting the realist orientation to those cases in which the scientific method applies. The A–T approach sees the limits in the subjects and their correlated methods, but also in the nature of the human intellect itself. Thus, A–T realism is a very far cry from a naïve realism, indeed. What does it mean to "operate in a basically reliable way, as the processes of nature do too"? For one thing, it means that the human intellect *is* a process of nature, in a broad sense. Just as, in nature, an acorn tends to develop into an oak tree, the human intellect tends to grasp reality. However, just as an acorn might fall into bad soil, or be poisoned by someone and thus not grow into a mature oak, the intellect similarly can fail to obtain truth because the person does not have the proper training, or knowledge, or is misled by someone, or uses the wrong method of inquiry, or simply gets misleading data from the world. In short, "basically reliable" means that the human, in rational inquiry, tends toward grasping the truth, but that tendency can be disrupted, just as any other natural process can be disrupted for a variety of reasons and therefore fail to develop in the usual manner. Thus, A–T realism recognizes both the normal human tendency to grasp reality via rational inquiry, but also the obvious capacity for human error. Importantly, in the A–T view, this pattern is true across all areas of rational inquiry, with appropriate caveats and cautions for the inherent differences among subjects of inquiry. There is no one gold standard that guarantees truth across substantially different domains, nor is there a case to be made for making a metaphysical or epistemological principle out of a single method.

Third, we have claimed that one of the advantages of the A–T approach is that it is a coherent system of ideas that can build all the way from the most fundamental metaphysical notions to philosophy of mind and even to empirical work in psychology. Some might wonder whether a philosophic underpinning based on the

A–T approach makes any difference to how one does psychology. Although the A–T approach is currently somewhat out of fashion, we have found aspects of the A–T approach very helpful in our work in philosophy of psychology (Spalding, Stedman, Hancock, and Gagné, 2014; Stedman, 2013; Stedman, Spalding, and Gagné, 2016; Stedman, Sweetman, and Hancock, 2006, 2009) and in various areas of the empirical psychology of human cognition (Gagné, Spalding, and Kostelecky, in press; Spalding and Gagné, 2013, 2015). Others have also recently found the A–T approach helpful across a wide spectrum of areas within psychology (see e.g., Butera, 2010; DeRobertis, 2011; Freeman, 2008; Prasada and Dillingham, 2006, 2009). And, of course, historically, there was a time when some scholars considered all of psychology to be consistent with the general A–T approach to psychology (see, e.g., Maher, 1909; Mercier, 1918).

Conclusions

We hope to have shown that scientific realism, as commonly adopted as a realist philosophical underpinning for psychology, does not, in fact, provide a coherent and comprehensive underpinning for the whole field of psychology. Rather than giving up on a realist underpinning, however, we propose that psychology would benefit from a rediscovery of the Aristotelian–Thomistic version of realism. We have argued that the A–T view could be applied to modern psychology in a way that avoids the problems of scientific realism, while providing a solidly realist underpinning.

References

Bridgeman, P. W. (1927). *The logic of modern physics*. New York: Macmillan.

Butera, G. (2010). Thomas Aquinas and cognitive therapy: An exploration of the promise of the Thomistic psychology. *Philosophy, Psychiatry, and Psychology, 17*(4), 347–366.

Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Oxford University Press.

Chakravartty, A. (2011). Scientific realism. *The Stanford encyclopedia of philosophy* (pp. 1–25). https://plato.stanford.edu/entries/scientific-realism/

Cronbach, L. T., and Meehl P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

DeRobertis, E. M. (2011). St. Thomas Aquinas' philosophical-anthropology as a viable underpinning for a holistic psychology: A dialog with existential–phenomenology. *Janus Head, 12*(1), 62–91.

Fotion, N. G. (1995). Logical positivism. In T. Honderich (Ed.), *The Oxford companion to philosophy* (p. 508). Oxford: Oxford University Press.

Freeman, W. J. (2008). Nonlinear brain dynamics and intention according to Aquinas. *Mind and Matter, 6*(2), 207–234.

Frost–Arnold, G. (2010). The no-miracles argument for realism: Inference to an unacceptable explanation. *Philosophy of Science*, *77*, 35–58.

Gagné, C. L., Spalding, T. L., and Kostelecky, M. (in press). Conceptual combination, property inclusion, and the Artistotelian–Thomistic view of concepts. In J. Hampton and Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology. Language, cognition, and mind.* New York: Springer.

Gergen, K. (1994). *Realities and relationships: Soundings in social constructionism*. Cambridge, Massachusetts: Harvard University Press.

Giere, R. N. (1988). *Explaining science: A cognitive approach*. Chicago: University of Chicago Press.
Green, C. D. (1992). Of immortal mythological beasts: Operationism in psychology. *Theory and Psychology, 2*, 291–320.
Hacking, I. (1983). *Representing and intervening*. Cambridge: Cambridge University Press.
Kitcher, P. (1993). *The advancement of science: Science without legend, objectivity without illusions*. Oxford: Oxford University Press.
Koch, S. (1959). Epilogue. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 3, pp. 729–788). New York: McGraw–Hill.
Kraus, M. W. (2013). *The psychology of the "psychology isn't a science" argument*. Retrieved from: https://www.psychologytoday.com/blog/under-the-influence/201308/the-psychology-the-psychology-isnt-science-argument
Langfeld, H. S. (1945). Symposium on operationism. *Psychological Review, 52*(5), 241–294.
Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science, 48*, 19–48.
Lyons, T. D. (2003). Explaining the success of a scientific theory. *Philosophy of Science*, 70, 891–901.
Madden, J. D. (2013). *Mind, matter and nature: A Thomistic proposal for the philosophy of mind*. Washington, D. C.: The Catholic University of America Press.
Maher, M. (1909). *Psychology: Empirical and rational (*sixth edition). London: Longmans, Green, & Co.
MacCorquodale, K., and Meehl, P. E. (1948). On the distinction between hypothetical constructs and intervening variables. *Psychological Review, 55*, 95–107.
Meehl, P. (1993). Four queries about factor reality. *History and Philosophy of Psychology Bulletin, 5*, 4–5.
Mercier, D-J. (1918). *The origins of contemporary psychology*. London: R. & T. Washbourne, Ltd.
Messer, S. B., Sass, L. A., and Woolfolk, R. L. (1988). *Hermeneutics and psychological theory: Interpretative perspectives on personality, psychotherapy, and psychopathology*. New Brunswick: Rutgers University Press.
Prasada, S., and Dillingham, E. M. (2006). Principled and statistical connections in common sense conception. *Cognition, 99,* 73–112.
Prasada, S., and Dillingham, E. M. (2009). Representation of principled connections: A window onto the formal aspect of common sense conception. *Cognitive Science, 33,* 401–448.
Psillos, S. (1999). *Scientific realism: How science tracks truth*. London: Routledge.
Putnam, H. (1975). *Mathematics, matter, and method*. Cambridge: Cambridge University Press.
Rogers, T. B. (1989). Operationism in psychology: A discussion of the contextual antecedents and an historical interpretation of its longevity. *Journal of the History of the Behavioral Sciences, 25*, 139–153.
Spalding, T. L., and Gagné, C. L. (2013). Concepts in Aristotle and Aquinas: Implications for current theoretical approaches. *Journal of Theoretical and Philosophical Psychology, 33,* 71–89.
Spalding, T. L., and Gagné, C. L. (2015). Property attribution in combined concepts. *Journal of Experimental Psychology: Learning Memory and Cognition, 41*(3), 693–707.
Spalding, T. L., Stedman, J., Hancock, C., and Gagné, C. L. (2014). Intentionality and the Aristotelian–Thomistic view of concepts. *Journal of Mind and Behavior, 35,* 245–262.
Stedman, J. M. (2013). Aristotle and modern cognitive psychology and neuroscience: An analysis of similarities and differences. *Journal of Mind and Behavior, 34,* 121–132.
Stedman, J., Spalding, T. L., and Gagné, C. L. (2016). Does functionalism offer an adequate account of cognitive psychology? *Journal of Mind and Behavior, 37,* 15–30.
Stedman, J. M., Sweetman, B., and Hancock, C. (2006). Neuroscience and the correspondence theory of truth. *Contemporary Philosophy, 27,* 84–92.
Stedman, J. M., Sweetman, B., and Hancock, C. (2008). Hermeneutics and scientific realism: Rival philosophical claims in psychology. *Contemporary Philosophy, 28*, 40–48.
Stedman, J. M., Sweetman, B., and Hancock, C. (2009). Neuroscience and the problem of abstraction. *Contemporary Philosophy, 29,* 84–92.
Tuomela, R. (1977). *Human action and its explanation: A study on the philosophical foundations of psychology.* Boston: D. Reidel Publishing Company.
Williams, R. N. (2015). Introduction. In R. N. Williams and D. N. Robinson (Eds.), *Scientism: The new orthodoxy.* (pp. 1–21). London: Bloomsbury Academic.

The Journal of Mind and Behavior
Summer and Autumn 2016, Volume 37, Numbers 3 and 4
Pages 219–246
ISSN 0271–0137

Behavior Analytic Pragmatism

J. Moore

University of Wisconsin – Milwaukee

According to pragmatism, the meaning of a philosophical topic is found in its implications and consequences for human affairs. Absent is any assumption that the topic represents some aspect of a metaphysical reality inferred to be beyond human experience and behavior. The present review suggests that the views of metaphysics and scientific verbal behavior found in contemporary pragmatism, with Richard Rorty as the example, are compatible with those found in the behavior analysis of B.F. Skinner.

Keywords: antirepresentationalism, levels of analysis, scientific verbal behavior

Pragmatism is a decidedly American viewpoint in philosophy whose development in the late nineteenth century is often attributed to Charles Sanders Peirce, William James, and John Dewey (Menand, 2002), and whose influence in the late twentieth and early twenty-first centuries is often attributed to the late Richard Rorty. As is well known, James and Dewey contributed to psychology as well as to philosophy. Indeed, pragmatism was instrumental in the development of American functionalism at the close of the nineteenth century, and is often linked with the development of behaviorism during the twentieth century. In regard to behaviorism, Leigland (1999), Moxley (2001, 2001/2002, 2002), and more recently Schoneberger (2016) have reviewed the relation between pragmatism, for example, as represented by Peirce and Rorty, and the behaviorism of B. F. Skinner, known as behavior analysis. The present paper seeks to continue the discussion of the relation between pragmatism and behavior analysis, and for purposes of illustration to compare the pragmatic dimensions of behavior analysis with those of cognitive science.

I thank Ted Schoneberger for helpful comments on an earlier draft. Correspondence concerning this article should be addressed to J. Moore, PhD, Department of Psychology, University of Wisconsin – Milwaukee, Milwaukee, Wisconsin 53201. Email: jcm@uwm.edu

Pragmatism as a Philosophical Orientation

Pragmatism has been described in many ways by many commentators, sometimes as much in terms of what it opposes as what it advocates. The following passage from Dewey (1926) is illustrative:

> [Philosophy's] primary concern is to clarify, liberate, and extend the goods which inhere in the naturally generated functions of experience. It has no call to generate a world of "reality" de novo, nor to delve into secrets of Being hidden from common sense and science. It has no stock of information or body of knowledge peculiarly its own; if it does not always become ridiculous when it sets up as a rival of science, it is only because a particular philosopher happens to be also, as a human being, a prophetic man of science. Its business is to accept and to utilize for a purpose the best available knowledge of its own time and place. And this purpose is criticism of beliefs, institutions, custom, policies with respect to their bearing on good. This does not mean bearing upon the good, as something itself formulated and attained within philosophy. For as philosophy has no private store of knowledge or of methods for attaining truth, so it has no private access to good. As it accepts knowledge of facts and principles from those competent in science and inquiry, it accepts the goods that are suffused in human experience. It has no Mosaic or Pauline authority of revelation entrusted to it. But it has the authority of intelligence, of criticism of these common and natural goods…. (pp. 407–408, italics in original)

Dewey here clearly continued in the tradition of Peirce and James by challenging the dominant philosophical thinking of his time and suggesting an alternative approach. Absent was an emphasis on a metaphysics that extended beyond the domain of human affairs. In its place was a firm commitment to analyses grounded in human experience and behavior.

A second and more modern illustration of pragmatism is the work of Richard Rorty. Much as Dewey had, Rorty (e.g., 1979, 1991) emphasized that pragmatism is therapeutic instead of constructive. It is therapeutic in the sense that it argues the job of philosophers is to clarify and refine an understanding of the processes according to which humans interact with the world, often in causal ways. For Rorty as for Dewey, the meaning and value of philosophical terms invoked in this endeavor follow from their implications and consequences with respect to human experience and behavior. Rorty rejected the idea that philosophers should be concerned with constructing a metaphysical story about how human minds create the world in which humans live, and then with justifying how this story reflects a reality that philosophers know in some privileged way. Rorty employed the umbrella term "antirepresentationalism" in conjunction with his pragmatic views. By this term he meant an account "which does not view knowledge as a matter of getting reality right, but rather as a matter of acquiring habits of action for coping with reality" (Rorty, 1991, p. 1).

Unpacking Rorty's words is useful here. Following Dewey (1926), Rorty (1991) challenged the traditional view of philosophers "who find it fruitful to think of mind or language as containing representations of reality" (p. 2). Thus, Rorty objected to assuming that philosophical language mirrors or represents an underlying reality in some Platonic, Cartesian, or Kantian sense, and that Truth was a matter of the fidelity of this representation, through the correspondence between words and the inferred, underlying reality. Rorty further objected to reducing philosophy to debates between such traditional dualisms as objective vs subjective, appearance vs reality, realism vs anti-realism, and so on. Indeed, Rorty argued that even entering into these debates implicitly accepts the legitimacy of such dualisms. The more useful position is to reject them as representationalist. Rorty took his antirepresentationalism from a wide variety of sources, some of which are listed here alphabetically: Darwin, Davidson, Dewey to be sure; Heidigger, Quine, Wilfrid Sellars, and even the later work of Wittgenstein. For example, in his early work such as the *Tractatus*, Wittgenstein (1922/1974) sought to develop a logically consistent "picture theory" of language. This work was central in the rise of logical positivism during the 1920s. In his later work, Wittgenstein (1953) virtually repudiated the earlier work, writing instead of language games, wherein language was a tool for speakers instead of a mirror. For the later Wittgenstein, speakers talk about the world and its constituents as part of getting along in life. Meaning consists in use, rather than in correspondence between a word and some entity from a domain beyond human affairs, where the entity was inferred to possess some essential quality according to some metaphysical doctrine. Whereas the early Wittgenstein was in the essentialist tradition, the later Wittgenstein was in the pragmatic tradition.

With respect to the human condition, in Rorty's hands pragmatism challenges a host of traditional philosophical assumptions regarding the relation between mind and body. As Rorty (1991) put it,

> [T]here is no harm in continuing to speak of a distinct entity called "the self" which consists of the mental states of the human being: her beliefs, desires, moods, etc. The important thing is to think of the collection of those things as *being* the self rather than as something which the self *has*. This latter notion is a leftover of the traditional Western temptation to model thinking on vision, and to postulate an "inner eye" which inspects inner states…. The important thing is to avoid taking common speech as committing one to the view that there is, after all, such a thing as the "True Self," the inner core of one's being which remains what it is independent of changes in one's beliefs and desires. There is no more a center to the self than there is to the brain. (p. 123, italics in original)

Here, Rorty argues against traditional assumptions that a human intention, belief, or desire is one thing but human beings are entirely different, such that they can be known only in some privileged way. Rorty sees these traditional assumptions as deceptive and not particularly helpful ways of conceiving of how humans actually interact with the world or learn more about themselves.

At the heart of pragmatism is a rejection of the traditional view of verbal behavior, according to which words are conceived as symbols that gain their meaning by referring or corresponding to something else. The something else is generally some object that is spoken about. The object is defined in terms of some essential metaphysical quality it is inferred to possess, and this quality identifies the object as belonging to some particular metaphysical category. Note here the assumptions underlying the traditional view: reality consists of the collection of such objects; humans come to know these objects and their essential qualities through their mental processes; human language represents this reality; Truth is a matter of the correspondence between words and the objects in reality; epistemology is a matter of justifying a story about how the mental states and processes in question yield statements about reality that can be agreed upon and thereby validated as facts; and so on. Pragmatism takes such assumptions to be simply a legacy of longstanding social–cultural traditions that are of little or no value in providing answers to important questions about (a) how humans actually interact with and adapt to the world and (b) how philosophy can contribute to an understanding of those interactions.

The Relation Between Pragmatism and Early Schools of Psychology

As noted in the present introduction, pragmatism was especially influential in the development of American functionalism in the last quarter of the nineteenth century, where functionalism may be distinguished from structuralism. Structuralism was an "Ivory Tower" approach to psychology, concerned with the supposed contents of consciousness — sensations, images, and feelings. For example, structuralists thought that if participants had the correct amount of training — perhaps as many as 10,000 training trials were necessary, the participants would then be able to introspectively discern 42,415 different sensations. However, nonhumans were not eligible to serve as subjects because whether they even had conscious minds was not clear, let alone how they could introspectively comment on the contents of those minds. Similarly, children were not eligible to serve as participants because although they did have minds, whether their introspective comments could be trusted was not clear. Applications in the world outside the research laboratory were not of concern. Evolution was not directly relevant. A representative research question in structuralism was: What is the texture of an individual's sensation of green? Structuralists were not bothered in the least that the pragmatic dimensions of such questions were doubtful.

In contrast, functionalism was more directly concerned with the adaptation of behaving organisms, especially humans, to their social and material environments. In keeping with its concern with adaptation, functionalism was infused with developmental, evolutionary thinking. Its analytic and explanatory concepts were dynamic and functional. Highlighting an area that might now be called service delivery or applied psychology, James (1892) emphasized that "All natural sciences aim at practical prediction and control and in none of them is this more the case than psychology to-day" (p. 148). He went on to argue that what "every educator, every asylum superintendent, asks of psychology is practical rules" that will help these professionals to improve the ideas, dispositions, and conduct of people in their charge (p. 148). Arguments about the philosophic grounds of mental phenomena are not especially useful to professionals in such endeavors.

A further example of early pragmatic, functionalist thinking is Dewey's (1896) well-known article on the concept of the reflex arc in psychology. As did other functionalists, Dewey suggested that the important goal in psychology was to explain adaptation. To explain adaptation, he argued psychologists needed a holistic account that integrated both physiological and mental processes on the part of the behaving organism. Such an account was obviously not restricted to the physiology of reflex processes. Moreover, the value of knowledge claims in psychology lay in their ability to engender prediction and control of behavioral events in the world of human affairs outside the laboratory, as James (1892) had earlier argued. All this is consistent with pragmatism as it may be broadly understood. Indeed, many functionalists were staunch pragmatists.

The Relation Between Pragmatism and Watson's Classical S–R Behaviorism

The beginning of behaviorism as an independent movement is sometimes traced to the work of John B. Watson (e.g., 1913, 1925). To be sure, Watson was reasonably familiar with the philosophy of his time, and Dewey was one of Watson's mentors during Watson's graduate school days at Chicago in the early 1900s. Watson also sought to reinterpret some traditional topics in psychology in terms of interactions with the environment, such as by rendering thinking as subvocal speech elicited by various internal and external stimuli (e.g., Watson, 1925, p. 215). On these grounds one might suspect Watson was at least somewhat familiar with a pragmatic orientation.

Important to emphasize, however, is that Watson was primarily interested in pursuing a natural science of behavior, not philosophy. For example, the two opening sentences of Watson's (1913) "behaviorist manifesto" are well known: "Psychology as the behaviorist views it is a purely objective branch of natural science. Its theoretical goal is the prediction and control of behavior" (p. 158). This sense of prediction and control lent itself to practical application not only as William James would have it, but also in the form of behavioral engineering to secure desired outcomes. The engineering goal of producing desired forms of behavior was directly in the spirit of Jacques Loeb, also one of Watson's mentors at Chicago. At the time of his manifesto, Watson had not yet committed to a behavioral technology based on S–R relations. When he did commit to that technology a few years later, the path for social progress based on a science of behavior seemed clear to him, and dominated his work in his remaining years in academic psychology and through the 1920s, when he sought to popularize his behavioristic approach. Nonetheless, although Watson was almost certainly aware of pragmatism, it appears to have influenced him less systematically than it did James and Dewey.

The Relation Between Pragmatism and Neobehaviorism

Smith (1986) has provided a comprehensive analysis of E. C. Tolman's and C. L. Hull's forms of behaviorism during the era of "Grand Learning Theories" in the second quarter of the twentieth century. To be sure, their forms of behaviorism, usually known as mediational neobehaviorism, differ greatly from each other. Nevertheless, pragmatism is linked in subtle ways with the rise of neobehaviorism.

First, neobehaviorists accepted observable stimulus (S) and response (R) variables as the principal data in their systems, which is in keeping with a pragmatism. However, neobehaviorists then postulated unobservable organismic variables (O) to mediate the relation between stimulus and response according to an S–O–R framework. The concept of mediation means that observable external stimuli activate or trigger one or more unobservable intervening structures that are hypothesized to be causally connected in some complex but systematic way to an ensuing observable response. Note that Watson's classical S–R behaviorism had no such mediating organismic variables. Appeal to the mediating variables was thought to be necessary to overcome the limitations of Watson's strict S–R framework in explaining the flexibility of behavior as well as its sequential organization.

Neobehaviorists then cast the mediating variables as theoretical terms, based on developments in the philosophy of science at about the same time. However, those mediating variables needed to be operationally defined. Thus, operationism became a central concern during this era, as a technique to produce agreement about the mediating variables and avoid a return to the vague, ambiguous ways of introspective structuralism. Agreement was achieved through linking the unobservable mediating variable with some specified, observable stimulus or response measure. Here then was one link between pragmatism and neobehaviorism.

Second, neobehaviorists were also concerned during this era with developing a coherent approach to psychological research and theorizing. Laws and theories were primary considerations based on their contributions to the hypothetico–deductive model of explanation that was developing in the philosophy of science under the influence of logical positivism. Predictions deduced from the laws and theories about actual observable behavioral data were paramount. Here again is a link between pragmatism and neobehaviorism, insofar as pragmatism asked for observable implications and consequences of theoretical endeavors in science.

As an aside, despite the neobehaviorist emphasis on observation, the pragmatic dimensions of the proposed laws and theories were sometimes questionable, in light of the uncertain ontology of some of the mediating organismic variables: What variables would practitioners actually manipulate? Hull's "oscillation factor" and "afferent neural interaction" are generally cited as problematic in this regard. In addition, questions were raised about experimental control: What interventions would actually produce desired and predictable outcomes? To be sure, Hull's Institute for Behavioral Relations at Yale was concerned with the practical application of conditioning principles to psychopathology (Dollard and Miller, 1950), but again there was little formal identification with pragmatism as a philosophical movement. The principal concern was with theory development rather than application.

Third and finally, neobehaviorists subscribed to the thesis of instrumentalism. Instrumentalism holds that the function of theories and the terms they contain is to generate testable predictions about events. On this view, the nature of a specific term in the theory, such as a mediating O variable in the S–O–R framework, is less important than its role in generating a verifiable prediction, following from the hypothetico–deductive model of theorizing and explaining. Here, an emphasis on using some observable behavioral measure to assess the predictive success of the mediating variables and theories suggests a third link between pragmatism and neobehaviorism.

In sum, pragmatism probably influenced the methods of classical behaviorism and neobehaviorism more than their content, such as by emphasizing objective methods, observable data, and a decision process rather than introspection. As well, the neobehaviorist reliance on operationism seems consistent with pragmatism. However, the story is more complex than it first appears, and more now needs to be said about the relation between modern pragmatism and behaviorism, especially in regard to Skinner's behavior analysis.

The Relation Between Modern Pragmatism and Behavior Analysis

The relation between modern pragmatism, with Richard Rorty as the representative, and Skinner's behavior analysis is of particular interest in the present review. Rorty speaks critically of Skinner in several instances, for example, when he disparages Skinner's work as mere methodological behaviorism (Rorty, 1979, p. 213), pointless fantasy (Rorty, 1991, p. 33), and positivistic reductionism (Rorty, 1991, pp. 110, 135). Regrettably, Rorty's charges are well wide of the mark.

For instance, in an informative review of Rorty (1991), Leigland (1999) addressed some of Rorty's charges against Skinner. Leigland pointed out that Skinner has consistently embraced a pragmatic interpretation of truth. One example is when Skinner (1974) said that "Scientific knowledge is verbal behavior.... It is a corpus of rules for effective action.... [A] proposition is true to the extent that with its help the listener responds effectively to the situation it describes" (pp. 241–242).

This passage indicates that Skinner did not subscribe to the traditional conception of knowledge as some representation of reality, but rather to a pragmatic conception in terms of how humans interact effectively with the world. Scientists may well collect facts as they go about their business, but those facts are construed as aids to effective action, not unassailable representations of a reality that exists beyond human experience and behavior.

In addition, Rorty's charges that Skinner subscribed to Baconian or Machian positivism simply suggest Skinner's preference to view scientific knowledge as aiding in adaptation. To be sure, Skinner sympathized with Bacon's (1623/1937) wishes that involved "shaping nature as on an anvil" (p. 413) and achieving outcomes that benefit humans through direct, practical action. Mach (1886/1959) put it similarly: "The ways even of science still lead to the mouth" (p. 23). For Skinner (1969), "The point of science … is to analyze the contingencies of reinforcement found in nature and to formulate rules or laws which make it unnecessary to be exposed to them in order to behave appropriately" (p. 166). Skinner's statement here is surely in the pragmatic tradition.

The key to understanding behavior analytic pragmatism lies in a behavioral view of verbal behavior. For Skinner (e.g., 1957), verbal behavior is operant behavior that develops through its effects on other persons. Once a suitable verbal repertoire is acquired, speakers' verbal behavior can influence the speakers themselves, just as it influences other persons. Whether it actually does is an empirical question and a function of further relations during the lifetime of a speaker. Words are rather arbitrary patterns of "sounds and marks" (to use Rorty's felicitous phrase) that arise according to conventional practices of a social group. Words have meaning through their participation in contingencies pertaining to that group. However, meaning for the speaker may be usefully distinguished from meaning for the listener. Meaning for the speaker is a matter of the contingencies that govern the emission of the word. Meaning for the listener is a matter of the contingencies according to which the word functions as a source of discriminative stimulation for the listener, recognizing again that speakers can sometimes be their own listeners. On this view, dictionaries don't give meanings of words. Rather, they give other words that mean the same thing (e.g., Skinner, 1957, p. 9). Grammar and syntax are special, higher-order features of verbal behavior. One such feature is agreement in case, tense, and number between or among individual responses. Another feature is the structural arrangement or sequence of individual responses. The higher-order features develop when listeners encourage speakers to take into account various aspects of the situation about which they speak, such as which of several conceivable actors is or are engaging in the action in question or the source or strength of the speaker's response, because listeners find those aspects useful to know. These features contribute to the discriminative value of the verbal behavior.

Nowhere in this treatment is there an endorsement of a reference or representational theory of language. In fact, Skinner (1957, pp. 114–129) explicitly rejected a reference theory. The term language simply identifies the set of conventional practices that prevail in a verbal community. To speak of linguistic rules is simply to speak of descriptions of those conventional practices. Nowhere is there an endorsement that words refer to or represent a reality of essential Platonic, Cartesian, or Kantian qualities in another dimensional system. Much of this is consistent with Rorty and the spirit of modern pragmatism.

When Dewey (1926) spoke of "the goods which inhere in the naturally generated functions of experience" (p. 407) and Rorty (1991) spoke of "acquiring habits of action for coping with reality" (p. 1), their words are consistent with Skinner's emphasizing that individuals interact with their world through their repertoire of operant behavior. The operant repertoire includes both nonverbal and verbal components. These components develop through the reinforcing outcomes of an individual's actions. Thus, for behavior analysts an understanding of how humans adapt to the world is a matter of understanding the function of their operant behavior. This concern is science not metaphysics, much as pragmatists from Dewey to Rorty argued. Again, behavior analysis may be seen as compatible with pragmatism.

Finally, both pragmatism and behavior analysis strive to avoid ontological commitments, albeit in their own ways. For example, Rorty repeatedly argues against the epistemic, quasi-instrumentalist claim that when a scientist's verbal behavior leads to prediction and control, the verbal behavior in question should be assumed to accurately represent some metaphysical reality. For Rorty, verbal behavior — whether that of scientist or poet — just isn't the type of phenomenon that represents anything, accurately or otherwise. Rather, verbal behavior is simply an instance of an organism's interacting with its world. The relation between the world and language is causal, not representational. As Rorty (1991) put it when advocating his antirepresentationalism,

> The anitrepresentationalist is quite willing to grant that our language, like our bodies, has been shaped by the environment we live in. Indeed, he or she insists on this point — the point that our minds or our language could not (as the representationalistic skeptic fears) be "out of touch with the reality" any more than our bodies could. What he or she denies is that it is explanatorily useful to pick and choose among the contents of our minds or our language and say that this or that item "corresponds to" or "represents" the environment in a way that some other item does not. (p. 5)

Skinner (1969) engaged questions of ontology in a similar fashion when he said that "The basic issue is not the nature of the stuff of which the world is made or whether it is made of one stuff or two but rather the dimensions of the things studied by psychology and the methods relevant to them.... The objection is not that these things are mental but that they offer no real explanation and stand in the way of a more effective analysis" (pp. 221–222). Suffice it to note that for both pragmatism and behavior analysis, matters of ontology develop from uncritical assumptions about verbal behavior. These matters may be set aside in favor of analyzing the events, variables, and relations responsible for both (a) a given instance of behavior, and (b) the verbal behavior about that instance of behavior, so the relevant concerns in the domain of human affairs may be more effectively engaged.

Pragmatism, Behavior Analysis, and Cognitive Science

Behavior analysts argue that mentalism, for example as evidenced in cognitive science, is the dominant explanatory orientation in contemporary psychology. Further, behavior analysts are opposed to mentalism. Worth noting, of course, is that mentalists are just as opposed to behavior analysis, as virtually any text in cognitive psychology will reveal. Thus, a more direct comparison of how behavior analysis and mentalism stand with respect to pragmatism is useful at this point. To lay the groundwork for the comparison, some characteristics of the behavior analytic view of science are examined first, followed by characteristics of the mentalist view.

For behavior analysts, science is in large measure the operant behavior of scientists, along with (a) the artifacts associated with the origin and the execution of their behavior and (b) the artifacts produced by their behavior. The operant behavior of scientists can be nonverbal or verbal. Nonverbal scientific behavior involves interactions with nature that include such material artifacts as test tubes, scales, microscopes, spectrum analyzers, gas chromatography devices, and so on. In turn, this behavior may yield new and improved material artifacts as well as techniques for employing them. Verbal scientific behavior involves such verbal artifacts or products as theories and explanations. In some instances, these verbal artifacts are formulated prior to and guide the investigation of some subject matter. In other instances, new and improved verbal artifacts arise during or after such investigations.

The reinforcers for engaging in scientific behavior fall on a continuum. At one end is the prediction and control of natural events. At this end is a concern with outcomes that have relatively immediate, practical benefits for humans. As reviewed earlier, Bacon and Mach were concerned with such outcomes. At the other end is "the discovery of uniformities, the ordering of confusing data, the resolution of puzzlement" (Skinner, 1979, p. 282). At this end is a higher-order concern with the more abstract products of science, such as theories and explanations that transcend particular instances of prediction and control. Sometimes these products are identified as "knowledge for its own sake," where that phrase signifies verbal products that are derived from more particular endeavors and whose generality has increased as a science progresses.

Primitive science presumably began with relatively primitive technologies that targeted particular outcomes: making clay pots to store food, making garments to keep warm, making houses to protect against the elements, making wheeled carts and wagons to transport people or goods, domesticating and selectively breeding animals to provide for the needs of the clan, selectively breeding plants to provide a predictable food supply, making hammers and knives to aid construction, making swords and spears and shields to defend against rivals or attack them. Cultures then developed rules for implementing these technologies. The rules became formalized as sources of discriminative control — typically verbal, which in turn allowed the knowledge of how to produce desired outcomes to be transmitted to future generations. As the rules became more abstract over time, the degrees of freedom for their application increased, and the rules then became useful across more and more situations.

The important point here is that theories and explanations may be understood as accounts built on a foundation of functional relations. However, the mere accumulation of results is no more valuable as a theory than a heap of stones is valuable as a house (e.g., Poincaré). Data become useful when organized and extended so as to suggest a coherent way to deal with a subject matter. To be sure, some degree of "speculation" may even be involved in the process, as Skinner (1974) suggested: "Speculation is necessary, in fact, to devise methods which will bring a subject matter under better control" (p. 17). The speculation, which is related to what Skinner called "interpretation," takes the form of applying (a) known principles derived from situations in which controlled experimentation was carried out to (b) further situations in which controlled experimentation is not feasible (Moore, 2008a, p. 306 ff.). Lyell's (1830–1833) uniformitarianism is an early example, and much modern work continues this tradition, for example, in (a) plate tectonics, where findings from research on the behavior of substances subjected to high pressures and temperatures are invoked to explain the movement of land masses on the surface of the earth and earthquakes; and (b) evolution, where findings from research on molecular processes in genetics are invoked to explain the origin of species.

As suggested above, behavior analysts call attention to the contingencies that control scientific behavior. Presumably, effective scientific behavior is controlled to a great extent by contingencies arising from operations and contacts with data, rather than from social and cultural contingencies that are more a matter of conforming to statements of revelation and authority. These social and cultural contingencies are linked to a supposed domain beyond the natural world and evidence minimal regard to the outcome of interactions with the natural world, which is a great problem. More is said later in this review about why this matter is critical.

In contrast, mentalism pertains to a particular way of pursuing the causal explanation of behavior. According to mentalism, an individual's intrinsic psychological make-up is taken for granted to include such nonbehavioral phenomena as mental acts, states, mechanisms, processes, entities, structures, faculties, and cognitions.

These states, processes, and structures belong to a domain that differs from the behavioral domain. The mental phenomena are unobservable and inferred to underlie observable behavior. They cannot be characterized in the same terms, and do not function according to the same principles as observable events, variables, and relations in the environment. Some representative terms for this nonbehavioral domain are mental, cognitive, and subjective — in short, the domain of "mind." Traditional mind–body psychophysical (i.e., substance) dualism is an example of mentalism, but not the only one.

These mental phenomena are not acquired or influenced by environmental events during an organism's lifetime (i.e., through experience) in any significant way. Rather, the phenomena are postulated to be evolutionary, innate, or maturational. Physiological measures are said to provide neural correlates and are evidence of the underlying mental phenomena, but do not define them. Rather, the phenomena are defined in terms of their functional characteristics, such as their capacities, contents, processing times, and so forth, rather than their physical realization or observable expression. Observable behavior is important for mentalism insofar as it provides objective evidence to support explanatory inferences about the causal properties of the phenomena, rather than because observable behavior is a subject matter in its own right, as in behaviorism. The phenomena afford competence, which makes the observed behavior possible in whatever situation the organism finds itself.

Importantly, according to the mentalism of cognitive science, researchers and theorists should explain behavior in terms of the functional properties and architecture of the underlying mental mechanisms, structures, states, and processes. Mentalists sometimes argue that their point of view follows from the history of science. According to mentalists, science progresses by inferring unobservable yet theoretically rich analytical and explanatory concepts, rather than by restricting analyses and explanations to observable events, variables, and relations. Frequently cited examples are atoms, electrons, cell theory, germ theory of disease, and receptor sites. None of these explanatory concepts were directly observed at the time they were first inferred. The mentalist argument is that they all illustrate why science should not be restricted to observable events, variables, and relations. The primary concern in science should be epistemological: to construct and justify a story of how the underlying structures yield competence. Prediction and, if necessary, control are held to be at best secondary, technological concerns about performance. Mentalists argue they are simply doing the same as all genuine sciences by inferring these underlying, unobservable phenomena. In turn, these inferences lead to the appropriate causal understanding at a theoretical level.

Mentalists typically contrast their position with any form of behaviorism, including behavior analysis. According to mentalists, behavioral statements simply describe observable environmental events, variables, and S–R relations in the manner of Watson (1925). Behavioral statements don't explain how behavior can come about. In addition, behavior is more flexible than is expected on the basis of observable S–R relations, and its sequential organization differs from expectations based on observable S–R relations. Therefore, mentalists hold that purported explanations of behavior in terms of how that behavior is related to features of the environment are incomplete at best and defective at worst because they don't specify the underlying, unobservable causal structures responsible for the performance. Therefore, behaviorism can't possibly be regarded as generating genuinely theoretical, explanatory knowledge, and can't possibly be regarded as scientific in any meaningful sense of the word.

Finally, mental explanations may even gravitate toward the S–O–R neobehaviorist framework that is prominent in the history of behaviorism, where the O stands for organismic variables that are inferred to mediate the relation between stimulus (S) and response (R) and to provide the desired richness and flexibility. For example, Neisser (1967) argued that "Whatever we know of reality has been *mediated*, not only by the organs of sense but by complex systems which interpret and reinterpret sensory information" (p. 3, italics added). However, for mentalism these mediating O variables are explicitly conceived as nonbehavioral and unobservable. Accordingly, mentalism argues that its explanatory scope supersedes that of any form of behaviorism (Moore, 2013a, 2013b).

In reply to mentalist concerns, behavior analysts agree that trying to explain all behavior in terms of observable S–R relations is surely inadequate, just as mentalists charge. However, behavior analysts do not seek to explain all behavior in terms of observable S–R relations. Although some behavior is indeed attributable to observable S–R relations, by far the most important and relevant form of human behavior is operant behavior, and a contingency is responsible for operant behavior, not observable S–R relations. An important consideration is whether mentalist criticisms of behavior analysis recognize the difference. Many — possibly most — do not.

On a deeper level, behavior analysts argue that mentalism adheres to correspondence theories of truth by virtue of the epistemological concern with justifying the metaphysical story about the supposedly underlying mental states and processes that mediate psychological functioning. Of less concern is any pragmatic interest in the prediction and control of actual behavioral events. Accordingly, mentalism is inherently concerned with formulating an account in terms of the essential properties of those states and processes. Miller's (1956) "Magical number seven plus or minus two" is a case in point. This concern is antithetical to pragmatism. Again, a concern with structure per se is not what makes mentalism unpragmatic. Talk about these structures, their operating characteristics, and their capacities that informs effective action is surely pragmatic. What makes mentalism unpragmatic is its concerns with (a) justifying the metaphysical talk about those structures and their operating characteristics, such that the structures can be said to correspond with reality; and (b) dismissing as unscientific any interest in how the talk contributes to effective action, such as through prediction and control. Justifying the concern in mentalism by saying it is conceptual or theoretical rather than technological is merely begging the metaphysical question.

To be sure, many psychological theories and explanations do contain terms and concepts that at first glance appear to be mental. Nonetheless, for behavior analysts, some of those terms and concepts are free from concern because they do not appeal to causal entities from a nonbehavioral domain. Rather, they take into account behavioral events, variables, and relations worthy of study in their own right. However, many other terms are unselfconsciously mental. Therein lie behavior analytic concerns. These judgments need to be made on a case-by-case basis. Behavior analysts refer to these judgments as involving the operational analysis of the verbal behavior in question (e.g., Skinner, 1945).

Thus, behavior analysts are concerned about mentalism on a pragmatic basis. More specifically, behavior analysts argue that the supposed properties to which mentalists appeal tend to obscure and indeed actively impede the search for the relevant relations between behavior and environment, allay curiosity by inducing the acceptance of fanciful "explanatory fictions" as causes, misrepresent the facts to be accounted for, and give false assurances about the state of our knowledge. Consequently, behavior analysts argue that mentalism tends to interfere with effective prediction, control, and explanation of behavior, despite mentalist claims to the contrary. Moreover, the mentalist conception of these unobservable phenomena implies that they and the behavior they cause arise and function independently of environmental circumstances, and nothing can be done to promote beneficial forms of behavior or to replace troublesome forms. Such a view is surely unpragmatic.

Pragmatism and Behavior Analytic Theories and Explanations

Worth emphasizing at this point is how strongly behavior analysts are committed to theories and explanations, rather than simply descriptions of observed events, as critics so often argue. For example, Skinner (1972) argued that "[T]he cataloguing of functional relationships is not enough…. Behavior can only be satisfactorily understood by going beyond the facts themselves. What is needed is a theory of behavior…. [E]xperimental psychology is properly and inevitably committed to the construction of a theory of behavior. A theory is essential to the scientific understanding of behavior as a subject matter" (pp. 301–302). The important issue for behavior analysts is the source of control in the contingencies governing the verbal behavior of theorizing and explaining. The source of control is important because it determines the discriminative contribution of the verbal behavior to effective action.

For Skinner, theories with extensive sources of control in neural, mental, and conceptual domains were pragmatically questionable. As Skinner (1950) put it in a well-known article, such theories offer "explanations which appeal to events taking place somewhere else, at some other level of observation, described in different terms, and measured, if at all, in different dimensions" (p. 215). In their place, Skinner advocated theories having other forms: "Beyond the collection of uniform relationships lies the need for a formal representation [sic] of the data reduced to a minimal number of terms…. But such a construction will not refer to another dimensional system and will not, therefore, fall within our present definition" (p. 215). Such theories are based on organizations of facts. They have a broad generality that transcends particular facts and allows them to contribute to effective action, such as through prediction and control.

Skinner actually wrote a great deal about his recommendations for a pragmatically based scientific epistemology. In one instance, Skinner (1964) put it as follows: "When I said 'explanation' I simply meant the causal account" (p. 102). Thus, an explanation entails the specification of a functional relation between behavior and manipulable or controllable variables. In another instance, Skinner (1953) argued that "Science is not concerned with contemplation. When we have discovered the laws which govern a part of the world about us, we are then ready to deal effectively with that part of the world. By predicting the occurrence of an event we are able to prepare for it. By arranging conditions in ways specified by the laws of a system, we not only predict, we control: we 'cause' an event to occur or to assume certain characteristics" (pp. 13–14). Again, such statements indicate that for behavior analysts, the principal concern of scientific epistemology is the extent to which a scientific statement promotes practical, effective action. As before, an assertion that the function of a theory is to generate some theoretical understanding divorced from the possibility of practical action, as in traditional theories concerned with competence and expressed in neural, mental, or conceptual dimensions, is pragmatically questionable.

Pragmatism, Behavior Analysis, and the Sources of Control over Scientific Verbal Behavior

As argued in the present review, behavior analysts emphasize sources of control over scientific verbal behavior. One source of control over such verbal behavior is surely the operations that a scientist performs and the outcomes of those operations. In Skinner's (1957) terminology, this source of control involves tact relations and extensions of those relations.

However, some verbal behavior said to be scientific is linked at least partly to social–cultural sources that are cherished for extraneous and irrelevant reasons. One such social source of control lies in Western culture, which is largely mentalistic if not palpably dualistic. After all, scientists are typically socialized in a mentalistic society. They conform in certain respects to the prevailing social–cultural traditions and institutions, which are surely mentalistic. A popular name for this source is "folk psychology." For example, religious practices and institutions routinely appeal to the Soul, commonly secularized as the Mind. The individual is seen as an originator of action or as an agent with "free will." These views are prevalent in books, articles, as well as statements by learned figures in the culture. Students get high grades in their schoolwork for reciting these views in their classwork. The social reinforcement associated with this source of control is enormous. Starting with Watson (1913), behaviorists have been concerned with it. For behavior analysts, the task is to set such mentalistic social influences aside in favor of effective action at the naturalistic level. To use the field of medicine as an example, surely therapeutic practices would have advanced more rapidly if physicians had critically examined the outcomes of treating illnesses with leeches, cuppings, and emetics instead of conforming to convention and remaining with these socially approved techniques.

A second source of social control lies in the mischievous linguistic practice of converting adjectives and adverbs to nouns, then assuming the noun stands for some entity in some other domain — neural, mental, conceptual — of which behavior is merely a symptom or an expression. Again, Skinner (1974) commented on this practice as follows: "When a person has been subjected to mildly punishing consequences in walking on a slippery surface, he may walk in a manner we describe as cautious. It is then easy to say that he walks with caution or that he shows caution. There is no harm in this until we begin to say that the walks carefully because of his caution" (p. 166). Here, caution is invoked as a mediating property in a mental domain, when it has simply been converted from an adjective into a noun, and then cited as a cause. Common terms for this process are reification and hypostatization.

A third source of social control is in the mischievous use of metaphors. Perhaps metaphors are useful in some sense. After all, some applications on computer networks are meaningfully said to function according to a client–host arrangement. Data are said to be stored on the "cloud." However, metaphors can also mislead. For example, memory is commonly accepted to be a process of storage and retrieval. However, if anything is stored, it is a changed organism, not a copy of an experience that is later retrieved. Part of the problem is the general acceptance in our culture of a reference (or as a pragmatist might put it, representational) theory of meaning in which various entities are invented to fill linguistic niches.

In Rorty's (1979, 1991) arguments in favor of construing verbal behavior in a particular way, he speaks approvingly of a number of authors in the post-modern tradition, including Nietzsche, Derrida, Foucault, and even Thomas Kuhn. When it comes to science, Rorty's writing often offers post-modern objections to the assumption that through a philosophically based scientific method, scientific verbal behavior may be taken as correctly identifying fundamental elements of reality, and that prediction and control justify the knowledge claims that arise from such endeavors. Rorty was especially concerned about scientism, which may be described as a "representationalist view of science as a privileged window on Reality. According to this view, science employs a special Method that provides information on the true nature of things" (Leigland, 1999, p. 495). Rorty further argued that our culture does well not to embrace a post-Kantian epistemology "taken for granted by most Western philosophers during the last two centuries" (Rorty, 1991, p. 118), according to which the Self consists of various layers that mediate our perceptions of and statements about physical reality. Rorty favored instead the position of *nonreductive materialism*, according to which the fundamental distinction is between the individual and the rest of the universe, recognizing that various neural and physiological processes are carried out within the individual's body. Leigland further commented that "Such a straightforward view of biological–behavioral–environmental interaction is quite compatible with … behavior-analytic science" (p. 492). On a behavior analytic view, science is a form of human interaction with and adaptation to the world, rather than a picture of a reality that philosophy tries to justify as lying somewhere else, in some other domain than that of human affairs.

A related question at this point is whether pragmatism and by extension behavior analysis should be considered as equivalent to an instrumentalist view of the role of verbal behavior in science. The position taken here is that pragmatism, behavior analysis, and instrumentalism are not equivalent. To be sure, all three positions can be said to use observations to determine whether predictions are accurate and action is effective. Indeed, many contemporary forms of science do the same. At issue is the source of control over a scientific statement, such that predictions are accurate and action is effective. Behavior analytic pragmatism asks whether the source of control can be refined, for example, by minimizing or even eliminating any control related to mentalistic social influences, such that predictions can be even more accurate and actions even more effective. Instrumentalism stops short of asking such questions. As a result, the present argument is that behavior analytic pragmatism cannot be equated with instrumentalism.

This matter is relevant because of the common instrumentalist orientation of "as if" (e.g., Vaihinger, 1924), for example, in statements that humans act "as if" their minds were computers with such and such capacities, or humans act "as if" their behavior were a function of some personality factor with such and such properties. This orientation invites mental theories and explanations, owing to the social prevalence of mentalism, and the liabilities of mentalism have already been reviewed. Again, if some theory or explanation proves useful, the basis for its utility needs to be examined, and possibilities for enhancement need to be explored. Because previous behavioral rather than so-called mental theories and explanations have proved effective, every reason exists to believe the same situation prevails again, and the reasons for the effectiveness of any theory or explanation are behavioral rather than mental. If so, further examination will reveal what those behavioral reasons are. Moreover, the social reasons for appealing to mental factors will become apparent. As noted, those social reasons include conformity to mentalistic social–cultural traditions, reification brought about by linguistic practices, and embrace of mischievous metaphors. The basis for effective action derives from understanding the variables and relations that participate in the event being described, rather than the aforementioned social influences. Moreover, the possibility exists that by refining the control from naturalistic factors, the action can be even more effective and ultimately of greater benefit in human affairs.

The position espoused above does not mean all verbal behavior that appears to be mentalistic should be automatically rejected. Indeed, Skinner (1964) commented quite the opposite when he said that "No entity or process which has any useful explanatory force is to be rejected on the ground that it is subjective or mental. The data which have made it important must, however, be studied and formulated in effective ways" (p. 96). The second sentence in Skinner's statement is central. Again, at issue is the source of control over the verbal behavior in question. All this is critical in behavior analysis, although not developed to the same degree in Rorty and pragmatism.

The important consideration is that verbal processes may be understood as behavioral rather than logical, referential processes, and as language games of speakers as they get along in their verbal and nonverbal worlds. To say that words have meanings and that some meanings contribute better to adaptation than do others is clearly reasonable. However, to say that some meanings are more legitimate than others because of that to which they refer concedes rather than rejects the premise of representation. In this regard, concepts are matters of discriminative control: generalization within the class of stimuli that sets the occasion for a term and discrimination between classes. Abstraction is similarly a matter of discriminative control, where the source of control lies in, say, one property rather than the totality of properties of the object, situation, or event with which the speaker is currently in contact. In all cases, an assumption that the concept corresponds to some unobservable, underlying category of reality determined by some metaphysical essence is unwarranted. Importantly, a behavioral understanding of verbal behavior puts everything in good order. Most especially, an operational analysis of the concept as an instance of verbal behavior reveals any extraneous sources of control. When revealed, these extraneous sources can be minimized, leaving refined and ultimately more effective verbal behavior to help individuals adapt to their world.

A final point concerns the source of control over terms commonly said to be mental. Five cases may be examined. In the first, terms said to be mental have a source of control that is actually in private behavioral events (e.g., Moore, 2008a, chapter 10). Private behavioral events are those events to which no one beyond the behaving individual has access, for example, because the events are within the behaving individual's skin. One subcategory of these terms pertains to the development and maintenance of verbal reports about internal sensations and feelings, as in statements about personal experiences involving pain, pleasure, or anxiety. Another subcategory pertains to the development and influence of covert operants, as in thinking or problem solving. Skinner (1964) commented on the importance of private behavioral events in the following passage: "*An adequate science of behavior must consider events taking place within the skin of the organism, not as physiological mediators of behavior, but as part of behavior itself.* It can deal with these events without assuming that they have any special nature or must be known in any special way. The skin is not that important as a boundary. Private and public events have the same kinds of physical dimensions" (p. 84, italics added). For behavior analysts, private events are behavioral and owe their origin and influence to environmental circumstances that individuals experience during their lifetimes. Once again, the importance of remaining in the domain of human affairs when engaging certain psychological or philosophical topics, rather than appealing to a mental domain, is evident. More is said about private behavioral events in a following section of this review.

In the second case, terms said to be mental have a source of control that is actually in physiology. An organism's physiology necessarily participates in any behavioral event. At issue is how the contribution of its physiology is to be incorporated in an explanation of its behavior. For behavior analysts, these terms pertain to physiological processes in the two gaps in a purely behavioral account. One gap is within a behavioral event, such as the time from an organism's contact with a stimulus until its response. A representative term here is recruitment. A second gap is between behavioral events, such as between an organism's experiences on one occasion and the effects of those experience on a later occasion. A representative term here is consolidation. After all, an organism's body is surely composed of physiological structures, and these structures surely do have operating characteristics and capacities that can be studied and known about.

In the third case, terms said to be mental have a source of control that is actually in behavioral dispositions. These terms pertain to the probability of a particular form of behavior in particular circumstances. Dispositional interpretations are the prominent substitute for all mental terms in various forms of philosophical behaviorism, from Carnap's logical empiricism to Ryle's conceptual analysis. For behavior analysts, some but not all mental terms may indeed be understood as dispositional. Representative terms here are those of propositional attitudes and the intentional idiom, such as belief, desire, and intention.

In the fourth case, terms said to be mental have a source of control that is actually in stimulus control relations. These terms pertain to the influence of antecedent environmental circumstances on behavior. Representative terms here are from the vocabulary of stimulus control: attention, discrimination, generalization. For example, the term discrimination identifies the fact of differential responding as a function of differential antecedent stimuli, not a mediating mental process of discriminat*ing* that causes the differential responding.

The terms in the preceding four cases have sources of control in tact relations or extensions of those relations. In contrast to these terms are those with sources of control that may be traced more to irrelevant and extraneous social factors, such as conforming to authority or uncritically accepting social conventions and cultural practices. These terms are evident in folk psychology, appeal to inappropriate metaphors, and follow from various linguistic practices, such as when adjectives and adverbs are converted into nouns and then the nouns are assumed to correspond to causal acts, states, etc., that exist in a mental domain. These terms are reified explanatory fictions said to belong to a domain that differs from the behavioral domain. Examples abound in the lexicon of traditional psychology. Memory is conceived as a mental storage and retrieval process, rather than as the reinstatement of a response as a function of the passage of time. Thinking is conceived as an autonomous mental process whose neural correlates are located in the prefrontal cortex, rather than as a behavioral process with either public or private dimensions that contributes to discriminative control. And so it goes. Rather than a function of the tact relation or its extensions, these terms are simply socially induced by language patterns according to what Skinner (1957) identified as intraverbal or echoic control. Terms with these sources of social control are troublesome because they ultimately lead to the counterproductive practices of mentalism and methodological behaviorism (Moore, 2011/2012, 2013a, 2013b).

Behavior Analysis and Private Behavioral Events

One of the topics that attracts a great deal of attention in traditional approaches to both philosophy and psychology is how to understand processes that go on inside individuals in some sense, such as when individuals talk about aches and pains or think. Whereas overt behavior may be easily seen and measured, processes inside the skin seem to require another mode of inquiry and analysis. As a result, the risk of metaphysical intrusion from extraneous sources looms large. In keeping with pragmatism, behavior analysts neither deny nor ignore events inside the skin, nor do behavior analysts remain strictly at the level of relations between observable stimuli and responses. In addition, behavior analysts do not accept traditional assumptions that unobserved processes must be mental. Rather, behavior analysts regard these processes as just as behavioral as observed processes. Thus, an important feature of behavior analysis involves private behavioral events.

As noted earlier, for behavior analysts, private behavioral events fall into two categories (Moore, 2008a, chapter 10). The first is when individuals talk about their sensations and feelings. The second is when individuals engage in covert behavior, such as thinking or solving a problem. Thus, private behavioral events may be undeniably relevant to an understanding of a given instance of behavior as an act in context, even though these events are inaccessible or unobservable from the vantage point of another. What behavior analysts argue is that such covert events may be explained using the same principles as overt behavior.

When it comes to talk about sensations and feelings, behavior analysts argue that listeners reinforce such talk on the basis of public features of the circumstances in which it occurs. Thus, speakers learn to say the pain they feel is a sharp pain when it is caused by a sharp object, a dull pain when it is caused by a dull object, a burning pain when it is caused by a hot object, and so on. There then can be generalization to other circumstances based on the similarity of the sensation. For example, speakers might learn to say they are experiencing butterflies in their stomachs when they experience a fluttering sensation resembling that of a butterfly on their arm.

When it comes to engaging in covert behavior like thinking, behavior analysts argue that behavior is acquired in overt form. Then, because of experiences in the environment, the behavior recedes to the covert form. It is executed by the same motor system, just reduced in magnitude. One common experience that leads behavior to become covert is punishment. After all, individuals learn to read aloud, but reading aloud is punished in the library. As a result, individuals learn to read covertly. Another factor is that engaging in covert behavior is often faster or more expedient. Skilled mathematicians can often solve a problem covertly — "in their heads" — faster and easier than laboriously writing out computations using paper and pencil. In such cases, the individuals are not doing anything essentially different from when they engage in overt forms of the behavior.

Importantly, the behavior analytic position on private behavioral events is neither mentalism nor methodological behaviorism. The position is not mentalism because (a) the events are in the behavioral domain, not a mental domain; (b) the responses are executed by the same response systems as overt responses, just reduced in magnitude; and (c) the origins and effects of private behavioral events on subsequent behavior are a function of environmental circumstances. Thus, for behavior analysts, private behavioral events are very different from the wide variety of causal mental or cognitive states and processes that traditional psychology posits as necessary for an explanation.

Similarly, the behavior analyst position on private behavioral events is not methodological behaviorism because behavior analysts speak directly about the functional relevance of covert behavior, even though the covert behavior is inaccessible to others. Behavior analysts do not try to gain agreement and make analytic or explanatory talk of phenomena from an unobservable mental domain scientifically respectable by appealing to observable data, as in a traditional operational definition. Thus, thinking is a form of behavior in and of itself. Thinking is not construed as traditional approaches have it as some underlying mental or cognitive process that is expressed in overt behavior and for which that overt behavior is an operational measure. To be sure, until technology improves and a second person can directly access the otherwise private events of a first, a second person infers the private events of the first, but for the first, the private events are no inference.

With respect to the first category of private behavioral event — verbal reports about sensations and feelings, what individuals feel are conditions of their bodies. The conditions felt are causal in the sense that they are discriminative stimuli for verbal reports about them. However, the conditions felt are not usefully regarded as causal for behavior. Rather, a more useful understanding is that environmental events, variables, and relations cause both (a) behavior and (b) the conditions felt. A causal analysis most usefully traces a verbal report about sensations and feelings back to the environmental circumstances to which the verbal report is related.

With respect to the second category of private behavioral event — covert behavior, such behavior is causal in the sense it can contribute to discriminative control over subsequent behavior, for example, through its participation as a link in a sequence of responses extended over time. However, covert behavior does not automatically and necessarily occur in every instance of behavior. When it does occur, it does not necessarily influence subsequent behavior. Rather, covert behavior may contribute to discriminative control through individuals' experiences during their lifetimes. The extent to which it actually does so depends on those experiences. These considerations distinguish the behavior analytic conception of covert behavior from the accounts of mediational S–O–R neobehaviorism.

The concept of private behavioral events offers a comprehensive and pragmatic scientific account of one form of human activity based on thoroughgoing behavioral principles, rather than an account based on the explanatory fictions of mediating mental states and processes. The fundamental issue is whether psychologists can explain in a naturalistic way how humans learn to talk about what they feel or how they learn to think. For example, suppose psychologists simply ignore such matters or deny they are relevant on the grounds they are not publicly observable. For behavior analysts, the problem is that psychologists who do so resort to a form of methodological behaviorism, which is hardly a solution.

Alternatively, suppose psychologists accept internal events as ontologically mental, endow the events with the desired causal status, and then render those events as operationally defined hypothetical constructs for the purposes of science. These moves are commonplace in psychology. However, behavior analysts argue that the moves create problems when it comes to explaining the behavior of (a) subjects in psychological research and (b) psychologists as they do science. First, with respect to explaining the behavior of research subjects, behavior analysts argue that the moves institutionalize causal mental phenomena from beyond the behavioral domain. Rendering the mental phenomena as operationally defined hypothetical constructs only makes the constructs proxies for explanatory fictions of dubious origin and ad hoc properties. In short, the moves fractionate rather than unify accounts of nature.

Second, with respect to explaining the behavior of psychologists as they do science, behavior analysts argue that the moves endorse an epistemological dualism. That is, the moves mean that psychologists take for granted that the explanation they provide of their own scientific behavior should be in mental, nonbehavioral terms, as when psychologists take for granted that appealing to operationally defined hypothetical constructs provides them with the epistemological leverage necessary to explain certain processes in their subjects. The moves promote a demonstrably ineffective and unpragmatic means of seeking prediction and control, despite the mentalists' claims of scientific legitimacy (Moore, 2011/2012, 2013a, 2013b, 2013c).

Pragmatism and the Relation between Psychology and Physiology

The matters considered to this point are relevant to a pragmatic understanding of several further issues in a science of behavior. One is the relation between psychology and physiology. On a traditional view, physiology is held to be the epistemological foundation for theories and explanations of behavior, such that theories and explanations are incomplete unless some underlying physiological mechanism is elucidated. To be sure, knowledge of an organism's physiology can be relevant if prediction and control of the organism's behavior is sought. Also relevant is a physiological technology for intervening in the situation at hand. The technology would involve the knowledge of how to intervene as well as the means to do so.

On a pragmatic view, one issue is whether knowledge of an organism's physiology is necessary for prediction and control of its behavior. The answer here is no — behavior is easily altered without knowing how independent and dependent variables are connected physiologically.

A second issue is whether behavior can be more easily altered, or whether more resources are available for altering behavior, if knowledge of an organism's physiology is available. In principle the answer here is yes.

Consider a child who is being taught to read in a classroom. A child who has learned to read differs physiologically from one who has not yet learned. If those physiological differences are known, then in principle direct interventions and manipulations could produce them. However, a series of questions follows. How likely will the knowledge of the relevant physiology of a human even be known, regardless of the setting in which it is sought, such that prediction and control of reading is possible? Even if it is known, how likely is it that the appropriate apparatus will be connected to the child in the classroom, such that a teacher has access to the apparatus and to the relevant physiological state of the child? How likely is it that the teacher can actually intervene physiologically in the necessary ways?

Notwithstanding the earlier "in principle" answer, practical considerations suggest answers to the questions above skew in the direction of not very likely. An alternative approach is suggested by viewing explanation as well as prediction and control in pragmatic terms. For behavior analysts, an explanation of behavior entails a functional account that leads at least in principle to prediction and control. On this view, interventions based on either physiological or environmental variables can yield prediction and control. The choice of which intervention to employ is pragmatic, based on the resources available at a particular time and place that will bring about the behavior of interest. So formulated, physiological interventions may be seen as complementary to environmental, rather than foundational: the more that is known from one domain to predict and control, the less is needed from the other. The decision of how to intervene and shape nature as on an anvil, as Bacon would have it, is pragmatic: Which type of intervention will work better, easier, faster, and so on, based on the current state of psychological knowledge and the technology that psychologists have available at the time and place they wish to influence behavior (Moore, 2002)? Physiological knowledge is not superior to or more fundamental than behavioral knowledge, nor does physiological knowledge provide the grounds for validating behavioral knowledge.

Thus, the physiology of the behaving organism is clearly a relevant topic in a science of behavior, but in a different way than much of traditional psychology argues. At issue for behavior analysts is the extent to which much of contemporary neuroscience with its unselfconscious cognitive orientation can contribute to predict and control, owing to its inherent mentalism. Skinner was doubtful:

> [C]ognitive constructs give physiologists a misleading account of what they will find inside. (Skinner, 1978, p. 111)

> Cognitive science is premature neurology. (Skinner, 1987, p. 111)

Additional understanding of the relation between psychology and physiology is gained by turning to Aristotle's classic treatment of causation. For behavior analysts the error in much of contemporary neuroscience and cognitive science generally is the conception of an organism's physiology as an efficient cause. More usefully, an organism's physiology may be understood as a material cause. The organism's neural, muscular, and hormonal systems mean that the organism is sensitive to the environmental circumstances that it experiences, such that environmental interventions will have their desired effect. Without that sensitivity, those interventions would be ineffective. Of course, direct physiological interventions may produce an equivalent effect, and if so, are as valuable as interventions based on environmental manipulations. This sense of causation differs from a sense in which those systems are endowed with endogenous power to produce the behavior in question, in a variation of efficient causation, and promotes a pragmatic view of the relation between psychology and physiology.

Pragmatism and Levels of Analysis in Psychology

The matters considered to this point are also relevant to a pragmatic understanding of another controversial issue, that of molar versus molecular levels of analysis (e.g., Moore, 1983). Early in the nineteenth century, E. B. Holt (1915) rejected Watson's molecular thesis that temporally extended forms of behavior consisted of a series of concatenated reflexes, strung together like beads on a string. Holt's molar views were then elaborated by his student Tolman, who proposed mediating states and processes, as in mediational S–O–R neobehaviorism, to explain the organization of behavior. For Skinner (1969), Tolman's proposal was clearly not a solution:

> [Tolman] put the "third" variables inside the organism, where they "intervened" between stimulus and response. There was no reason to do this except to maintain something like the old reflex-arc pattern. His intervening variables quickly assumed the function of mental processes (as they were essentially designed to do), and it is not surprising that they have been warmly taken up by cognitive psychologists. (p. 28)

An alternative is to think of molar versus molecular levels of analysis in terms of the temporal context of independent and dependent variables, where molar implies long term temporal relations and molecular short term temporal relations (Moore, 2008a, 2008b). Thus, the pragmatic question would be: According to which time scale — short, intermediate, long, or even some combination — are prediction and control more usefully achieved? As stated, the pragmatic aspects of the question are clear, and the answers will follow from conducting research. The answers need not be identical for all behavioral processes, whether operant or respondent; for all subjects or participants, whether white rats, white Carneau pigeons, or humans; for all discriminative stimuli, whether lights or tones; for all consequences, whether food, water, and money as a positive reinforcer or shock avoidance as a negative reinforcer; or for all punishers, whether electric shock or loss of money. Attempts to formulate laws as metaphysical representations of reality and then to explain behavioral events by asserting that they are instances of those laws are simply legacies of essentialist thinking outlined earlier (Moore, 2008b). Accordingly, in many instances those attempts fall victim to unpragmatic thinking, with its attendant liabilities. If research does find uniformities across species, variables, and relations, so much the better, but the data are the arbiter.

Summary and Conclusions

In conclusion, pragmatism and behavior analysis have much in common. Of particular concern for both is an account of how verbal behavior contributes to adaptation. A problem arises with traditional assumptions when words are assumed to refer to items of reality, and truth is assumed to be a matter of how faithfully words correspond with reality. Rather, a more useful position is that the analysis of verbal behavior reveals its sources of control. Some valuable sources of control are derived from observations and extensions of those observations.

In contrast, other sources of questionable value are derived from social influences: social–cultural traditions, reification, and inappropriate metaphors. These social sources are cherished for extraneous and irrelevant reasons, and mislead inquiry. Overall, to engage in science is to engage in operant behavior. Such behavior is maintained by a range of outcomes, from (a) prediction and control of events in our lives to (b) making sense of those events by seeing order and identifying the factors that participate in them. Questions about the relation between neuroscience and psychology and levels of analysis for psychological data are inherently pragmatic questions about scientific effectiveness, based on such criteria as technological knowledge and resources available to the scientist, rather than on a metaphysically reductive epistemology. A healthy and informed interaction between pragmatism and behavior analysis benefits both parties.

References

Bacon, F. (1937). De dignitate et augmentis scientarium. In R. F. Jones [Ed. and Trans.], *Essays, Advancement of Learning, New Atlantis, and other pieces* (pp. 377–438). New York: Odyssey. (Original work published 1623)

Catania, A. C., and Harnad, S. (Eds.). (1988). *The selection of behavior: The operant behaviorism of B. F. Skinner: Comments and controversies*. Cambridge: Cambridge University Press.

Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review, 3*, 357–370.

Dewey, J. (1926). *Experience and nature*. Chicago: Open Court.

Dollard, J., and Miller, N. (1950). *Personality and psychotherapy*. New York; Mc Graw–Hill.

Holt, E. B. (1915). *The Freudian wish and its place in ethics*. New York: Holt.

James. W. (1890). *Principles of psychology* (two volumes). New York: Holt.

James, W. (1892). A plea for psychology as a "natural science." *Philosophical Review, 1*, 146–153.

Leigland, S. (1999). Pragmatism, science, and society: A review of Richard Rorty's *Objectivity, Relativism, and Truth: Philosophical Papers, Volume 1*. *Journal of the Experimental Analysis of Behavior, 71*, 483–500.

Lyell, C. (1830–1833). *Principles of geology* (three volumes). London: John Murray.

Mach, E. (1959). *The analysis of sensations*. New York: Dover. (Original work published 1886)

Menand, L. (2002). *The metaphysical club*. New York: Farrar, Strauss, and Giroux.

Miller, G.A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Moore, J. (1983). On molarism and matching. *The Psychological Record, 33*, 313–336

Moore, J. (2002). Some thoughts on the relation between behavioral neuroscience and behavior analysis. *The Psychological Record, 52*, 261–280.

Moore, J. (2008a). *Conceptual foundations of radical behaviorism*. Cornwall–on–Hudson, New York: Sloan.

Moore, J. (2008b). A critical appraisal of contemporary approaches in the quantitative analysis of behavior. *The Psychological Record, 58*, 641–664.

Moore, J. (2010). Philosophy of science, with special consideration given to behaviorism as the philosophy of a science of behavior. *The Psychological Record, 60*, 137–150.

Moore, J. (2011/2012). Methodological behaviorism as a radical behaviorist views it. *Behavior and Philosophy, 39-40*, 145–202.

Moore, J. (2013a). Mentalism as a radical behaviorist views it–Part 1. *Journal of Mind and Behavior, 34*, 133–164.

Moore, J. (2013b). Mentalism as a radical behaviorist views it–Part 2. *Journal of Mind and Behavior, 34*, 205–232.

Moore, J. (2013c). Methodological behaviorism from the standpoint of a radical behaviorist. *The Behavior Analyst, 36(2)*, 197–208.

Moxley, R. (2001). Sources of Skinner's pragmatic selectionism in 1945. *The Behavior Analyst, 24*, 201–212.

Moxley, R. (2001/2002). The selectionist meaning of C. S. Peirce and B. F. Skinner. *The Analysis of Verbal Behavior, 18*, 71–91.
Moxley, R. (2002). Some more similarities between Peirce and Skinner. *The Behavior Analyst, 25*, 201–214.
Neisser, U. (1967). *Cognitive psychology*. New York: Appleton–Century–Crofts.
Rorty, R. (1979). *Philosophy and the mirror of nature*. Princeton, New Jersey: Princeton University Press.
Rorty, R. (1991). *Objectivity, relativism, and truth: Philosophical Papers, Volume 1*. Cambridge: Cambridge University Press.
Schoneberger, T. (2016). Behavioral pragmatism: Making a place for reality and truth. *The Behavior Analyst, 39,* 219–242.
Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review, 52,* 270–277, 291–294.
Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review, 57*, 193–216.
Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton–Century–Crofts.
Skinner, B. F. (1964). Behaviorism at fifty. In T. W. Wann (Ed.), *Behaviorism and phenomenology* (pp. 79–97). Chicago: University of Chicago Press.
Skinner, B. F. (1969). *Contingencies of reinforcement.* New York: Appleton–Century–Crofts.
Skinner, B. F. (1972). *Cumulative record*. New York: Appleton–Century–Crofts.
Skinner, B. F. (1974). *About behaviorism.* New York: Knopf.
Skinner, B. F. (1978). *Reflections on behaviorism and society*. Englewood Cliffs, New Jersey: Prentice–Hall.
Skinner, B. F. (1979). *The shaping of a behaviorist*. New York: Knopf.
Skinner, B. F. (1987). *Upon further reflection*. Englewood Cliffs, New Jersey: Prentice–Hall.
Smith, L. D. (1986). *Behaviorism and logical positivism*. Stanford, California: Stanford University Press.
Vaihinger, H. (1924). *The philosophy of "as if."* London: Routledge. [English translation by C. K. Ogden]
Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*, 158–177.
Watson, J. B. (1925). *Behaviorism*. New York: Norton.
Wittgenstein, L. (1974). *Tractatus logico-philosophicus* [D. F. Pears and B. F. McGuiness, Trans.]. New York: Routledge. (Original edition published in 1922)
Wittgenstein, L. (1953). *Philosophical investigations*. New York: Macmillan.

The Journal of Mind and Behavior
Summer and Autumn 2016, Volume 37, Numbers 3 and 4
Pages 247–270
ISSN 0271–0137

Reciprocity and Reputation: A Review of Direct and Indirect Social Information Gathering

Yvan I. Russell

University of Göttingen and *Middlesex University*

Direct reciprocity, indirect reciprocity, and reputation are important interrelated topics in the evolution of sociality. This non-mathematical review is a summary of each. Direct reciprocity (the positive kind) has a straightforward structure (e.g., "A rewards B, then B rewards A") but the allocation might differ from the process that enabled it (e.g., whether it is true reciprocity or some form of mutualism). Indirect reciprocity (the positive kind) occurs when person (B) is rewarded by a third party (A) after doing a good deed towards somebody else (C) — with the structure "A observes B help C, therefore A helps B." Here too, the allocation differs from the process: if there is underlying cognition, then indirect reciprocity is based on some ability to keep track of the reputations of others (to remember that "B helped C"). Reputation is a kind of social impression based on typicality, derived from three channels of experience (direct encounters, bystander observation, and gossip). Although non-human animals cannot gossip verbally, they can eavesdrop on third parties and learn vicariously. This paper ends with a proposal to investigate the topic of social expertise as a model for understanding how animals understand and utilise observed information within their social groups.

Keywords: reputation, reciprocity, cooperation, expertise

Reputation as an Animal Concept

In our daily lives, we often cogitate on matters of gossip and reputation. People get upset over a bad reputation. This was humorously illustrated in Anton Chekhov's 1883 short story "A Slander" ("Клевета"), whereupon a prestigious schoolmaster, attending his daughter's wedding, smacks his lips in approval of some delicious food prepared in the kitchen — and then later is obsessively chagrined after discovering that his innocent lip-smacking noise was heard by someone in an adjoining room, leading to a widespread rumour that he was adulterously kissing the female

I would like to thank two anonymous reviewers for their helpful comments. I would also like to thank Professor Raymond Russ (University of Maine) for extensive feedback and guidance. Correspondence concerning this article should be addressed to Dr. Yvan I. Russell, Department of Psychology, Middlesex University, The Burroughs, London NW4 4BT, United Kingdom, Email: yvanrussell@gmail.com

cook who was in the same room when he made the noise (Chekhov, 1883/1921). Stories like this illustrate "gossip" as a pejorative term, something that might unjustly cause embarrassment and suffering.

However, we can use the word "gossip" in a non-pejorative sense too: simply as a mechanism of social information exchange. Most human conversations are dominated by social gossip, which suggests that gossip has important functions (Dunbar, 2004). Gossip "allows individuals and communities to accumulate behavioural evidence about others and to form and refine judgements about their vices and virtues" (Emler, 1994, p. 133). In other words, gossip gives us important information. If a person has a bad reputation, then one might be inclined to avoid that person (lest one suffer the way that others have suffered). If a person has a good reputation, then one might be inclined to approach that person (to benefit the way that others have benefited). If you know nothing about a person's reputation, then you approach that person as a blank slate with no predictive information on whether you will encounter positive, neutral, or negative consequences. If the stakes are high, then it is useful to take advantage of information gathered by others. Here, we can bring in a biological concept: the producer–scrounger effect, where the thief (scrounger) takes advantage of the "behavioural investment of another (producer) to obtain a limited resource" (Barnard and Sibly, 1981, p. 543). Usually, this concept is applied to phenomena such as kleptoparisitism (stealing food from one who made the — perhaps risky — effort to acquire the food; this is a low-effort way to obtain food, and a loss for the other, e.g., Spencer, Russell, Dickins, and Dickins, 2017). We can apply the concept of producer–scrounger to social information gathering, where the "limited resource" is information. Imagine that you need information that will help decide whether to approach a man called Mr. Enemy. Imagine further that you witnessed Mr. Enemy injuring Mr. Friend. Here, Mr. Friend made the — perhaps risky — effort to "produce" information for you. By "scrounging" information produced by Mr. Friend (seeing him get injured), you have gained valuable information (that you should perhaps avoid Mr. Enemy) whilst avoiding personal injury yourself. There is an advantage to gathering information by proxy. The cognitive mechanism here is analogical reasoning: "if Mr. Enemy hurt Mr. Friend, then he will probably hurt me too." I return to the topic of "if–then" social reasoning (de Waal, 2003) later.

Like others before me (e.g., Dunbar, 2004), I am interested in gossip and reputation in an evolutionary context. I view it as important to adapt the concept of reputation in such a way as to accommodate the capabilities of the whole animal kingdom (Russell, 2007). The word "reputation" is normally used to describe an exclusively human activity: "person A gives testimony about person B to person C, teaching C something new about B." In everyday colloquial usage, "reputation" refers to this transitivity: information verbally passing along the gossip network from one person to another. Language is the crucial ingredient of this kind of information flow, and without it, information does not flow past those who perceived the signals first-hand. Verbal information becomes highly advantageous for cooperation when the size of a social group exceeds the point where an individual can rely on first-hand knowledge (see Greif, 1989, for an historical example). A good reputation is beneficial. As Alexander (1987, p. 95) wrote: "The concept of status implies that an individual's privileges, or its access to resources, are controlled in part by how others collectively think of him (hence, treat him) as a result of past interactions (including observations of interactions with others)." With this advantage in mind, it makes sense that people are motivated to behave well when they know they are being watched (Emler, 1990; Engelmann and Fischbacher, 2009). To adapt the concept of reputation for non-human animals, I start by recognising that reputation involves more than language. Information about others can be also gained from sheer observation (well within the capabilities of animals). Therefore, I have previously (Russell, 2007; Russell, Call, and Dunbar, 2008) defined reputation in an animal-inclusive manner as *knowledge about an individual's typical behaviour based on a knowledge of that individual's past behaviour*.

In many situations, the word "individual" can be replaced with "entity," because people routinely assign reputations to groups of individuals (even whole nations), corporate entities in the business world, or even insentient objects and phenomena. The attribution of reputations to groups is in several ways a parallel (but not identical) process to the attribution of stereotypes (cf. McGarty, 2002 and Spears, 2011). Reputation has been defined in many diverse ways by a large number of different researchers (e.g., see the reputation model developed for the business world by Carmeli and Tishler, 2005, who define reputation as the external perception of a business in terms of distinctiveness and prestige). The word "reputation" itself (like many words) can seem semantically opaque about its actual referent. It is not really a concrete "thing," but rather it is information which somehow has a life of its own beyond individual brains. Here, I focus on *individual* reputations. Reputations are not an inherent property of an individual, but are subjective attributions made by others (Obreiter, Fähnrich, and Gianluca, 2005; Pollock and Dugatkin, 1992). Among non-human animals, reputation can only exist completely outside the sphere of language. Animals respond deftly to motivational cues and signals from other animals (Krebs and Dawkins, 1984; Smith and Harper, 2004), but reputation is relevant only if knowledge about a particular individual's past behaviour is remembered and influences current behaviour towards that individual. Animals may learn the typical behaviour of others in three ways (Smith and Harper, 2004): (1) direct reputation (personal encounters), (2) indirect reputation (observing events as uninvolved bystander), and (3) reported reputation (gossip) [cf; Ostrom, 2003, pp. 43-44]. Whilst verbal gossip is surely uniquely human, the other channels of information (direct reputation and indirect reputation) are usable by animals to varying extents (depending on the species and its cognitive abilities; see discussion in Russell, 2007).

The actual content of human gossip varies widely. In humans, it is often about personality traits (such as Chekhov's unfortunate schoolmaster being labelled an adulterer) or about episodes that lead to personality attributions (that same schoolmaster being heard to make a kissing sound). Social psychological research on the topic of reputation has tended to focus on information related to personality attributes (Emler, 1990). In order to apply the concept of reputation to animals, we need to focus on more basic behaviour. Below, I will focus on the most basic "moral" behaviours. In doing so, I adopt the view that biology and morality are intertwined (cf. Alexander, 1987). Theorists such as Alexander (1987) and Binmore (2005) promote an empirical and naturalistic view of morality: instead of prescribing rules based on abstract principles, we can study what humans actually do. We can view morality as being based on social contracts, and success measured by the establishment of *equilibria* (see below about the Nash equilibrium). However, as Hamilton (1975) wrote, contractual morality has an in-built uncertainty: "It is very frequently necessary for one party to execute his half of a bargain without any way of being certain that the other party will later stick to his" (p. 150). Below, I review the concepts of direct reciprocity, indirect reciprocity, and reputation as mechanisms that help ensure that the second half of the bargain is met.

Direct Reciprocity

There are many forms of dyadic (two-person) reciprocity (see Dugatkin, 1997), but here I describe the most basic form (Binmore, 2005; Dugatkin, 1997; Sigmund, 2010): "When individual A copies what individual B does. Hence, if B gives, A gives back; if B fails to give, then A defects in return" (failure to reciprocate is called a "defection"). If a reciprocal relationship lasts for multiple rounds, then it can take the appearance of a feedback circuit or loop: A pays B, then B pays A back. Kolm (2000) averred that reciprocity is classifiable in two different ways: according to allocation (the actual budget of given and received items) and process (the mechanism that enabled it — for example a psychological motivation). In this paper, we will reflect on both allocation and process.

Stable cooperative relationships are formed as a summation of repeated interactions, the exact pattern of which is unique to a particular dyad (Bowles and Gintis, 2011; Hinde, 1976). Mutually beneficial relationships are established after a progressive reduction of uncertainty between two actors about the benefits that arise after one actor signals an intention to benefit the other (Markl, 1985). A feedback circuit is perhaps not the best metaphor, because being a sender/receiver is a role (rather than a characteristic) of an organism (Markl, 1985). For both parties to continue to respond to each other, there needs to be some two-way payoff; otherwise, "nothing in the world can keep receivers dancing like puppets on the strings of the senders' signals — unless it is to their own advantage, too, to be manipulated" (Markl, 1985, p. 165). Responsiveness can be called "tightness" (Markl, 1985): how tight the loop is between signal and response (and whether a response occurs at all). When direct (pairwise) reciprocity occurs between sentient animals in natural settings, the "circuit" is created through behavioural episodes occurring at fairly unpredictable intervals between organisms who may or may not transact again. Each dyad, furthermore, does not stand in isolation, but is embedded within the complexity of an ecological niche with its connate social network (Alexander, 1987; Clutton–Brock, 2009; Hinde, 1976; McGregor, 2005; Nowak and Highfield, 2011; Ostrom, 2003). Reciprocation does not necessarily consist of costs and benefits for each side; benefits may flow with both parties gaining rather than losing, becoming what is termed pseudoreciprocity or by-product mutualism (Alexander, 1987; Clutton–Brock, 2009; Dugatkin, 1997). It is useful to focus on the simpler constituents of prosociality as a step towards understanding the larger social/cooperative structure of animal and human societies (Alexander, 1977, 1987; Clutton–Brock, 2009; Dawes, 1980; Dugatkin, 1997; Hinde, 1976; McGregor, 2005; Nowak and Highfield, 2011; Sigmund, 2010, etc.). Reciprocity is widely regarded as a key mechanism in human sociality (Fehr and Gächter, 2000; Kolm, 2000). Episodes of reciprocation — so common across cultures, between friends or strangers, ingrained in our social norms — are elementary units of our cooperative societies which provide public goods, and which contribute to the survival of members of our species (Alexander, 1987; Dawes, 1980; Nowak and Highfield, 2011).

Markl (1985, p. 170) identified four scenarios of payoffs for dyadic relationships: (1) both actors benefit (cooperation), (2) the sender benefits but not the receiver (exploitative), (3) the receiver benefits but not the sender (also exploitative), and (4) neither benefit (a wasted effort). Direct reciprocity falls into the first category, but only if reciprocation occurs (otherwise it is exploitation). We can analyse these relationships using *game theory*: where a player's probability of payoff is contingent on the behaviour of others (see Binmore, 2005; Dugatkin, 1997; Sigmund, 2010). The colloquialism "I'll scratch your back, you scratch mine" is often invoked as a one-liner summary of reciprocity. To prevent oneself from suffering defection, it is helpful to avoid one-shot encounters and benefit from repeated encounters with reliable individuals. Binmore (2005, p. 10) elaborated:

> Rational reciprocity can't work unless people interact repeatedly, without a definite end to their relationship in sight. If the reason I scratch your back today is that I expect you will then scratch my back tomorrow, then our cooperative arrangement will unravel if we know that there will eventually be no tomorrow.

However, the possible extent of calculation is limited. Imagine that a human being (the scratcher) has his emotional life and culture all stripped away and what remains is a coldly calculating person who thinks only of the payoff. For example, this robotic-type might say to himself: "if I scratch his back for ten minutes, then there is a 95% chance that he will scratch my back for ten minutes within a week." No person actually thinks quantitatively in this manner (precision would be impossible). However, we are influenced by this kind of rationalistic dynamic on a functional level. We might make the analogy of a rat in a Skinner box, being influenced by the principles of operant conditioning (e.g., Guttman, 1953). The actual quantitative explanation of the rat's behaviour is only calculable by the scientist standing outside the box taking measurements. The rat itself cannot understand the operant principles governing its own behaviour. We humans might regard ourselves intellectually superior to a rat — but psychologists (e.g., Simon, 1955, 1983) have known for a long time that cognitively we just do not follow the economic rules of "rational man"; instead, we put in just enough mental effort to attain some desired outcome (Gigerenzer, 1997) because we are not privy to the full information that would enable us to maximise our benefits at every step of our daily behaviour (Simon, 1983). Furthermore, we are highly imperfect reasoning machines, subject to numerous biases (Ayton, 2010). We humans are typically like the rat in the Skinner box, and this includes situations where we respond to costs and benefits of reciprocal interactions (Binmore, 2005; Ostrom, 2003), much like how the unreflective rat in the Skinner box responds to the benefits of pushing a lever. That non-human animals show reciprocation behaviours is well established, although the proximate mechanisms (processes) are debated (see Clutton–Brock, 2009).

Game theory is a system for investigating how payoffs differ according to the strategy adopted. Payoffs can be anything. For chimpanzees, payoff might literally be the receipt of "back-scratching" (i.e., social grooming, see Russell and Phelps, 2013). For humans who play economic games, the payoff might be money (Dawes, 1980). In direct reciprocity among moneyless organisms, the payoff might be your future reproductive success (Trivers, 1971). PAYOFF is a generic concept. Accordingly, game theorists refer to *utility* (and its unit of measurement, *util*), a generic unit of payoff that results from a given decision (Binmore, 2005; Simon, 1955): as a currency, an util can be anything (whether it's reproductive success, food, actual money, etc.) and even when undefined it can be used as a variable in biological or cognitive modelling (see Bowles and Gintis, 2011). In evolutionary game theory, costs and benefits can be numbered as "fitness units" addable or subtractable from a baseline fitness (Sober and Wilson, 1998). In our everyday thinking, we lack the perfect information that allows robotic-like total rationality — so instead we rely on our limited information and use cognitive shortcuts (Ayton, 2010; Gigerenzer, 1997; Ostrom, 2003; Simon, 1983; Sober and Wilson, 1998; Sutherland, 1992). For example, the take-the-best strategy (Gigerenzer, 1997) is a proposal that binary decisions (choosing one or the other) are made using as few cues as possible (cf. Simon, 1955). It is plausible that a "take-the-best" strategy is applicable to binary decisions in the social realm too (e.g., decide to interact with someone or not).

Sigmund (2010) delineated the mechanics of direct reciprocity in game theory terms (below described verbally, not mathematically). First, consider the Nash equilibrium (see also Binmore, 2005 and Ostrom, 2003) by imagining that player 1 can play two possible strategies — *e1* and *e2* — against player 2. This is a probability: if I play strategy *e1* this time, what is the probability that I will play the same strategy (*e1*) *next* time? It all depends on player 2, who (for example) has two possible strategies of her own: *f1* and *f2*. The Nash equilibrium is all about your "best response" to the other person's strategy. Suppose player 1 chooses *e1* and player 2 chooses *f1* — should player 1 stick to *e1* or switch to *e2*? If the payoff is higher by playing *e1* (instead of *e2*) in response to *f1*, then player 1 will likely keep on playing *e1* (his best response to *f1*). Remember, though, that the other's dyadic game *also* consists of two strategies. Player 2 might have her own Nash equilibrium — for example that *f1* is the best response to *e1*. Accordingly, the players can form an *equilibrium pair* and keep going in that same pattern which would prove beneficial for both. Suppose, however, the player 2 changes her strategy to *f2*. This might change the payoffs for player 1 and perhaps his new best response is now *e2* (or perhaps it stays the same). Thus, the game changes according to the behaviour on both sides of the dyad. Establishing equilibrium pairs through repetition is key to establishing a cooperative relationship within a dyad. Based on this, there have been many models of dyadic cooperation in the literature (see Binmore, 2005; Bowles and Gintis, 2011; Dugatkin, 1997, Sigmund, 2010). Games, of course, can have multiple equilibria, not just two (Binmore, 2005). Furthermore, there is huge variability of human behaviours, meaning that Nash equilibria are often reached at the group rather than individual level (Ostrom, 2003). As Alexander (1987) wrote, successful sociality is about "flexible strategizing" (p. 9). Such games, as described above (and laboratory experiments designed to test them), cannot begin to capture all of the messy complexity of real life (Binmore, 2005; Ostrom, 2003): nonetheless, such games are useful tools for understanding the principles that explain behaviour.

Imagine a different scenario where an organism is not only two, but many — living in a finite population where individuals might have three possible dyadic strategies (Sigmund, 2010): (1) always cooperate, (2) always defect (i.e., never give anything), (3) be choosy and do tit-for-tat (cooperate when meeting a co-operator, defect when meeting a defector). There have been many agent-based computer simulations where individual agents are programmed to use only *one* strategy each (e.g., Nowak and Sigmund, 1998). A typical such simulation comprises a series of iterations (one generation after another) and, after the program starts running, we assess the simulation by looking at how the proportions of types (cooperate/defect/choosy) alter over time. There are many questions to ask here. What type of agent will succeed in such a simulation? Will the population be overtaken by defectors? Will the co-operators prevail? Or, is choosiness the only path that allows cooperation to flourish? You can look at a population and see how the percentages of each changes over time (e.g., see Bowles and Gintis, 2011; Nowak and Sigmund, 1998; Sigmund, 2010). For example, if the simulation begins with 100% co-operators, then it is fairly easy for defectors (who might appear suddenly due to mutation) to take over; but the defectors cannot dominate the population if discriminators are present. It all depends on proportions (Sigmund, 2010; Sober and Wilson, 1998): the initial mix of types, and how the interactions cause some groups to benefit over others (e.g., co-operators diminish in the population because they are giving too much away and not getting anything in return). The idea is that those who have successful strategies replicate (produce offspring using the same strategy) while those using unsuccessful strategies head towards extinction (if you don't get enough favours, you don't live to replicate). Therefore, strategies are said to "evolve" (increasing, decreasing, or staying about the same across iterations). How does a scientist decide the outcome of this complex mix? The possible outcomes can be derived from the *replicator equation* (described mathematically in Bowles and Gintis, 2011 and Sigmund, 2010). The replicator equation helps to predict how quickly a particular strategy will grow within this finite population, and the answer is that "a strategy *ei* will spread or dwindle depending on whether it does better or worse than average" (Sigmund, 2010, p. 31). The equation takes the average payoff to a particular strategy (e.g., a co-operator) and subtracts from that the average payoff of all individuals using all strategies (e.g., co-operators, defectors, and discriminators). Who tends to win, then? Generally, it depends on the numbers (how much percentage of each exists in the first place), but the discriminating strategies often win out. Defectors lose out when discriminators notice they are defecting; co-operators lose out because they are not choosy — but they can flourish when there are few or no defectors around (Bowles and Gintis, 2011; Dugatkin, 1997; Sigmund, 2010). Sometimes one can even get *rock-paper-scissor* dynamics (Sigmund, 2010): the proportions of each strategy oscillate, whereupon every strategy takes turns in being the most common.

A primary lesson here is that it pays to be choosy (Sober and Wilson, 1998). Referring to the environment of ancestral humans, Bowles and Gintis (2011) wrote that "those who failed to distinguish between long-term or short-term or one-shot interactions would be at a significant fitness disadvantage" (p. 96). In other words, treating everybody as a trusted friend will not necessarily benefit you. Given this risk, there must be some underlying principle that explains why people habitually act pro-socially, even to strangers.

Reputation and Indirect Reciprocity

Add a third person to a dyadic interaction and a triad emerges (Faust, 2007). When a three-way interaction occurs, it is nearly impossible for each actor to engage in precisely 33% of the interaction. There is inevitably some imbalance, with two of the actors more deeply involved than the third (cf. chimpanzee grooming patterns described by Russell, 2007). Sometimes the third actor is not directly involved at all, but is merely watching the interaction. The proportion of non-involved individuals will increase further as the group size increases. This situation (being a triad or higher) sets the scene for indirect reciprocity. As Alexander (1987, pp. 94–95) described it:

> I regard indirect reciprocity as a consequence of direct reciprocity occurring in the presence of interested audiences — groups of individuals who continually evaluate the members of their society as possible future interactants from whom they would like to gain more than they lose (this outcome, of course, can be mutual).

In real life, cooperation is multidirectional. This is true throughout nature: all the way from the level of RNA hypercycles to that of human teamwork (see Bourke, 2011). Many models of cooperation have focused on the dyad (e.g., Bowles and Gintis, 2011; Dugatkin, 1997; Ostrom, 2003; Trivers, 1971). These dyadic models are predicated on the phenomenon of *trust* (see Kohn, 2009) — as well as using your knowledge of past behaviour to guide your current behaviour (for reviews and discussions, see Alexander, 1987; Dugatkin, 1997; McElreath, Clutton–Brock, Fehr, Fessler, Hagen, Hammerstein et al., 2003). Evidence from human studies shows that face-to-face contact is very important in establishing trust (Ostrom, 2003); but face-to-face contact cannot always happen. As a population grows larger, the probability of repeated interactions is reduced (because a person encounters strangers more and familiars less). In this case, being choosy is an essential strategy, because an indiscriminately generous person in a mixed population will always end up with a lower payoff than defectors (see Sober and Wilson, 1998, pp. 19–23). If learning whether to trust someone depended solely on direct encounters, then helpers are vulnerable to defection when helping a stranger the first time (Pollock and Dugatkin, 1992). This is a problem that can be by-passed if the helper has prior knowledge of how the potential recipient behaved in the past towards others. Reputation is useful here. It is that knowledge source.

Alexander (1977, 1987) proposed that indirect reciprocity ("A observes B help C, therefore A helps B") — a system that rewards the generous and punishes the selfish — is a defining mechanism of human moral systems (also see Binmore, 2005; Bowles and Gintis, 2011; Dugatkin, 1997; McElreath et al., 2003; Nowak and Highfield, 2011; Sigmund, 2010). Indirect reciprocity can occur in other forms too, such as "A helps B, B helps C, C helps A" (Alexander, 1987, p. 81), a form which will not be covered here. Another name for indirect reciprocity is "vicarious reciprocity" (Sigmund, 2010). The population-level benefit of indirect reciprocity might be simply summarised by saying that "everyone may gain when social beneficence is prevalent" (Alexander, 1987, p. 210; cf. Kohn, 2009). However, indirect reciprocity is *not* a synonym for "generalized reciprocity" (Alexander, 1987, p. 85). According to indirect reciprocity, when people are good, the strategy is ultimately self-serving despite the up-front costs (Alexander, 1987). The benefits of indirect reciprocity towards a well-regarded individual can manifest in at least three ways: (1) direct compensation from all or part of a group (e.g., when someone is deemed a hero), (2) more opportunities to engage in fruitful interaction due to being approached by third parties who witnessed the generosity, and (3) the generosity ultimately benefits the group to which the generous person is a part — and perhaps even benefiting that person's own descendants (Alexander, 1987, p. 94).

Importantly, the terms indirect *reputation* and indirect *reciprocity* should not be confused. The former refers to an information source and the latter refers to the moral/social system that is enabled by the information source (Alexander, 1987; Nowak and Sigmund, 1998). Reputation is construable as one component of an interacting system where repeated social dilemmas (Dawes, 1980; Ostrom, 2003) are worked through when reputation feeds into trust, which feeds into the probability of reciprocity, leading hopefully to the best collective outcome possible (see Ostrom, 2003, pp. 49–61). Indirect reciprocity requires that group members monitor each other's reputations, ideally creating conditions where generous individuals prosper and selfish individuals suffer (Alexander, 1987; McElreath et al., 2003; Nowak and Sigmund, 1998; Pollock and Dugatkin, 1992). Over the years, a series of agent-based computer simulations have been developed to explore this possibility using an image scoring paradigm (Brandt and Sigmund, 2005; Nowak and Sigmund, 1998; reviewed in McElreath et al., 2003). "Image score" is a numerical measure of how generous an individual person has been to others. The aim of these simulations was to explore the conditions under which image scoring individuals (those who preferentially give rewards to those with sufficiently high image scores) would dominate a population that also consists of defectors (never help anyone) and unconditional givers (help others indiscriminately). The main conclusion from these models is that helping is an evolutionarily stable strategy (the population resists being overwhelmed by defectors; see Parker and Smith, 1990) only if the majority of the population consists of *strict* image scorers (Brandt and Sigmund, 2005; Nowak and Sigmund, 1998). The models inspired a series of real-life human experiments which showed that people actually do spontaneously consider reputation when deciding whom to reward; and that they behave more cooperatively in order to preserve their good reputations (Engelmann and Fischbacher, 2009; Wedekind and Milinski, 2000, etc.). Important to note, however, is that indirect reciprocity is not the only mechanism for preventing defections. Many large-scale human endeavours come about through institutionalisation — entailing the creation of formal organisations where things are put in writing and mechanisms designed to put principles above individual proclivities are in place (Alexander, 1987; Fehr and Gächter, 2000; Urpelainen, 2011). For example, think about the massive amount of planning and cooperation needed to successfully operate a highly complex entity such as London Heathrow Airport (Wicks, 2014): multiple levels of organisation are needed to manage more than 76,000 people — each with a unique set of skills and stipulations — in their respective roles all geared towards the simply-stated (yet highly tricky to coordinate) goal of managing airplane arrivals and departures (an average of 1400 per day). Heathrow is a conspicuous example — but there are countless other types of organisations, large and small, that would be difficult or impossible to run without coordinated (and often highly regimented) action between strangers. Institutions often save us the trouble at needing to gather social information helping us decide with whom to work. In a place like Heathrow Airport, one does not usually need to know the reputations of those with whom one cooperates in order to get a plane to fly: people know each other's roles by default (baggage handler, pilot, etc.) and can therefore successfully collaborate with complete strangers constantly. Defecting is minimised through a set of rules and punishments (see Fehr and Gächter, 2000 for a review of reciprocity and punishment in the workplace; cf. Sober and Wilson, 1998). We might consider Heathrow Airport as a highly codified, almost reputation-irrelevant zone. This is one extreme on a spectrum of social situations. Another extreme is a setting consisting of familiars only: the kin, the friends, the neighbours, etc. This is where an abundance of information about the people one can interact with lies: not only that of an individual person, but the relationships between those individuals. Non-human animals, of course, usually exist (in the natural world) only amongst familiars (Hinde, 1976). It is we humans who cast the net wider.

Let us think again about the *replicator dynamics*, this time for indirect reciprocity (Bowles and Gintis, 2011; Sigmund, 2010), referring (as before) to the simulated population consisting of co-operator, defector, and discriminator. Now, the discriminator (the choosy one) needs to rely on the image score, which in a computer simulation can be as simple as 0/1 (known to be either generous in the previous round or not). The replicator dynamics equation here needs to incorporate the payoff for the reciprocator, who gives out a benefit only if the recipient has a good reputation, or when no information is available (in other words, reciprocators cooperate except when encountering a bad reputation). How can a discriminating strategy evolve in this setting in order to produce a simulation where cooperation is dominant? This depends, first of all, on whether reputations are knowable. As Sigmund (2010) wrote, "if the probability… to know the co-players' past is too small (i.e., if there is not much scope for reputation), then cooperation cannot evolve.… [A] cooperative population consisting of these two types of altruists (some conditional and some not) exists, if the average level of information within the population is sufficiently high" (p. 86). Once reputation becomes possible, then it all depends on the numbers: what percentage of the population is occupied by either co-operators, defectors, or discriminators. Obviously, too high a percentage of defectors will not allow cooperation to flourish — and too high a percentage of "gullible" co-operators will simply allow the defectors to take over unbridled. What is needed is a high-enough proportion of discriminators to be the bulwark against the takeover by the selfish. A cooperative population can exist, even if there are many defectors and many gullibles, as long as the largest group happens to consist of discriminators (see Sigmund, 2010 for the mathematical treatment and implications of varying parameters). Now, let us think again about the knowability of reputation. Thinking across all animals, this depends very much on cognitive ability. Examples like Heathrow irresistibly remind us of insect societies, such as leafcutter ants (Nowak and Highfield, 2011) where success is implemented by seven different castes (anatomically differentiated) to carry out specialised tasks (within a world of chemical signalling). Clearly, complex cooperation arises non- or minimally cognitively across all facets of life (Bourke, 2011). This is why it is important to clarify the issue of when and why indirect reciprocity needs deliberation and when it does not.

The Cognitive Substrate of Indirect Reciprocity

In encountering animal studies, the temptation is often to infer human-style cognition. It is obvious that animals collect information (McGregor, 2005). As an information source, indirect reputation will flow ubiquitously from any animal communication network where it is possible to eavesdrop without being directly involved (Markl, 1985; McGregor, 2005).[1] The question, if we are thinking across the animal kingdom, is: What depth of processing occurs in animals living within these communication networks (Russell, 2007)? Evolutionary and psychological explanations of cooperative behaviour are interrelated (Sober and Wilson, 1998, pp. 203–205). It is useful to think of biological explanation the way Tinbergen (1963) delineated, in which every biological explanation can be construed in four ways: (1) phylogenetic (how it evolved), (2) ontogenetic (how it develops), (3) functional (why it evolved), and (4) proximate (the actual mechanism that enables it). In the animal kingdom, all examples of indirect reciprocity will have a functional explanation (number 3 above). Kolm's (2000) prescription for reciprocity can be applied to indirect reciprocity: thinking separately about allocation (the actual budget of given and received items) and process (the mechanism that enabled it). The questions of interest to psychologists tend to be those of process, that is, the proximate mechanisms (usually favouring cognitive explanations, with a special bias towards assuming conscious awareness). It is often difficult to write about non-sentient evolutionary processes in a way that does not sound like one is writing about characters in a play. This is why it is important to reiterate Tinbergen's "four whys." The issue of what is happening in the animal's mind when it engages in indirect reciprocity is a proximate-level description. At its simplest, indirect reciprocity is describable in terms of a dyadic interaction

[1] Here, I use the word "eavesdrop" to refer to any modality, whether it is from sight, sound, chemical senses, or other means.

(e.g., "A → B") that would not have occurred unless the actor had been primed by having observed an earlier interaction involving the recipient (e.g., "A had observed B → C"). Of course, this structure can also describe punishment (e.g., "A attacked B, because earlier B attacked C"). Three conditions are necessary for (positive) indirect reciprocity to occur: (1) favours occur in a setting observable by third parties, (2) the third party is motivated to reward the rewarder, and (3) the third party is influenced by indirect reputation.

The first condition, observability of behaviour, is the prerequisite for all phenomena involving reputationally based cooperative behaviour. Among humans, behaviour might be observed second- or third-hand through verbal gossip, whereas pre-linguistic animals are limited to direct observation. A wide variety of species have evolved observational skills that effortlessly detect cues and signals emitted from others regardless of where the emitter was aiming (it is possible, of course, that a signal is emitted without an intended direction). In this context, a field of *public information* evolves in the mind's eye of the species — the populations of which are now able to interpret "inadvertent social information" (Danchin, Giraldeau, Valone, and Wagner, 2004). The second condition concerns the proximate motivation of the third party in repaying the favour on behalf of the recipient. There are four ways to partition this, as explained below.

1. *There is no motivation (on a cognitive level).* Above, indirect reciprocity was described in its simplest form: an interaction occurring as a consequence of the actor being primed by a prior interaction involving the recipient and a third party. Defined this way, indirect reciprocity can be identified anywhere the above causality is established, regardless of the level of cognition of the actors. Some examples are found among cleaner fish and their "clients" (Bshary, 2002) ["client" refers to the recipient of prosocial behavior, e.g., other species of fish]. Here, clients observe the interactions of cleaner fish towards the third parties; those cleaners who defect (eat living flesh in addition to the dead flesh that they are supposed to be cleaning off) are avoided by the clients more than cleaners who do not cheat (Bshary, 2002). Indirect reciprocity is likely restricted here to a functional rather than an explanation. How much the fish actually understands the third-party interactions is open to debate, but the point shown by Bshary is that such events can be identified in cases where sophisticated cognitive abilities are unproven (it is possible, of course, that fish are cleverer than we think — but we are safest for now in assuming that indirect reciprocity is happening with minimal cognition in this class of animal).

2. *There is a selfish motivation.* Alexander (1987) suggested that individuals may reward the rewarder simply as a by-product of their desire to interact with someone known to be cooperative. If cleaner fish clients (Bshary, 2002) were human-style conscious beings, then their motivations might be regarded as selfish: they reward non-cheating cleaner fish by offering themselves, in the process rewarding both themselves (being cleaned) and the cleaner (who obtains food).

Traditional economic theories predict humans to behave this way, but empirical results show that humans in economic games behave selfishly only part of the time (Binmore, 2005; Schram, 2000). This implies that humans have something to gain by behaving pro-socially (see next point). Interestingly, chimpanzees seem more likely than humans to behave in a self-interested manner consistent with traditional economic theories (Jensen, Call, and Tomasello, 2007).

3. *There is a motivation to behave pro-socially "for its own sake."* Schram (2000) identified three reasons that humans might pursue a conscious pro-sociality. The first is the "warm glow of giving" where an individual cooperates because it feels good to do so (see also Binmore, 2005; Pradel and Fetchenhauer, 2010; Sober and Wilson, 1998, pp. 267–271). As mentioned above, such feelings could arise by association with past positive outcomes. This means that being generous (e.g., giving to a charity to help starving children) can actually be selfish: "If you were an egoist, you would help the starving, but your ultimate motive would be to make yourself feel good" (Sober and Wilson, 1998, p. 244). The second reason is fairness, where an individual cooperates on the contingency that the partner co-operates. This is the basis of a number of tit-for-tat cooperation models (Dugatkin, 1997; Trivers, 1971), which in humans involves knowing reciprocity norms (Binmore, 2005). Operating in this manner requires that an individual engage in mental score-keeping: keeping track of all past activities of one's trading partners (see also Call, 2002; Schino and Aureli, 2009). Theoretical models have introduced many variations on this general idea (e.g., contrite tit-for-tat; see Dugatkin, 1997). Among apes, mental score-keeping is unproven but possible (Call, 2002; Schino and Aureli, 2009). Whether apes can really recognise reciprocity norms is unclear (cf. Tomasello, Carpenter, Call, Behne, and Moll, 2005). The third reason identified by Schram (2000) was other-regarding: concern for the well-being of someone else (perhaps emotionally charged). All of the above reasons could apply to indirect reciprocity.

4. *There is a motivation based on consciously known cooperative gains.* This occurs when an individual understands collaborative behaviour (Schram, 2000): that a goal cannot be accomplished unless individuals work together (cf. Dawes, 1980; Dugatkin, 1997; Ostrom, 2003). This is not a difficult task if two individuals are standing side-by-side co-operating to achieve an immediate goal. Even chimpanzees can do this if trained, but only if the reward is forthcoming to both. If one chimpanzee does not receive an immediate observable reward, then that individual will probably not cooperate (Jensen, Hare, Call, and Tomasello, 2006). In the case of indirect reciprocity, this would require an explicit understanding concerning how indirect reciprocity maintains the overall pro-sociality of the group. In real-life terms, this might be seen among human societies where community spirit plays a salient role in people's lives, for example, in the Indonesian farming community that Schweizer (1989) studied, where the concept of neighbourly harmony was an influential factor in social and religious life. This prevented

families from adopting practices that maximised financial income at the expense of community members (cf. Dawes, 1980). Two examples were cited: (a) farming families chose labourer hiring practices that distributed wealth more evenly, and (b) families never eschewed their obligation to host a *slametan*, a religiously-based feast that was thought to spiritually benefit the whole community (despite the expense to the family). These examples represent a form of indirect reciprocity, because generous behaviour is rewarded (but not in a dyadic tit-for-tat manner). Although some of this pro-social behaviour is probably motivated by a fear of ostracism (Schweizer, 1989), everyone is aware of the relationship between one's own behaviour and the community's well being (cf. Urpelainen, 2011).

As illustrated, there are many paths to indirect reciprocity. It is a behavioural example of the evolutionary principle of functional equivalence, meaning that it is possible to identify a number of different proximate mechanisms which "all deliver roughly the same behaviors in the same circumstances" (Sober and Wilson, 1998, p. 206). Indirect reciprocity is a beneficial strategy in particular circumstances, and different animals have evolved different levels of necessary cognition. Among humans, each of the above levels of motivation should occur. Moreover, selfish and pro-social motivations can co-exist and intermingle (for example, "person A might have initially helped B for selfish reasons, had been rewarded, and then started to care for B's welfare without wanting a reward," cf. Sober and Wilson, 1998, pp. 217–222, 242–250, 319–321). Although we congratulate ourselves on being able to adopt the most cognitively sophisticated strategies in the animal kingdom, much of the time we are likely getting by with a minimum of cogitation (Newell and Shanks, 2014): indirect reciprocity may occur unplanned, be a by-product of a selfish motivation, or follow genuinely pro-social sentiments motivated either by a feel-good factor or from a self-aware intention to contribute to the common good (see Ostrom, 2003). Among non-humans, there is still much to learn about how this works. To address this question, the best approach is probably to emulate approaches equivalent to Byrne and Whiten (1997) when they searched for deception in the animal kingdom (see also Byrne, 2003): find as many instances as possible where deception occurs and only later start worrying about characterising the cognition (if any) that is involved. We can do the same for indirect reciprocity.

Thus, when indirect reciprocity occurs, premeditation may not be necessary. It should be useful to identify instances of indirect reciprocity in nature as a possible context where individuals are heeding each other's reputations. The evidence for indirect reciprocity among non-human primates is sketchy, but there are possible candidates in the literature. One example is the possible revenge system described by Aureli, Cozzolino, Cordisch, and Scucchi (1992) in macaques, where individuals would attack the family member of an aggressor. Another example is in a catalogue of "triplet interactions" by Mori (1983) on free-ranging chimpanzees: sequences of behaviours where a dyadic interaction was soon followed by a different dyadic interaction involving one of the previous interactants (see Mori, 1983, Table 8). Twenty different types of triadic interaction were identified, two of which may qualify for indirect reciprocity. One interaction type was where "a first appeasing chimpanzee became the recipient in the second interaction" (Mori, 1983, p. 58, #4 in Table 8). The other one was a revenge incident (#9 in Table 8) similar to that reported by Aureli et al. (1992). Reputation probably plays a role in such interactions. For a researcher interested in reputational thinking, the key issue is whether an individual can compare relationships between self and other to relationships among third parties (Russell, 2007). At best, we can take a "bird's eye view" and imagine living within a complex social structure and be aware of our place in it. Most of the time, we are not doing this explicity.

Social Expertise

It would be valuable to apply cognitive–psychological models of expertise to the social and non-human sphere. I will start with what Markl (1985, p. 165) said about the recipient of animal signals:

> Of course, addressees are not sitting around in extra-evolutionary space offering [unmodifiable] releasing mechanisms just waiting to be manipulated; in fact, it is well known that there is hardly anything that can be more easily modified both by evolution and individual experience — where we call it focusing of attention — than reaction thresholds and response selectivity of releasing mechanisms or sensory-neural pattern recognition devices.

Not learning means not surviving. "Inflexibility or preprogramming would be the worst possible strategy in the face of conflicts of interest, competition, the importance of cooperation, and other aspects of sociality" (Alexander, 1987, p. 9). The history of interactions that leads to something we can call a "relationship" (Hinde, 1976) is also the story of successive learning experiences, gauging and re-gauging expectations (Markl, 1985). This might lead to something we can call expertise.

Social expertise is something that Humphrey (1976) compares to chess: a game played with a reactive partner, where competence depends on accumulated knowledge, ability to keep track of changeable circumstances depending on the opponent's behaviour, and planning ahead according to what others may do. The Machiavellian intelligence hypothesis (Byrne and Whiten, 1997) construed social expertise as a skilled manipulation of others for personal gain. This requires "mind reading" ability (Byrne and Whiten, 1997), which is generally the skill of visualising the point of view of another's perception and intentions, seeing how intervening variables alter such intentions, and being able to identify deceptive behaviours (cf. Tomasello et al., 2005). This view of expertise can be nested within a broader framework called the social brain hypothesis (Dunbar, 2003; Dunbar, Gamble, and Gowlett, 2014), where function is emphasized: the burden of maintaining an optimal size of one's personal social network within larger social groups. According to this view, ecological factors put pressure on individuals to form large groups in order to enhance survival (Bourke, 2011; Dunbar, 1988), which in turn creates selective pressure for the evolution of increased size in the areas of the brain that facilitate social expertise (Dunbar, 2003). There appears to be a widely held opinion that the key skill of manipulation is an ability to anticipate another's behaviour based on a talent for mind reading (a.k.a. theory of mind or experience projection) [e.g., Byrne and Whiten, 1997; Dunbar, 2003; Humphrey, 1976]. The term "mind reading," of course, has been used to describe intention reading even in less cognitively advanced animals, referring to an animal's innate reactions to certain cues and signals (cf. Krebs and Dawkins, 1984). What makes mind reading "expert" is the knowledge base that the animal draws upon in order to behave proficiently. As Donald (2001) wrote, such mental feats "demand considerable memory, since each individual must have a 'slot' in the tracker's mind, which must be kept up to date" (p. 129). Monkeys and apes accumulate considerable knowledge about their conspecifics in at least three domains (Call, 2002): (1) information about how individuals behave, (2) the quality of their relationships with others, and (3) the quality of relationships among third parties. This type of knowledge base permits an individual to engage in social manipulation: where a manipulator induces a conspecific to behave in a certain way in order to accomplish a goal desired by the manipulator. For example, begging to receive food is a form of dyadic manipulation ("A induced B to do X") [Call, 2002]. A more complex skill is triadic manipulation (Call, 2002), where a manipulator induces a conspecific to behave a certain way towards a third party ("A induced B to induce C to do X"). Both dyadic and triadic manipulations are known as "social tool use," an expert skill that requires accumulated knowledge about the typical behaviour of others, along with some ability "to generate hypotheses about who interacts with whom, when and how" (Call, 2002, p. 178).

Even dyadic manipulation requires some form of indirect reputation — because the only way for A to learn the causal chain between a conspecific and a desired outcome is to observe how B behaves towards something in the environment, and to see how that behaviour leads to the outcome. What differentiates this from technical tool use (observing how a tool behaves towards the environment) is the fact that the social tool involves an animate being (Call, 2002), and hence there is a built-in source of uncertainty (whether B will behave the way anticipated). Triadic manipulation entails the same observational learning as dyadic manipulation, except that the causal chain now has two sources of uncertainty (whether either B and C will behave the way anticipated). How might an individual overcome such uncertainty in order to make the social manipulation work? Pure luck is obviously a factor (but one that likely underpays). What might facilitate higher payoffs is predictive ability, where individual A has learned — by experience — what behaviours to expect from individuals B and C during the relevant events. To illustrate this advantage, consider the following hypothetical situation about food sharing.

Imagine that individual A observes that C is monopolising a local food source (i.e., can prevent others from getting it). A desires the food, but knows from past experience that C will not offer any food if approached. However, A knows two other things:

(1) if B approaches C, then C always gives food to B

(2) if A (self) approaches B, then B always gives food to A

Here, triadic manipulation can occur if A induces B to approach C for food. When B obtains the food and carries it away, this provides an opportunity for A to beg for food from B. Additionally, suppose that individuals D and E are also nearby, but that A knows three other facts:

(3) if D begs for food from C, then C will refuse

(4) if A begs for food from E, then E will refuse

(5) if E begs for food from C, then C will give food to E

Obviously, it is pointless for A to engage D or E in social manipulation because D cannot obtain food from C, and E won't give it to A (even if getting it from C). If A knows this, then A will approach B and nobody else. This might comprise a form of declarative memory (Anderson, 1983): learned factual knowledge stored as long-term memory traces (interconnections). Facts in the long-term memory can be organised into themes, whereupon a number of thematically related facts are interconnected. This is one basis for an expert memory. It might be useful here to refer to general theories of expertise — in order to illuminate mechanisms that might also apply in the social domain. On the topic of non-social human expertise, there is a long and rich history of psychological testing and theorising (de Groot and Gobet, 1996; Ericsson, Charness, Feltovich, and Hoffman, 2006; Gobet, Chassy, and Bilalić, 2011; Russell, 2011; Sternberg, 1997). Some primate researchers have made detailed comparisons between physical and social reasoning, in an attempt to delineate commonalities and contrasts in the required intellectual abilities between species (e.g., Call, 2002). The cognitive mechanisms of human expertise have been characterised in many different ways over the years, but there is general agreement that expert skill acquisition involves deliberate practice, learning a large number of relevant patterns, cultivating a long-term memory base where memory traces are flexibly accessed, and understanding how to respond appropriately to meaningful patterns (Anderson, 1983; Gobet et al., 2011; Russell, 2011; Sternberg, 1997). In the food-sharing example presented above, individual A knew the reputations of individuals B, C, D, and E based on past experience. As mentioned earlier, typicality is key (cf. Emler, 1990). Another key issue in expertise is the amount of accumulated knowledge (Sternberg, 1997), and this is applicable in the social realm too.

Table 1

Third Party Knowledge of Possible Direct Benefits Based on Indirect Observation of Focal Animal's Current Behaviour to Someone Else

Knowledge Base (friend)
A watches what B (friend) is doing to F: If B → √ → F, then B → √ → A 100% of time If B → X → F, then B → √ → A 100% of time
Knowledge Base (semi-friend)
A watches what C (semi-friend) is doing to F: If C → √ → F, then C → √ → A 100% of time If C → X → F, then C → √ → A 70% of time and C → X → A 30% of time
Knowledge Base (non-friend)
A watches what D (non-friend) is doing to F: If D → √ → F, then D → √ → A 100% of time If D → X → F, then D → √ → A 20% of time and D → X → A 80% of time
Knowledge Base (enemy)
A watches what E (enemy) is doing to F: If E → √ → F, then E → √ → A 100% of time If E → X → F, then E → X → A 100% of time

Note: This is a kind of rudimentary classification system; here, the individual has a non-verbal "knowledge base" about each category of friend, semi-friend, non-friend, and enemy, which facilitates an appraisal of what is likely to happen. It implies knowledge about each of the individuals involved. Thus, you are less likely to approach an enemy for food because he will likely refuse you even if he has been seen feeding someone else. At one extreme, there is the friend (100% chance of feeding you if he fed someone else). At the other extreme, there is the enemy (0% chance). There are also two other situations (semi-friend, non-friend) which we can regard as representing two points along a continuum between the extremes. The X refers to hostile behaviour (e.g., attack). The check mark (√) refers to friendly behaviour (e.g., feed). Arrows indicate the direction of these behaviours. For example, B → √ → F means that individual B is friendly to individual F; and B → X → F means that B is hostile to F.

Employing an "if–then" syntax in social reasoning (de Waal, 2003), Table 1 (above) is a conjectural framework showing how a person (possibly also an animal) might employ this syntax when facing others who represent four different grades of social relationship: (1) a friend, (2) a semi-friend, (3) a non-friend, and (4) an enemy (all numbers are notional). These grades do not presume formal labels in the mind (cf. Spears, 2011). What these labels represent will, however, somehow have real-life denotation. Specifically, Table 1 shows a scenario where an animal begins to recognise differences in direct experience that *correlate* with observed third-party interactions. Observer A will know how every individual (friend, semi-friend, non-friend, enemy) is likely to behave towards A, after observing how these individuals behaved towards others. In all cases, in the table, friendliness begets friendliness. Differences arise in what happens after the conspecifics are observed being hostile towards others. The friend is the easiest to comprehend and trust, because the behaviour is friendly to A 100% of the time. The second easiest to comprehend is the enemy, who, if hostile to others, is hostile to A. The behaviour of the non-friend is less predictable: if hostile to others, the non-friend is usually — but not always — hostile to A. In the case of the semi-friend, there is also unpredictability: if hostile to others, the semi-friend is usually — but not always — friendly to A. The rate of hostility is low (only 30%), which means that the observer should regard friendliness as the default expectation. The if–then syntax is the basis upon which an observer develops an understanding about the correlations between direct reputational experience (e.g., how B behaved towards me) and indirect reputation (e.g., how B behaved towards others). The knowledge of this correlation is an impression inside the animal's mind, established during a personal history between the observer and the other animal (information by direct reputation), and intuitively cross-checked against the other's interaction with others (information by indirect reputation).

Thinking of conspecific behaviour in a probabilistic manner (as above) is useful because some kinds of information are important for survival, such as avoiding attack: hostile behaviour necessitates that the observer be vigilant. As Dunbar (1988) wrote, the "amount of visual monitoring that an animal does is primarily a function of its nervousness, and reflects the animal's need to keep track of the movements of the more dominant individuals in order to avoid being attacked unawares" (p. 115). Situations like this are where it is advantageous to have an additional channel of information (e.g., indirect reputation) in addition to personal encounter (direct reputation). As Sober and Wilson (1998) wrote: "two sources of evidence are better than one, as far as reliability is concerned" (p. 307). When facing the conspecifics, as presented in Table 1, the observer would be vigilant when facing the enemy and non-friend (because they might attack); and non-vigilant when facing the friend and semi-friend (because they are unlikely to attack). For everyone but the enemy, expectations are based on some degree of trust (cf. Kohn, 2009). It is beneficial for an observer to be non-vigilant most of the time, because this frees the observer's attention to focus on other things (e.g., feeding) [Dunbar, 1988]. Accuracy of assessment would be valuable here, enabling the observer to know when to relax, and with whom to associate. However, small sample sizes could create misleading impressions. To know that the semi-friend is friendly 70% of the time (as per Table 1), the observer should perhaps witness at least ten occurrences during which friendliness occurred seven times. If the observer has witnessed only one occurrence and it was hostile, a misleading impression has been formed — making the observer unnecessarily vigilant in the semi-friend's presence (an example where a "larger sample size," i.e., more encounters, would be useful). It seems clear that the direct and indirect experience would have unequal influence on that social impression. There is surely what Sober and Wilson (1998) call "D/I asymmetry": the direct (D) experience will likely be more reliable than indirect (I) experience. In gaining indirect knowledge, you may not have seen all of the relevant events that make an accurate impression; furthermore, the occurences are towards other people (not you). We can also make a note about the third form, gossip: it is a cheap form of information (cf. Smith and Harper, 2004) for a couple of reasons. Firstly, it is the most subject to distortion (you did not observe the events yourself, but learned about them through at least one other person's cognitive filters). Secondly, gossip is much easier to fake (because people lie) than information gained from observing behaviour directly, in particular when honest (difficult-to-fake) signals are being displayed (e.g., visual cues of health). So, comparing the different "channels," direct (D), indirect (I), and reported (R) reputation, in terms of value (e.g., reliability and accuracy), it is plausible that D > I > R.

Also, we should remember that our social impressions will be riddled with inaccuracies — a social version of the inaccuracies uncovered in decision theory (cf. Ayton, 2010). Furthermore, the social impression is likely highly distorted by emotional processing. As Schino and Aureli (2009) argue, cooperative behaviour amongst animals is likely mediated by a kind of "emotional bookkeeping" (rather than a rational and cognitive bookkeeping). This is likely in humans too (McElreath et al., 2003), because people generally are not rational actors maximizing their benefits without emotion (Gigerenzer, 1997; Simon, 1955, 1983; Sober and Wilson, 1998; Sutherland, 1992). Binmore (2005) proffered that emotion "evolved to help police primeval social contracts, and they remain useful to us for this purpose" (p. 83). A behaviourist interpretation of emotional bookkeeping is that organisms are motivated by emotional rewards and punishments that get associated with specific interactions with particular individuals (see discussion in Sober and Wilson, 1998, pp. 256–260). To me, this sounds like the basis for acknowledging at least a rudimentary form of social expertise in animals (cf. Helton, 2005), applicable to the concepts of reputation and reciprocity, direct and indirect.

References

Alexander, R. D. (1977). Natural selection and the analysis of human sociality. In C. E. Goulden (Ed.), *Changing scenes in the natural sciences, 1776–1976* (pp. 283–337). Philadelphia: Academy of Natural Sciences.
Alexander, R. D. (1987). *The biology of moral systems*. New York: Aldine de Gruyter.
Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, Massachusetts: Harvard University Press.
Aureli, F., Cozzolino, R., Cordisch, C., and Scucchi, S. (1992). Kin-oriented redirection among Japanese macaques: An expression of a revenge system? *Animal Behaviour*, *44*, 283–291.
Ayton, P. (2010). Judgement and decision making. In H. Kaye (Ed.), *Cognitive psychology* (second edition, pp. 417–453). Milton Keynes: The Open University.
Barnard, C. J., and Sibly, R. M. (1981). Producers and scroungers: A general model and its application to feeding flocks of house sparrows. *Animal Behaviour*, *29*, 543–550.
Binmore, K. (2005). *Natural justice*. Oxford: Oxford University Press.
Bourke, A. F. G. (2011). *Principles of social evolution*. Oxford: Oxford University Press.
Bowles, S., and Gintis, H. (2011). *A cooperative species: Human reciprocity and its evolution*. Princeton: Princeton University Press.
Brandt, H., and Sigmund, K. (2005). Indirect reciprocity, image scoring, and moral hazard. *Proceedings of the National Academy of the United States of America*, *102*, 2666–2670.
Bshary, R. (2002). Biting cleaner fish use altruism to deceive image-scoring client reef fish. *Proceedings of the Royal Society of London Series B: Biological Sciences*, *269*, 2087–2093.
Byrne, R. (2003). Tracing the evolutionary path of cognition. In M. Brüne, H. Ribbert, and W. Schiefenhövel (Eds.), *The social brain: Evolution and pathology* (pp. 43–60). Chichester: John Wiley & Sons.
Byrne, R., and Whiten, A. (Eds.). (1997). *Machiavellian intelligence II: Extensions and evaluations*. Cambridge: Cambridge University Press.
Call, J. (2002). Social knowledge and social manipulation in monkeys and apes. In C. S. Harcourt and B. R. Sherwood (Eds.), *New perspectives in primate evolution and behaviour* (pp. 173–196). Otley, West Yorkshire: Westbury.
Carmeli, A., and Tishler, A. (2005). Perceived organizational reputation and organizational performance: An empirical investigation of industrial enterprises. *Corporate Reputation Review*, *8*, 13–30.
Chekhov, A. (1921). A slander [C. Garnett, Trans.]. In *The tales of Chekhov, volume X: The horse stealers and other stories* (pp. 221–226). New York: The Macmillan Company (originally published 1883).
Clutton–Brock, T. (2009). Cooperation between non-kin in animal societies. *Nature*, *462*, 51–57.
Danchin, E., Giraldeau, L.-A., Valone, T. J., and Wagner, R. H. (2004). Public information: From nosy neighbours to cultural evolution. *Science*, *305*, 487–491.
Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, *31*, 169–193.
de Groot, A., and Gobet, F. (1996). *Perception and memory in chess. Heuristics of the professional eye.* Assen, Netherlands: Van Gorcum.
de Waal, F. B. M. (2003). Social syntax: The if–then structure of social problem solving. In F. B. M. de Waal (Ed.), *Animal social complexity* (pp. 230–248). Cambridge, Massachusetts: Harvard University Press.
Donald, M. (2001). *A mind so rare: The evolution of human consciousness*. New York: W. W. Norton & Company.
Dugatkin, L. A. (1997). *Cooperation among animals: An evolutionary perspective*. New York: Oxford University Press.
Dunbar, R. I. M. (1988). *Primate social systems*. London: Croon Helm.
Dunbar, R. I. M. (2003) The social brain: Mind, language, and society in evolutionary perspective. *Annual Review of Anthropology*, *33*, 163–181.
Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology*, *8*, 100–110.
Dunbar, R., Gamble, C., and Gowlett, J. (2014). *Thinking big: How the evolution of social life shaped the human mind.* New York: Thames & Hudson.
Emler, N. (1990). A social psychology of reputation. *European Review of Social Psychology*, *1*, 172–193.
Emler, N. (1994). Gossip, reputation, and social adaptation. In R. F. Goodman and A. Ben-Ze'ev (Eds.), *Good gossip* (pp. 117–138). Lawrence, Kansas: University Press of Kansas.

Engelmann, D., and Fischbacher, U. (2009). Indirect reciprocity and strategic reputation building in an experimental helping game. *Games and Economic Behavior*, *67*, 399–407.

Ericsson, K.A., Charness, N., Feltovich, P. J., and Hoffman, R. R. (Eds). (2006). *The Cambridge handbook of expertise and expert performance*. Cambridge: Cambridge University Press.

Faust, K. (2007). Very local structure in social networks. *Sociological Methodology*, *37*, 209–256.

Fehr, E., and Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, *14*, 159–181.

Gigerenzer, G. (1997). Bounded rationality: Models of fast and frugal inference. *Swiss Journal of Economics and Statistics*, *133*, 201–218.

Gobet, F., Chassy, P., and Bilalić, M. (2011). *Foundations of cognitive psychology*. London: McGraw–Hill.

Greif, A. (1989). Reputations and coalitions in medieval trade: Evidence on the Maghribi traders. *The Journal of Economic History*, *69*, 857–882.

Guttman, N. (1953). Operant conditioning, extinction, and periodic reinforcement in relation to concentration of sucrose used as reinforcing agent. *Journal of Experimental Psychology*, *46*, 213–224.

Hamilton, W. D. (1975). Innate social aptitudes of man: An approach from evolutionary genetics. In R. Fox (Ed.), *Biosocial anthropology* (pp. 133–153). London: Malaby Press.

Helton, W. S. (2005). Animal expertise, conscious or not. *Animal Cognition*, *8*, 67–74.

Hinde, R. A. (1976). Interactions, relationships, and social structure. *Man*, *11*, 1–17.

Humphrey, N. K. (1976). The social function of intellect. In P. G. Bateson and R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). Cambridge: Cambridge University Press.

Jensen, K., Call, J., and Tomasello, M. (2007). Chimpanzees are rational maximizers in an ultimatum game. *Science*, *318*, 107–109.

Jensen, K., Hare, B., Call, J., and Tomasello, M. (2006). What's in it for me? Self-regard precludes altruism and spite in chimpanzees. *Proceedings of the Royal Society of London Series B: Biological Sciences*, *273*, 1013–1021.

Kohn, M. (2009). *Trust: Self-interest and the common good*. Oxford: Oxford University Press.

Kolm, S.-C. (2000). The theory of reciprocity. In L.-A. Gérard–Varet, S.-C. Kolm, and J. Mercier Ythier (Eds.), *The economics of reciprocity, giving and altruism* (pp. 115–141). Basingstoke: Palgrave Macmillan.

Krebs, J. R., and Dawkins, R. (1984). Animal signals: Mind reading and manipulation. In J. R. Krebs and N. B. Davies (Eds.), *Behavioural ecology: An evolutionary approach* (second edition, pp. 380–402). Oxford: Blackwell.

Markl, H. (1985). Manipulation, modulation, information, cognition: Some of the riddles of communication. In B. Hölldobler and M. Lindauer (Eds.), *Experimental behavioral ecology and sociobiology* (pp. 163–194). Sunderland, Massachusetts: Sinauer.

McElreath, R., Clutton–Brock, T. H., Fehr, E., Fessler, D. M. T., Hagen, E. H., Hammerstein, P., Kosfeld, M., Milinski, M., Silk, J. B., Tooby, J., and Wilson, M. I. (2003). Group report: The role of cognition and emotion in cooperation. In P. Hammerstein (Ed.), *Genetic and cultural evolution of cooperation* (pp. 125–152). London: The MIT Press.

McGarty, C. (2002). Stereotype formation as category formation. In C. McGarty, V. Y. Yzerbyt, and R. Spears (Eds.), *Stereotypes as explanations: The formation of meaningful beliefs about social groups* (pp. 16–37). Cambridge: Cambridge University Press.

McGregor, P. K. (2005). *Animal communication networks*. Cambridge: Cambridge University Press.

Mori, A. (1983). An ethological study on chimpanzees at the artificial feeding place in the Mahale mountains, Tanzania — with special reference to the booming situation. *Primates*, *23*, 45–65.

Newell, B. R., and Shanks, D. R. (2014). Unconscious influences on decision making: A criticial review. *Behavioral and Brain Sciences*, *37*, 1–61.

Nowak, M., and Highfield, R. (2011). *Super cooperators*. Edinburgh: Canongate.

Nowak, M. A., and Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*, 573–577.

Obreiter, P., Fähnrich, S., Jens, N., and Gianluca, M. (2005). How social structure improves distributed reputation systems: Three hypotheses. *Lecture Notes in Computer Science*, *3601*, 229–236.

Ostrom, E. (2003). Toward a behavioral theory linking trust, reciprocity, and reputation. In E. Ostrom and J. Walker (Eds.), *Trust and reciprocity: Interdisciplinary lessons from experimental research* (pp. 19–79). New York: Russell Sage.

Parker, G. A., and Smith, J. M. (1990). Optimality theory in evolutionary biology. *Nature*, *348*, 27–33.
Pollock, G., and Dugatkin, L. A. (1992). Reciprocity and the emergence of reputation. *Journal of Theoretical Biology*, *159*, 25–37.
Pradel, J., and Fetchenhauer, D. (2010). Why most theories get it wrong: Altruistic intentions as an explanation of the evolution of genuine altruism. In U. J. Frey, C. Störmer, and K. P. Willführ (Eds.), *Homo novus – A human without illusions* (pp. 79–92). Berlin: Springer Verlag.
Russell, Y. I. (2007). *Reputations and polyadic interactions among great apes.* Doctoral dissertation, School of Biological Sciences, University of Liverpool.
Russell, Y. I. (2011). Prehistoric stone tools, chess expertise, and cognitive evolution: An experiment about recognizing features in flint debitage. *Journal of Evolutionary Psychology*, *9*, 249–269.
Russell, Y. I., Call, J., and Dunbar, R. I. M. (2008). Image scoring in great apes. *Behavioural Processes*, *78*, 108–111.
Russell, Y. I., and Phelps, S. (2013). How do you measure pleasure? A discussion about intrinsic costs and benefits in primate allogrooming. *Biology and Philosophy*, *28*, 1005–1020.
Schino, G., and Aureli, F. (2009). Reciprocal altruism in primates: Partner choice, cognition, and emotions. *Advances in the Study of Behaviour*, *39*, 45–69.
Schram, A. (2000). Sorting out the seeking: The economics of individual motivations. *Public Choice*, *103*, 231–258.
Schweizer, T. (1989). Economic individualism and the community spirit: Divergent orientation patterns of Javanese villagers in rice production and the ritual sphere. *Modern Asian Studies*, *23*, 277–312.
Sigmund, K. (2010). *The calculus of selfishness*. Princeton: Princeton University Press.
Simon, H. A. (1955). A behavioural model of rational choice. *Quarterly Journal of Economics*, *69*, 99–118.
Simon, H. A. (1983). *Reason in human affairs.* Stanford: Stanford University Press.
Smith, J. M., and Harper, D. (2004). *Animal signals*. Oxford: Oxford University Press.
Sober, E., and Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behavior.* Cambridge: Harvard University Press.
Spears, R. (2011). Group identities: The social identity perspective. In S. J. Schwartz, K. Luyckx, and V. L. Vignoles (Eds.)., *Handbook of identity theory and research* (pp. 201–224). New York: Springer.
Spencer, R., Russell, Y. I., Dickins, B. J. A., and Dickins, T. E. (2017). Kleptoparisitism in gulls *Laridae* at an urban and a coastal foraging environment: An assessment of ecological predictors. *Bird Study*, *64*, 12–19.
Sternberg, R. J. (1997). Cognitive conceptions of expertise. In P. Feltovich, K. M. Ford, and R. R. Hoffman (Eds.), *Expertise in context: Human and machine* (pp. 149–162). Menlo Park, California: AAAI/MIT Press.
Sutherland, S. (1992). *Irrationality*. London: Pinter & Martin.
Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, *20*, 410–433.
Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*, 675–691.
Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*, 35–57.
Urpelainen, J. (2011). The origins of social institutions. *Journal of Theoretical Politics*, *23*, 215–240.
Wedekind, C., and Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, *288*, 850–852.
Wicks, R. (2014). *Heathrow airport*. Yeovil, Somerset: Haynes.

The Journal of Mind and Behavior
Summer and Autumn 2016, Volume 37, Numbers 3 and 4
Pages 271–286
ISSN 0271–0137

A Non-Representational Understanding of Visual Experience

Kaplan Hasanoglu
Emmanuel College

This paper argues that various phenomenological considerations support a non-representational causal account of visual experience. This position claims that visual experiences serve as a non-representational causally efficacious medium for the production of beliefs concerning the external world. The arguments are centered on defending a non-representational causal account's understanding of the cognitive significance of visual experience. Among other things, such an account can easily explain the inextricable role that background beliefs and conceptual capacities play in perceptually-based external world belief-formation processes, the fact that visual mental states constrain beliefs because of their presentational phenomenology, and the phenomenon known as the transparency of visual experience.

Keywords: perceptual experience, hallucination, phenomenal properties

In this paper, I will show that certain phenomenological considerations support a *non-representational causal account* of visual experience.[1] This is a position in the philosophy of mind that claims that visual experience serves as a non-representational causally efficacious medium for the production of beliefs concerning the external world. So, it challenges both what Brewer (2006) calls *the content view,* which is the popular philosophical position that claims that visual experience is constituted by representational contents, as well as less fashionable alternatives such

---

Correspondence concerning this article should be addressed to Kaplan Hasanoglu, Ph.D., Philosophy Department, Emmanuel College, 400 The Fenway, ADM 357, Boston, Massachusetts 02115. Email: hasanogluk@emmanuel.edu

[1]Here and throughout I use the term "phenomenological" in the familiar sense that pertains to both *what it is like* to experience the world, as well as any related conclusions. Introspective reports are notoriously unreliable, given the way that they can be easily influenced by, say, folk psychology and other dubious theories. Therefore, in what follows I will methodically and self-consciously attempt to avoid — or, in Husserlian terminology, "bracket" — as many questionable theoretical commitments as possible. I am not alone in thinking that this kind of investigation is both still possible and potentially fruitful — see, for example, Gallagher (2005) and Thompson (2007).

as naïve realist (disjunctivist), sense-data, and adverbialist accounts. My aim is not to argue directly against any of these positions, but rather to present a non-representational causal account as another position worth taking seriously.

In what follows, when I talk of a perceptual experience's "cognitive significance," I will be referring solely to the role it plays in external world belief formation and sustenance, rather than in helping to render some external world beliefs justified and/or reasonable. That perceptual experience has the former role is fairly uncontroversial. For example, at the moment you have various (dispositional) beliefs about your surroundings, and it seems plausible to say that your current perceptual experiences play a crucial role in causing and sustaining those beliefs. A non-representational causal account understands this (roughly) as follows: in veridical cases, a non-representationally constituted internal mental state helps to produce a (reliably) true belief; whereas in non-veridical cases it obviously does not.

As a way of framing my task, let me mention and then somewhat abruptly set aside two possible objections. First, throughout this paper I will be unapologetically assuming the existence of reflectively accessible mental states as constituents of both veridical and hallucinatory experiences. That such states exist has famously come under repeated attack (making my assumption here the source of one objection). However, rightly or wrongly, the philosophical debate between content theorists, naïve realists (disjunctivists), sense-data theorists, adverbialists, etc., presupposes the existence of such states: the disagreement between them surrounds their status in hallucinatory cases, and/or the extent to which a proposed account of them in hallucinatory cases can be extended to the mental states involved in veridical cases. The purpose of this paper is to present a non-representational causal account as another horse in that particular race, so I will not here question the existence of such states. This may turn out to be a problematic theoretical commitment (see footnote 1 above); but given my aim, it is a necessary one.

Next, even if they do exist, it might appear prima facie problematic to posit such mental states as a non-representational causal intermediary involved in belief-formation, at least if we are also to grant that perceptual experiences play some role in *justifying* external world beliefs. Natural causes aren't reasons, after all, and in what follows I will only speak of experiences as causes. However, for one thing, though I can't defend the claim here, I agree with Shaun Gallagher that: "To have a belief is not to have an all-or-nothing mental representation, but to have some more-or-less-complete set of dispositions to act and to experience in certain ways" (2005, p. 214). So, in my view, what it means to say that a belief is an intentional entity is, at best, far from clear. If nothing else, this muddles the issue: What exactly is the problem here for a non-representational account supposed to be? And even if we ignore issues surrounding how to understand the intentionality of beliefs, it is also crucial to note that there is no current consensus in the philosophical literature regarding how to understand *justification,* generally speaking. This matters here because the stand one takes on, say, the internalist vs. externalist and/or the foundationalist vs. coherentist debate will have obvious implications for one's views about the justificatory role of perceptual experience. So-called *dogmatists* like Pryor (2000) and Heumer (2001) are (what I would consider to be) internalists who maintain that perceptual experiences provide prima facie justification for certain external world claims in virtue of their constitutive representational content. By contrast, Davidson (1983) famously defends a coherentist position that maintains that perceptual experiences cause but do not justify external world beliefs. Given such complexities, I have chosen to try to divide and conquer in this paper, and hence will leave off discussing perceptual experience's justificatory role for another occasion. Of course, punting on this issue might *still* seem unfair, since one reason for adopting a content theorist's position is its well-recognized theoretical elegance on this front — constitutive representational content makes the justificatory role of perceptual experience regarding beliefs with the same content a straightforward matter. But such support for a content theorist's view is obviously defeasible, and a complicated meta-methodological question surrounds how to weigh this kind of theoretical support against what I will argue below is recalcitrant phenomenological data. So, in sum, although it is fair to point out that a non-representational causal account owes us an explanation of the relationship between perceptual experience and justified beliefs about the external world, I would argue that it is unfair to rule it out initially based upon the presence of this explanatory gap. Indeed, if nothing else, the phenomenological investigation involved in what follows may turn out to be preparatory for filling that gap.

Phenomenology and the Cognitive Significance of Perceptual Experience

So, let me now clarify my aim. The phenomenological case I make for a non-representational causal account shall be centered, in particular, on defending its understanding of the role of reflectively accessible internal mental states in external world belief-formation, with a focus on visual experience. (I do think the view generalizes to other sense-modalities, but I will not defend that more general claim here.)

To motivate my position, it will help to begin by considering recognized visual perceptual illusions. In such cases, we are easily able to distinguish (i) what it means to take the experience at face value; (ii) what, all things considered, we take to be the case on the basis of the experience; and (iii) the inextricable role background beliefs play in having it be that (i) and (ii) come apart for such cases. Suppose that one is looking at a stick half-submerged in a glass of water. One natural (though not undisputed) thought is that if one were to take that experience at face value, one would take the stick in question to be broken in two. However, because one has background beliefs about (for example) the physics of light, one does not, all things considered, take the stick to be that way.

For cases like this, the distinction and interplay between (i)-(iii) seems relatively clear. What about more ordinary experiences? To help address this matter, I will now introduce some terminology. Take *the Good Case* to be a visual veridical case that is not an example of a veridical hallucination.[2] By contrast, *the Bad Case* should be understood to be a perfect (non-veridical) hallucination that is indistinguishable through reflection from some related Good Case. Finally, I will reserve the term *ordinary experience* for an experience of a typical human subject with relatively good vision that is either some Good Case involving unobstructed middle-sized objects in the viewed vicinity, or else the related Bad Case.

Consider, now, an ordinary experience of a nearby red sphere. I admit that by endorsing the claim that there is a red sphere, I *seem* to be doing nothing other than endorsing the experience itself, understood as an internal mental state, and in a way that intimately depends on how the mental state phenomenally presents things from my perspective. At the very least, this might seem to support the content view. However, against this, one should first acknowledge that if my background beliefs were to change significantly enough, a contrary claim would be endorsed upon enjoying a phenomenally identical experience. Examples that illustrate this are familiar: depending on my background beliefs, on the basis of the same internal visual state (here and henceforth understood as a structured complex of phenomenal properties), I might believe that I'm looking at a red wall, or else one that is white but bathed in red light. According to a non-representational causal account, the best explanation for this is that there is no one set of accuracy conditions concerning the external world that an experience can intrinsically call its own (see Travis, 2004, for an argument against the content view that seems to rest on a similar point).

What might be a typical example of such a background belief for ordinary experiences? Well, it is obviously part of my conception of spatiotemporal objects that they remain relatively stable through time: that is, we don't think of a mind-independent object as the kind of thing that will wink out of existence once we stop looking at and/or interacting with it. In other words, the size, shape, etc. of an object is what I will henceforth call an *experientially transcendent property.* Generally speaking, a property of an object x is an experientially transcendent property at time t — where t represents a time period during which we experience x — when, ceteris paribus, x also has that property at times other than t. Mind-independent objects are relatively stable, then, because their size, shape, etc. are experientially

---

[2] As the label suggests, veridical hallucinations are hallucinations that happen to be accurate, even though the cause of the perceptual experience is not the relevant object that is in fact in the perceiver's surroundings. For (a well-worn) example, I may hallucinate a dagger before me while there just so happens to be a dagger of that very same sort before me. For useful discussion, see Johnston (2004). For discussion of the related notion of veridical illusions, see Johnston (2006).

transcendent properties. According to a non-representational causal account, for cognitively sophisticated adults, beliefs about experientially transcendent properties are thus the kind of background beliefs that help explain why we believe what we do when we experience things like red spheres. On this view, the visual state does not itself present a red sphere, and it is, rather, background beliefs such as those concerning experientially transcendent properties that pick up the slack in the relevant belief-formation processes. Perceptual experience, properly construed, ought to be relegated to a mere non-representational causal role in that process.[3]

However, rival positions such as the content view can certainly accommodate the role that background beliefs play in such belief-formation processes.[4] Moreover, one might think that a non-representational causal account should be rejected on straightforward phenomenological grounds, since it seems unable to explain the way that perceptual mental states constrain belief. So, for example, it is fairly obvious that the mental state I'm enjoying at the moment constrains my beliefs in many ways — for one thing, it prevents me from believing that I'm walking down a busy street. Crucially, it evidently does this because of its *presentational phenomenology* — the mind-independent world is presented *by the mental state* I'm enjoying right now as being a way that my current (dispositional) beliefs reflect, and this feature seems intrinsic to it. (For discussion of this, see Siegel, 2010, chapter 2). Obviously, then, in order to defend it I need to provide an explanation of how internal mental states constrain beliefs on behalf of a non-representational causal account.

To do so, I will now argue that for a typical human subject who regularly enjoys Good Cases, the mental state involved in the Good Cases — to introduce one last piece of terminology — is *specific-object-involving*. A visual state is specific-object-involving when and only when one enjoys it while causally interacting in a non-deviant way with the mind-independent object(s) that play(s) an essential role in helping to create it. The idea here is that, since it is a Good Case, the mind-independent objects that are actually before you play an essential and non-deviant role in helping to create your current token visual state, similar to how the actions of a painter

[3]By contrast, in an influential paper, Horgan and Tienson (2002) defend a position that evidently would involve denying that background beliefs are relevant as to why we take certain experiences to be of the external world. They maintain, or so it appears, that intentional content involving the external world is intrinsic to the phenomenal properties of the experience of, say, a red sphere in one's immediate vicinity. But, as far as I can tell, the force of their arguments rests on introspection alone, and hence can be straightforwardly undermined by the considerations offered throughout this paper. In short, I would argue that Horgan and Tienson simply gloss over the implications of the obvious role that background beliefs, capacities, etc. play in mediating the relevant belief-formation processes, and hence in helping to explain the cognitive significance of the relevant perceptual states.

[4]To defend their view in light of the above kind of holism, the proponents of the content view might draw an analogy with belief. After all, even if one belief by itself doesn't fix what will be justified and/or endorsed upon maintaining it, this alone need not mean that the belief doesn't have determinate content.

are essential to helping to create a masterpiece. So, the visual state you are enjoying at this moment is specific-object-involving, since it exists now only because the objects right in front of you that have helped to create it also exist.[5] This is of course not to deny that such a mental state might have been created some other way. Instead, it is a non-counterfactual claim about what is in fact essential for creating your current token mental state.

I will now spend a fair bit of time offering various qualifications and developments of this idea, including laying out some of its more important implications. Once all of that is in place, I will be in position to provide a non-representational causal account's understanding of the way in which visual experiences constrain beliefs. First, the claim here should not be confused with a claim about what *constitutes* the mental state. To give a helpful analogy, consider that your parents played an essential role in your creation. But although (more specifically) their past actions were essential to creating you, neither your parents nor their actions are literally a part of you. In like fashion, on this view, although they are essential in helping to create them, mind-independent objects should not be thought of as constitutive of the internal visual states enjoyed in Good Cases.

Next, maintaining this is quite compatible with the so-called *abstractness* of perception noted, for example, in Tye (1995) — the fact that in other Good Cases numerically distinct objects can help to produce numerically distinct mental states of ostensibly the same phenomenal type. Since each *token* mental state's phenomenal properties are created, in part, by the mind-independent objects present, all that follows is that any identity of phenomenal *type* will itself be determined, albeit only in part, by the similarities of the distinct, specifically perceived, mind-independent objects themselves. So, right now the experience I'm having is specific-object-involving simply because this particular computer that I am now interacting with is playing an essential role in creating its token phenomenal properties. In other words, my current visual state is *this-computer-involving*. All that these latest considerations require me to go on to admit, then, is that the computer in front of me thereby creates a token mental state whose intrinsic phenomenal type is identical with the type of any numerically distinct mental state produced, in part, by the perception of any other sufficiently similar computer.

If possible, as a phenomenological exercise, it would be helpful to now examine one's own visual experience and try this last sort of assertion on for size ("My current visual experience is *this-X-involving*"; where "X" refers to some unobstructed middle-sized object in the viewed vicinity). My prediction is that it will strike one as a very natural assertion to make. On the other hand — to now lay my cards on the table — if the reader finds that there is absolutely nothing to such a characterization, then the arguments that follow will have little force.

---

[5]Here and throughout I have assumed that there is no causal over-determination involved in what in fact creates the mental state. There would be if, for example, it was a veridical hallucination.

Next, and quite crucially, matters are not all that different when it comes to the mental state involved in the Bad Case. You are probably not vividly hallucinating anything right now, but on some other occasion you might. How does a non-representational causal account treat such cases? In a word, that mental state is *derivatively* specific-object-involving. To get at what this means, notice first that in order for the mental state in question to be a perfect hallucination it has to *seem real;* that is, it has to be an experience that is taken to be veridical by the subject suffering from it at the time in question. But that means that *for those who regularly enjoy Good Cases*, the hallucination has to have the same general sort of phenomenal properties as the mental states that were, in fact, created in part by the presence of mind-independent objects the subject has in fact already encountered. It is only a perfect hallucination because it has an intrinsic phenomenal character of a basic sort already possessed by the mental states involved in the already experienced Good Cases. Intuitively, for those who regularly enjoy Good Cases, it is hard to understand how one could have a perfect hallucination — one indiscriminable through reflection from some related Good Case — that instantiates intrinsic phenomenal properties substantially different from the basic type of phenomenal properties created by Good Cases on other, previous occasions. Suppose, then, I am suffering from a perfect hallucination of a computer. In that case, the mental state can't be this-computer-involving in the sense mentioned above, simply because I am not causally interacting with any computer. However, since the hallucination would seem real, whatever caused it to come about would have to generate phenomenology that copied my normal, everyday experiences of the external world, including things like computers. It is in that sense that the mental state in question would be derivatively specific-object-involving.

To clarify, perhaps another analogy can help. For, although it shares observable properties with the original, you can only understand a perfect forgery of some painting in a similar manner. Dali painted *The Persistence of Memory,* but he did not paint a certain perfect forgery of it. And yet, by virtue of its observable properties, a certain painting is in fact a perfect forgery of *The Persistence of Memory* because of its derivative relationship with Dali's original. That is, we have to reference Dali's actions with respect to the original, if we are to understand why it is correct to describe a painting with certain observable properties as a perfect forgery of *The Persistence of Memory,* rather than as (of) some other painting. With important qualifications, hallucinations can be understood similarly. Admittedly, no analogous ontological priority is involved here. That is, we arguably don't need to first experience an object in order to hallucinate something involving it — fanciful dreams of fanciful things are common enough. So, plausibly, we can have perfect hallucinations involving objects we've never actually seen before. And even in cases where the requirement for prior perception would make sense — say, a vivid dream involving one of my brothers — the situation hallucinated may be novel. I may dream of Taner doing something I've never seen him do before, wearing clothes I've never seen him wear before, etc. Similarly, I may dream of a banana or another commonly perceived object as seen from a novel angle, in weird lighting, etc. Nevertheless, perfect hallucinations of novel situations and/or of never before seen objects are also derivatively specific-object-involving. No mere smattering of paint on a canvas is a forgery; similarly, no mere play of sensations is a perfect hallucination. In order for a painting to be a perfect forgery, it must be a convincing reproduction. Similarly, in order for an experience to be a perfect hallucination, the subject must take it to be veridical at the time she experiences it. *For subjects who regularly enjoy Good Cases*, this means that in order to understand what makes an experience a perfect hallucination, we must first understand the basic phenomenal character of the mental states involved in their prior everyday veridical experiences. So, if I am going to have a perfect hallucination of a banana (or else of some object I've never actually seen before), then even if the situation, perspective, etc. is novel, it still must be a mental state that shares a basic phenomenal character with the mental states involved in the actual situations I *have* already experienced. In particular, the hallucinated object(s) must still seem to be behaving in generally believable ways. For example, you can't have a perfect hallucination of objects that, ceteris paribus, unpredictably flick in and out of existence, spontaneously melt into one another, etc. If I was suffering from a hallucination where things were too weird or otherwise off in some fundamental sense, then assuming I was in an otherwise normal state of mind, I wouldn't take the experience to be veridical. Instead I would probably think something like: "This is too weird, I must be hallucinating." Therefore, it would not be a perfect hallucination. In sum, for those who regularly enjoy Good Cases, it is the prior experiences they have had of actual environing mind-independent objects that first determine what it means for things to seem real or not. As a result, it will be its fundamental relationship to those same prior experiences which will determine whether or not my novel dream of a banana, my brother, or even some object I've never seen before, etc., counts as a perfect hallucination. That is what makes even perfect hallucinations involving novel situations and/or never before seen objects derivatively specific-object-involving.

Interestingly, this all relates in important ways to the phenomenon that is often called the *transparency* or *diaphanousness* of perceptual experience, which is something usually thought to support a certain version of the content view. (For discussions of transparency, see Dretske, 1995, p. 62; Harman, 1990, p. 39; Moore, 1903, p. 450; and Tye, 2000, pp. 51-52). Some philosophers have argued that when we reflect on the subjective properties of our perceptual experiences, all we seem to find is what our experiences present as being the case in the mind-independent world. To focus on what it is to subjectively experience blue, for example, is just to focus on the blue thing that one's experience is presenting as being in front of one. What this is typically thought to show is the non-existence of non-representational phenomenal properties; that instead our experiences only have (sense-modality-specific) representational properties (see Harman, 1990; Tye, 1995, 2000). But one might also employ transparency as a means for claiming, on phenomenological grounds, something more general: that the experience itself has representational content that bears upon the external world. The claim would be that were we to try to find non-representational features of our experiences (such as their non-representational causal features), all we would end up noticing is how they present mind-independent reality.

However, even granting that perceptual experience is transparent, an entirely separate question is how we should interpret this phenomenon (see Stoljar, 2004). Indeed, the proponent of a non-representational causal account can explain transparency as follows. In the Good Case or its related Bad Case, when we examine our internal mental states we only find the world because that mental state in question is one that makes things seem real, where *for those who have regularly experienced Good Cases* this means that the mental state in question has the general sort of non-representational phenomenal properties created by the relevant mind-independent objects one has already experienced. On this view, then, transparency (along with the property of seeming real) is not something intrinsic to an experience, but is rather determined (in part) by the prior environing objects encountered as one has more or less successfully made one's way through the world.

This allows me to respond to another possible objection. I have in mind a worry that derives from the recent work of William Fish (2009). Fish is a naïve realist. He maintains that the phenomenal character of veridical experience is constituted by a subject's acquaintance with the properties of the experienced object(s) [Fish, 2009, p. 14]. So, for example, when I have a veridical experience of a red tomato, Fish would claim that the phenomenal redness that constitutes the mental state is a property of the tomato that is actually before me. Good Cases thus involve an "irreducible" mental relation with certain mind-independent objects (Fish, 2009, p.14, n.19). Of course, perfect hallucinations would involve no such relation. To account for them, Fish thus argues rather strikingly that perfect hallucinations lack phenomenal character altogether (p. 93). The reason Fish makes this bold move, it seems, is because on his view phenomenal character is something constituted by the above-described, irreducible mental relation. And so, the reasoning seems to go, since that mental relation is lacking in perfect hallucinations, phenomenal character must also be lacking (Schellenberg, 2013, p. 50). Fish admits that perfect hallucinations seem real. But, he argues, this is because of their "cognitive effects," rather than their phenomenal character (p. 94). In particular, they seem real only because they produce "the same beliefs or judgments that a veridical perception of that kind would have produced" (p. 94). For this reason, he would obviously deny that perfect hallucinations seem real because of their derivative phenomenal character.

Fish's fully developed views are complex, and his arguments are characteristically sophisticated and subtle. They are also stated within the context of the rather involved debate over naïve realism. Suffice it to say, then, that I cannot give Fish a fair treatment here. In any case, there is a simple and glaring problem for him. As Susanna Siegel bluntly puts it in one of her responses to Fish: "The idea that hallucinations lack phenomenal character is at odds with the crudest deliverances of introspection" (Siegel, 2010, p. 49, n. 19; for a more detailed critique, see Siegel, 2008. See also Martin, 2013; Pautz, 2013; and Schellenberg, 2013). As I would put it, a very basic phenomenological investigation supports the claim that perfect hallucinations possess phenomenal character. As a result, one would need extremely compelling reasons for denying that hallucinations actually possess such character.

I don't think that Fish provides such reasons. I have argued that Good Cases are specific-object-involving, but in a sense that amounts to a non-deviant *causal relationship* rather than a constitutive and "irreducible" relationship of acquaintance between the mental state and the object(s) before one. I am no naïve realist. And *if,* in order to maintain naïve realism, one must also deny that perfect hallucinations have phenomenal character, then I would argue that that is too high of a price to pay. (Whether this conditional actually holds is not something I can properly address here.) In sum, arguably, my position is superior to Fish's in the following way: I can explain a sense in which the mental states in Good Cases are world-involving, while also acknowledging that perfect hallucinations possess phenomenal character.

There is one final important implication of my overall position that is worth discussing at length. To get at it, we should note that one ubiquitous example of an object involved in Good Cases will be one's physical body — though, of course, one's body is special in that it is an object that one *lives through* rather than experiences at a distance. I mention this here because, since one's lived physical body is a unique item situated in the universe, a non-representational causal account is thus nicely poised to provide a way of *individuating* visual mental states. To individuate something is to provide an adequate account of what makes it unique. As I will understand it, such an account must have an actual and a counterfactual component.[6] It must both provide a true description of just the mental state in question *and* also fail to allow for (nearby?) possible worlds where there is more than one mental state that satisfies that same description (though there can be possible worlds where a different mental state satisfies the description). The first requirement assures that the properties mentioned in the individuating account are uniquely true of the mental state in question; the second assures — to a degree arguably suitable for my purposes — that the fact that those properties *uniquely* pick it out in the actual world is not an accident. (There are potentially going to be many different ways to individuate mental states in this way.)

[6]I am not sure how to understand individuation, generally speaking. However, here I am only concerned with individuating mental states adequately enough for the purposes of understanding the cognitive significance of perceptual experience. For that project, I would argue that the account offered here suffices.

To consider a different sort of case, on this view one adequate way to individuate Barack Obama is as 44th President of the United States. This is because he is in fact the only one that held that office *and* there are no (nearby?) possible worlds where there is more than one 44th President of the United States, given how we now understand that institution. (And, as predicted, there are other ways to individuate Barack Obama. For example, one might cite specific enough aspects of his unique personal history.)

So, as far as any visual states enjoyed in Good Cases are concerned, my claim here is that one way to individuate them involves citing the mind-independent objects (including one's body) that have helped to create them. To get at why, it will be useful to explore some related yet inadequate attempts to individuate such states. For example, notice that if you were to merely focus on a certain visual state's intrinsic phenomenal properties as a way of individuating it, mentioning nothing about its etiology, the account would be too coarse-grained: you would here run the risk of wrongly identifying it with a numerically distinct but phenomenally type-identical experience that is or might be enjoyed by some subject (including oneself on a different occasion). In other words, there is a possible world (that is perhaps the actual world) where there is more than one visual state with those same phenomenal properties. Merely adding the actual time that you enjoyed the state runs into similar problems: here you would run the risk of misidentifying it with a numerically distinct phenomenally type-identical experience that is or might be enjoyed by someone else at that particular time. That is, there is a possible world (that may be the actual world) where there is more than one visual state with those phenomenal properties enjoyed at that time. This suggests that one way to make the account sufficiently fine-grained is to add an adequately specified *place* where it is enjoyed. So, you might try to home in on the subject's unique place and time where he enjoys the state. My claim, then, is that one way of adequately specifying such a place and time is by citing the specific environing objects that have helped to create the mental state, including their precise location in relation to the subject's body. As long as this citation is specific enough in this way, referring to the specific object(s) rather than object type(s), it seems it will allow us to adequately individuate the mental state in question. There is no one other than myself (the body that is currently enjoying this visual state) viewing this particular computer from this particular angle relative to my body. And it certainly stretches the imagination to consider a case where there is more than one such visual state enjoyed at a certain time by an embodied subject like myself. Roughly, then, I would argue that an adequate principle of individuation for visual states enjoyed in Good Cases seems to be this: at any time t, visual states x and y are identical if and only if they have the same intrinsic phenomenal properties and the same mind-independent etiology (where we suitably specify that etiology according to one's unique bodily-oriented perspective). Notice, then, that on this view subjects involved in Good Cases can individuate their own mental states by making use of this principle, through a fairly unsophisticated phenomenological exercise like the one rehearsed a few sentences ago.

Typically, in discussions of perceptual experience, mind-independent objects are thought to be only extrinsically related to internal mental states, as an aspect of their causal ancestry. The object is thus not thought of as something given in the experience itself, since the latter is understood as a relatively independent downstream event in the overall process. This common way of looking at visual experience thus treats a mental state enjoyed in a Good Case, considered as such, as on an ontological (and perhaps also epistemological) par with a phenomenally indistinguishable hallucination. It is a way of looking at perceptual experience that is also friendly to the methodology enshrined by the Argument from Hallucination, where one first examines perfect hallucinatory cases and then attempts to spread that understanding to veridical cases, as well. However, according to the account defended here, we cannot rely on this outlook or this methodology if we wish to understand the cognitive significance of perceptual experience. If that is our aim, then we must acknowledge that in Good Cases visual mental states are merely one component of a larger process; a process that also involves background beliefs *and* the mind-independent objects (including one's physical body) partly responsible for creating the state. To say that mind-independent objects are only extrinsically related to these internal mental states is thus misleading, since it ignores the absolutely essential role that such objects play in helping to create the token mental state involved in a Good Case, and also the role they can play in such a case in an attempt to individuate one's own visual states through a phenomenological exercise. It also, for that matter, fails to take into account the derivative status of the mental states involved in Bad Cases. On the contrary, then, it is not despite but *because* a mind-independent object (including one's body) plays the role that it does in an internal mental state's causal ancestry in Good Cases, that we must include the former as essential to understanding the latter. So, since my current mental state is this-computer-involving, it follows that it is the position of the computer that is actually in front of my physical body at the moment that helps to determine the phenomenal properties of my current internal mental state; moreover, to the extent that this computer's position is something I am able to control through bodily action, that mental state likewise determines the position of the computer. To deny the computer's essential role in making my current mental state what it is, therefore, would be just as absurd as denying that my parents played an essential role in creating me, or that a painter played an essential role in creating a masterpiece, etc. Even though mind-independent objects are not, on this view, literally parts of the visual states that they help to create, since they are essential to their creation and also something we can cite to individuate such states via a phenomenological exercise, we cannot and should not ignore them when attempting to understand the latter; or, at least, we cannot and should not if our goal is to understand the cognitive significance of visual experience.

The phenomenological linchpin that brings these ideas home is this: when you are involved in a Good Case you simply cannot refer to an experience *as some isolated internal mental event or state* except by way of a mental act which *abstracts from* the lived, contextualized, body-and-world-involving process that is the Good Case as a whole, understood as a kind of successful dance with various immediately environing objects. According to a non-representational causal account, an act of abstraction like this is precisely what I engaged in above, when I recognized that my current mental state is this-computer-involving.

Now for the punchline: according to a non-representational causal account, the mental states involved in veridical experiences and perfect hallucinations constrain beliefs in the way they do simply because those mental states are specific-object-involving, and derivatively specific-object-involving, respectively. As a result, whatever specific mind-independent object is involved in helping to create a token mental state of a certain type (understood as a structured complex of phenomenal properties) will obviously thereby determine what beliefs result, albeit relative to a fixed set of background beliefs, capacities, etc. The mental state that is (derivatively) this-computer-involving will thereby not produce a belief that I am looking at a lion, relative to a fixed set of background beliefs, capacities, etc., simply because computers and lions are very different kinds of objects that consequently thereby help to produce very different structured complexes of phenomenal properties. In the Good Cases involving computers or lions, the different way in which such mental states constrain my beliefs, therefore, rests on nothing other than the mind-independent differences between computers and lions. The same is true in the Bad Cases derivatively involving computers or lions. In other words, a hallucination that is derivatively lion-involving will not make me believe that a computer is present relative to my current fixed set of background beliefs, capacities, etc. simply because the lions that I have already experienced have been different enough in a mind-independent sense from the computers I have already experienced, and hence have helped to produce the appropriately different structured complexes of phenomenal properties that thereby help to produce the appropriately different beliefs.

However, it must be acknowledged that being an object such as a computer is not simply a matter of having, say, a certain mind-independent molecular structure, but is also a function of the object in question's (socially reinforced) role in our practices. So, if we just so happened to give a certain mind-independent swarm of atoms a different role in our practices (say, as an object of worship rather than as a computing device), it would of course no longer be a computer. One might say, then, that the computer before me is only a computer because we represent it as a computer.

A similar point has been forcefully made by McDowell, who famously uses it to defend a certain version of the content view. In *Mind and World* he argues that we only experience the world as we do because we already possess and bring to bear various conceptual capacities. So, on his view, we can experience something as a computer only because we have such capacities already in place; capacities which (as I understand McDowell) have already holistically situated what it means to be a computer in a way that allows for the experience to rationally link up with other beliefs, experiences, actions, etc. As McDowell writes:

> By virtue of the way in which the conceptual capacities that are drawn into operation in an experience are rationally linked into the whole [conceptual] network… the subject of the experience understands what the experience takes in (or at least seems to take in) as part of a wider reality, a reality that is all embraceable in thought but not all available to this experience. (1996, pp. 31–32)

However, although it is quite correct that being a computer is in part a function of this kind of representation (and hence of what conceptual capacities we bring to bear when we visually perceive the world), this can be straightforwardly handled by a non-representational account in the manner already discussed. Namely, one can claim that what is already in place and hence thereby contributes to that representing is itself a function of our background beliefs, conceptual capacities, etc. rather than something intrinsic to the mental state.

To reinforce this last point and connect it up with the latest worry, suppose that I gain the concept of a Douglas fir.[7] Then suppose I subsequently see one in a set of circumstances (lighting, level of sobriety, etc.) that are otherwise identical with some set of prior circumstances in which I was in the presence of the same tree, but where I didn't possess the concept. Suppose, next, that (partly) on the basis of the latest experience I believe that a Douglas fir is present, but that I didn't do so on the basis of the previous experience. There are two possibilities here: either the later mental state's intrinsic phenomenal properties are type-identical with those that constituted the earlier mental state, or they are not. Consider the first option. In that case, one seems free to maintain that it is the very same type of perceptual mental state involved in both cases. (Why can't we slice things up in this way?) The proponents of a non-representational causal account seem free, then, to give that common type of mental state a non-representational gloss. They could maintain that since the intrinsic phenomenal properties did not change, the type of perceptual mental state involved in the relevant belief-formation process also did not change, even after I gained the concept of a Douglas fir. Instead, it is only the relevant background beliefs, capacities, etc. that are now different. According to a non-representational causal account, then, it would be the latter difference alone which would explain why I now believe that a Douglas fir is present, whereas before I didn't.

Similarly, even if for the sake of argument we granted that the later experience of the Douglas fir was constituted by different phenomenal properties, it still would be far too hasty to maintain that that change in particular is what fundamentally explains the difference in cognitive significance. Given that background

[7]This example could obviously be adapted to apply to the computer vs. object of worship case, *mutatis mutandis*. In particular, rather than talking about gaining a concept, we could talk about an alteration of the relevant conceptual capacity that I bring to bear in perceiving the swarm of atoms now identified as a computer.

beliefs, capacities, etc. inextricably mediate the belief-formation process, another possibility is that as before it is only a change in such beliefs, capacities, etc. that explains why I now believe that a Douglas fir is present, whereas before I didn't. In other words, it may well be that the change in phenomenal properties is simply explanatorily irrelevant.

More generally speaking, there is little doubt that changing what we believe and how we think (including the gaining/losing of conceptual capacities and related dispositions) can change how things perceptually seem to us — that there is so-called *cognitive penetration* involved here. But, if nothing else, there is still a debate to be had regarding what precise role such cognitive penetration should play in any attempt to explain the cognitive significance of perceptual experience.[8]

Finally, a related point that should be acknowledged is that when we see a computer before us, we do so via a sense modality that makes its own causal contribution to the process. As a result, in Good Cases, *how* a computer appears to us will, in part, be a function of peculiarities related to how our brain and eyes work. But it still can be true in Good Cases that *what* we see are mind-independent items before us that themselves have relational properties like *looking a certain way to a certain subject in a certain circumstance*, and that these relational properties of that mind-independent object *also* help to determine how we causally interact with it in a non-deviant fashion via that particular sense modality. So, when in the Good Case I see and point at a computer that is actually in front of me and say something like "This computer looks (is) black," I am pointing at a mind-independent object that looks the way it does in part because of how my brain and eyes work. And I call it a *computer* because of the particular role that that swarm of atoms plays in our practices. But because it is a mind-independent swarm of atoms that looks that way and plays that role (rather than something mind-dependent, like a dream or an after-image), it is nothing other than a mind-independent object which also helps to produce my current internal mental state.

[8]Recent work by Fish and Johnston also plausibly accommodates the role that conceptual capacities play in shaping the presentational phenomenology of perceptual experience, without espousing the existence of conceptual content. Reviewing the details here would be rather involved, and thus would perhaps take us too far afield. In any case, if nothing else the prominence of Fish's and Johnston's work alone shows that there is at least a debate to be had here, and hence that the above, McDowell-inspired objection is not decisive. See Fish (2009, pp. 67–74) and Johnston (2006, pp. 282–285). The point I am making above also helps to illustrate, I think, why the Burgean externalist views on mental content cannot be of assistance to the proponent of the content view. After all, Burge's thought experiments can be construed as supporting the idea that one's conceptual capacities are partly constituted by one's environment. But, one could of course admit that and just say that what the external environment thereby helps to constitute is the conceptual capacities that enter into the relevant belief-formation processes, rather than the content of the experience as such. Unfortunately, I lack the space to address the argument for perceptual content found in Burge (2010).

Conclusion

My hope is that the above discussion has shown how a non-representational causal account can readily explain various aspects of our visual experiences. However, this does not by itself show a non-representational causal account to be correct, or even the most plausible view to maintain. Making that case would require showing it to be the *best* explanation of such phenomena, preferably along with other basic features of perceptual experience. Although I have intimated at times how it might proceed, that project must be reserved for another occasion.

References

Brewer, B. (2006). Perception and content. *The European Journal of Philosophy, 14*, 165–181.
Burge, T. (2010). *Origins of objectivity*. New York: Oxford University Press.
Davidson, D. (1983). A coherence theory of truth and knowledge. In D. Henrich (Ed.), *Kant oder Hegel?* (pp. 423–438). Stuttgart: Klett–Cotta.
Dretske, F. (1995). *Naturalizing the mind.* Cambridge, Massachusetts: MIT Press.
Fish, W. (2009). *Perception, hallucination, and illusion.* New York: Oxford University Press.
Gallagher, S. (2005). *How the body shapes the mind.* Oxford: Oxford University Press.
Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives, 4*, 31–52.
Huemer, M. (2001). *Skepticism and the veil of perception.* Lanham, Maryland: Rowman and Littlefield.
Horgan, T., and Tienson, J. (2002). The intentionality of phenomenology and the phenomenology of intentionality. In D. Chalmers (Ed.), *Philosophy of mind: Classical and contemporary readings* (pp. 520–533). New York: Oxford University Press.
Johnston, M. (2004). The obscure object of hallucination. *Philosophical Studies, 120*, 113–183.
Johnston, M. (2006). Better than mere knowledge? The function of sensory awareness. In T. S. Gendler and J. Hawthorne (Eds.), *Perceptual experience* (pp. 260–290). Oxford: Oxford University Press.
Martin, M. G. F. (2013). Shibboleth: Some comments on William Fish's *Perception, Hallucination & Illusion*. *Philosophical Studies, 163,* 37–48.
McDowell, J. (1996). *Mind and world.* Cambridge, Massachusetts: Harvard University Press.
Moore, G. E. (1903). The refutation of idealism. *Mind, 12*, 433–453.
Pautz, A. (2013). Do the benefits of naïve realism outweigh the costs? Comments on Fish, *Perception, Hallucination and Illusion*. *Philosophical Studies, 163,* 25–36.
Pryor, J. (2000). The skeptic and the dogmatist. *Nous, 34*, 517–549.
Schellenberg, S. (2013) Perceptual content and relations. *Philosophical Studies, 163,* 49–55.
Siegel, S. (2008). The epistemic conception of hallucination. In A. Haddock and F. Macpherson (Eds.), *Disjunctivism: Perception, action and knowledge* (pp. 205–224). Oxford: Oxford University Press.
Siegel, S. (2010). *The contents of visual experience.* Oxford: Oxford University Press.
Stoljar, D. (2004). The argument from diaphanousness. In M. Escurdia, R. Stainton, and C. Viger (Eds.), *Language, mind, and world: Special issue of the Canadian Journal of Philosophy* (pp. 341–390). Edmonton, Canada: University of Alberta Press.
Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of the mind.* Cambridge, Massachusetts: Belknap Press.
Travis, C. (2004). The silence of the senses. *Mind, 113*, 57–94.
Tye, M. (1995). *Ten problems of consciousness*. Cambridge, Massachusetts: MIT Press.
Tye, M. (2000). *Consciousness, color, and content*. Cambridge, Massachusetts: MIT Press.

The Journal of Mind and Behavior
Summer and Autumn 2016, Volume 37, Numbers 3 and 4
Pages 287–310
ISSN 0271–0137

What Does Neuroscience Research Tell Us about Human Consciousness? An Overview of Benjamin Libet's Legacy

Jimmy Y. Zhong

Georgia Institute of Technology

This paper presents an overview of the key neuroscience studies investigating the neural mechanisms of self-initiated movements that form the basis of our human consciousness. These studies, which commenced with the seminal works of Benjamin Libet and colleagues, showed that an ensemble of brain areas — localized to the frontal and medial regions of the brain — are involved in engendering the conscious decision to commit a motor act. Regardless of differences in neuroimaging techniques, these studies commonly showed that early neuronal activities in the frontal lobules and supplementary motor areas, interpreted by some to be reflective of unconscious processes, occurred before one was conscious of the intention to act as well as of the act itself. I examine and discuss these empirical findings with regard to the need to analyze the contents and stages of awareness, and devise paradigm-specific models or theories that could account for inconsistent findings garnered from different experimental paradigms. This paper concludes by emphasizing a need to reconcile the principles of determinism with the notions of free will in future development of consciousness research and theories.

Keywords: consciousness, awareness, prefrontal cortex, supplementary motor area

"Consciousness" is a term that is hard to define. In the simplest sense, it pertains to the subjective state of sentience or awareness that accompanies us throughout the day whenever we are awake and performing our daily tasks (Searle, 1992, 1993). At a higher or more intricate level, it is an essential mental phenomenon

Parts of the review were written for a presentation of the Libet-type experiments in a seminar on neuroethics conducted by Scott D. Moffat (Georgia Tech). I thank him for highlighting the mind–body problem in consciousness research and several important empirical studies. I also thank two anonymous reviewers who offered advice for the relevance of important philosophical concepts to a well-rounded discussion of the experimental results attained by Libet and colleagues. Correspondence concerning this article should be sent to Jimmy Y. Zhong, Cognitive Neuroscience of Aging Lab, Center for Advanced Brain Imaging (CABI) 135, 831 Marietta Street NW, Atlanta, Georgia 30318. Email: jzhong34@gatech.edu

that applies to both humans and the vast majority of living creatures, creating a platform for the emergence of higher mental faculties like attention, perception, cognition, and memory. Every day we experience multitudes of perceptual sensations, subjective feelings, and streams of thoughts such as hearing the ringing of our alarm clocks, realizing the urgency to get up to go to work, deliberating about the tasks that await us, etc. How the brain makes sense of all these different experiences and bind them into a unified conscious state based on physical/neural processes has been designated as the "hard problem" of consciousness (Chalmers, 1995). Complementing this "hard problem" are the "easy problems" of consciousness, which aim at explaining the dynamics of consciousness by investigating the physical, functional, and/or computational properties of the brain (Baars, 1988; see Chalmers, 1995, for a list of mental phenomena [e.g., responding to stimuli, attention, verbal report, motor control] that are subsumed under such easy problems).

In the domain of neuroscience, many researchers have sought to study consciousness in relation to neural and brain-related processes. The primary focus has been on the easy problems of consciousness, as represented by a search for the neural correlates of consciousness (Crick and Koch, 1998) and establishing a biological framework for visual consciousness (Crick and Koch, 2003). Partly due to the dearth of scientific techniques that could unambiguously reveal the phenomenal aspect of consciousness (i.e., what it subjectively feels like to have an organized or integrated experience of reality [Chalmers, 1995; Nagel, 1974; Searle, 1992]), extant neuroscientific studies showed that the relationship between our conscious decisions and neurophysiological activities is tenuous and susceptible to different interpretations.[1] Keeping these issues in mind, this paper highlights the neuroimaging studies that commenced with the experiments of Benjamin Libet and his colleagues in the 1980s, the criticisms directed against Libet's methods, and the ensuing Libet-type experiments that were done after certain modifications to the original paradigm. By detailing the seminal neuroimaging studies that documented the neural precursors of motor acts, this paper aims to present a historical overview of the Libet-type experiments, and the implications of their findings for understanding conscious decisions and acts. It is vital to note that the study of consciousness via the Libet-type experiments adhered to the philosophical notion of *access consciousness* (Block, 1995), which considers any organism to be conscious as long as it can convert sensory or perceptual information into use for guiding decisions and behaviors. Consequently, there has been a lack of insight into what these experiments mean with respect to the phenomenal aspect of consciousness, and hence this paper proposes some ideas for how the neural

[1]Note that the currently available neuroscientific methodologies and tools are designed for examining brain anatomy and brain-related processes, but not for explaining phenomenal consciousness, which primarily pertains to subjective experience and the feeling of agency ("what is it like?") from an organism's perspective (Nagel, 1974).

precursors of motor acts can be interpreted with regard to some representative constructs of awareness.[2]

The neuroscience techniques used in the studies reviewed pertain to the non-invasive techniques of electroencephalography (EEG) and functional magnetic neuroimaging (fMRI), as well as to the invasive technique of microelectrode recording. In combination, these techniques greatly facilitated the discovery of pre-movement neuronal activity and engendered new questions about how this type of activity should be interpreted with regard to the plausible contents and stages of awareness.

Libet et al.'s Experiments

The seminal studies that investigated the cortical mechanisms of motor acts in humans were performed by Benjamin Libet and his colleagues in the early 1980s at the medical school of the University of California, San Francisco (UCSF) [Libet, Gleason, Wright, and Pearl, 1983; Libet, Wright, and Gleason, 1982]. In the first study in 1982, Libet and colleagues, using electroencephalography (EEG) and electromyography (EMG), examined the event-related potentials (ERP) of six participants performing a simple perceptual–motor task that warranted phasic movements of their wrists or fingers, which were linked to an EMG machine. This task required the participants to fix their gaze at the center of a cathode ray oscilloscope (CRO) clock placed 1.95 m away. A light spot revolved clockwise at the perimeter of the clock at a rapid pace of 2.56 seconds per revolution, and participants were instructed to flex their fingers or wrists after the first revolution whenever they felt a spontaneous urge. They were also asked to note down the spatial location of the light spot on the clock face at the same time they performed the flexing motions. The time that was observed by participants on the clock face corresponded to their consciously self-reported time, while the time that was recorded by the electromyogram pertained to the actual time when participants performed their spontaneous motor acts.

Libet et al. (1982) matched those timings with the ERP of their respective participants and made several discoveries that were startling for their time: they found three types of "readiness potential" (RP) [originally called *Bereitschaftspotentials*, as first discovered by Kornhuber and Deecke, 1965] emanating from the supplementary motor area (SMA) [Brodmann's area (BA) 6] *before* participants' self-initiated motor acts. A Type I RP emerged between 1500 and 1000 milliseconds (ms) before movement, and was generally present among participants who

[2]In line with Libet's (1993) postulate that the neuronal activity before and after movement onset marks the physical correlate of consciousness, this paper identifies awareness as the emergent property of this neuronal activity — that is, as the condition in which there is direct availability of information (perceptual and/or cognitive) for global control (i.e., control of behavior and verbal report) [Chalmers, 1995].

reported some pre-planning of the motor act. A Type II RP emerged at about 550 ms before movement, when the participants were consistently informed to let the urge to move come naturally on its own (i.e., no motor pre-planning). Finally, a late-occurring Type III RP emerged at about 200 ms before the motor act. The Type III RP was further shown to occur earlier than the subjectively reported sensation of a skin stimulus that was randomly delivered after one revolution of the spot of light on the CRO clock (see also Libet, 1989, 1999). This additional finding showed that *W* should not be conceived as being functionally equivalent to the subjective report of being aware of a simple sensory stimulus. Taken together, these three different types of RP showed that cerebral activity emanating from the supplementary motor area could portend the commission of a spontaneous or voluntary motor act rather than simply occuring at the same time as the act itself.[3]

To verify their findings, Libet and colleagues reapplied their experimental design in 1983 using more precise measures of time and ERP onsets, and managed to replicate their findings. This successful replication led to Libet's (1985) proposal that the onset times of the three different types of RP were representative of different cerebral processes. Namely, the Type I RP was regarded as representing pre-intentionality or a general preparedness that was not essentially automatic; the Type II RP was regarded as the harbinger of the cerebral processes that initiated the act before any subjective awareness of it; and the Type III RP was regarded as representing the conscious wish or will (*W*) to make the act. In an attempt to relate consciousness to the emergence of these different types of RPs, Libet (1985) further proposed that our conscious will could function as a *conscious veto* that blocks the consummation of any spontaneous act originating from preparatory cerebral processes at about 150 ms before the act (after removing the time taken for efferent commands to reach the hand muscles). He supported this proposal based on findings of veto RPs from previous experimental sessions in which participants suppressed their intention to act about 100 to 200 ms before the prearranged times at which they were otherwise supposed to act (see Libet et al., 1983). The main negative potential of these veto RPs tended to flatten or reverse at about 150 to 250 ms before the preset time. This implicated that preparatory cerebral processes could be interfered with or vetoed within the same time period that a conscious decision would emerge before a spontaneous act. By describing the conscious will as functioning in the form of a veto, one can imagine a person with conscious will as having the capacity to act intentionally toward the control of a decision.[4]

---

[3]Libet's conception of a "voluntary" motor act refers to a physical act performed out of one's volition or intention. Having an intention connotates a commitment to perform a physical act directed toward or in response to objects, stimuli, and/or events (see, e.g., Davidson, 1963, for a discussion of "intention"). In Libet's experiments, the presence of the Type III RP can be conceived as an internal event that engendered the intention or conscious will (*W*) to act (see main text for details).

[4]Even though it can be argued that the suppression of a motor act — as entailed by exercising one's "conscious veto" — cannot directly represent the presence of intentionality or agency, the operational definition of veto still lies upon a desire or act to reject a decision. More importantly,

Repercussions

Despite replicating significant findings and conducting control experiments that demonstrated the preponderance of neuronal activity before spontaneous motor acts, Libet's experiments generated great debate among scientists and philosophers, and many expressed harsh criticisms of his methods concerning the timing of conscious awareness (see, e.g., Breitmeyer, 1985, Glynn, 1990; Gomes, 1998, 2002). Notably, with respect to the CRO clock, Gomes (1998, 2002) called into question Libet's assertion that the self-reported time associated with perceiving the light spot's location can be directly coupled with the onset of conscious motor decision. He argued that some latency should be expected for the conscious perception of the light spot on the clock face, and that this latency may vary in an unspecified manner across trials depending on the light spot's position. However, he did not recommend any way to record this unknown latency.

Other than Gomes' criticisms, further concerns of imprecision in the timing of awareness pertained to the observation that not all participants in Libet-type experiments were able to unambiguously differentiate between the temporal onsets of conscious decision and motor act (Pockett and Purdy, 2010) and that a "smearing artifact" might account for Libet et al.'s (1983) findings (Travena and Miller, 2002). This smearing artifact pertained to the averaging of EEG components that would make the main negativity vertex of the average waveform appear earlier than the average decision time despite the presence of EEG components from some trials that occurred after decision-making (for details see Travena and Miller, 2002, p. 164).

Technical considerations aside, at the conceptual level, there is a pre-existing opinion of the conscious will or veto as nothing more than an epiphenomenon or illusion, in contrast to perceiving the conscious veto as a representation of the conscious will in action (Wegner, 2002). In an analysis of Libet's experiments, Wegner (2002) suggested that the experience of will might be nothing more than a "loose end — one of those things, like the action, that is caused by prior brain and mental events" (p. 55). Central to Wegner's notion of the illusory conscious will is that "the experience of consciously willing an action is not a direct indication that the conscious thought has caused the action" (p. 2). This statement highlights his principle of *exclusivity* — that is, a thought (or its ensuing action) cannot be proven to be conscious unless it is associated exclusively with a conscious cause (or causes). Despite arguing that the conscious will may just be an

---

the notion of conscious will as being tied to the exercise of one's intention is what researchers in other Libet-type experiments endorsed (see discussions of these experiments in the main text). Henceforth, subsequent mentions of "conscious will" in this paper refer to a conscious intention to perform a decision or motor act.

illusory perception of control over one's actions, Wegner concurred with Libet on the notion that unconscious neural processes could cause intentions and actions (see Wegner and Wheatley, 1999).[5]

On the other hand, Mele (2009) disagreed with both Wegner and Libet with regard to the causal role of unconscious neural processes. Principally, he characterized Libet's Type II RP as a potential precursor or antecedent to a decision to act rather than the actual cause of that decision, and suggested that the Type II RP might be better understood with regard to "urges to (prepare to) flex soon, brain events suitable for being proximal causal contributors to such urges, motor preparation, and motor imagery" (p. 56).[6] Based on such accounts, Mele eschewed giving a concrete answer as to whether or not all our decisions are conscious, and instead emphasized the causal role of intentions in generating actions regardless of whether or not we consciously experience the process of intention formation that incorporates decisions.[7]

In addition, there were similar views endorsed by Levy (2005), who agreed with Mele with respect to the pertinence of intentions for generating actions, but disagreed with Libet over the implication of the "conscious veto." Levy viewed the control of actions or behaviors strictly in the form of conscious volitions or intentions (brought about by exercising the "conscious veto") as "Libet's impossible demand," saying that the presupposition of a conscious control system would warrant an additional control system at a higher level (ostensibly conscious as well) to control it, and that this process might repeat itself perpetually, causing "an infinite regress of controllings" (Levy, 2005, p. 67). Therefore, Levy argued that our free will — as naturally characterized by the freedom to pursue our actions irrespective of external events — does not need to depend on decisions, volitions, or intentions that are irrevocably conscious (cf. Rosenthal, 2002, presented below). Principally, he stressed that the course of controlling our decisions and actions should not be regarded as inherently conscious, and that non-conscious mechanisms should be assigned functional roles in the emergence of a conscious state. Even though Levy's (2005) advocacy of non-conscious mechanisms was not at odds with Libet's (1985) interpretations of the Type I and Type II RPs as being

[5]To associate an outcome closely with its cause, Wegner proposed two other principles in addition to *exclusivity*: (i) *priority*: the thought must occur before the action; and (ii) *consistency*: the thought must be congruent with the ensuing action (Wegner and Wheatley, 1999).

[6]Note that a *decision*, according to Mele (2009), is defined as the mental act of forming an intention that is settled on executing a plan of action. Ensuing use of the word in the main text shall adhere to this definition.

[7]Mele (2009) emphasized the role of intentions in generating actions through an interesting example of turning on a switch: "A subject's wanting to flex soon and his experience of wanting to flex soon are not the same thing.… My flipping a light switch — not my *experience* of flipping it — is a cause of the light going on. Analogously, a subject's wanting to flex soon may be a cause of his flexing even if his experience of wanting to flex soon is not" (pp. 32–33).

closed to subjective awareness, he failed to relate his proposed non-conscious mechanisms to any specific readiness potential or neural signal that preceded the emergence of a conscious decision or act. This was perhaps due to his critical view of the empirical evidence produced by Libet and colleagues, which he dismissed as offering any serious challenge to the necessity of free will — a necessity that he thought could be proven on conceptual grounds.

All of the above criticisms, however, did not negate the validity of Libet et al.'s (1982, 1983) original findings, which had been accepted and praised by many of the world's leading neuroscientists (see Libet, 2002). Libet, on many occasions, was also able to offer sound counterarguments against his critics to justify his methodology and interpretations (see, e.g., Libet, 2000, 2002). In the philosophical domain, his works were notably supported by Rosenthal (2002), who argued that the neural signals preceding conscious volition or intention (namely, Type I and Type II RPs) could be identified as the direct indicators of unconscious states/events or their approximate physical correlates (i.e., neural signals that did not represent unconscious events per se but occurred simultaneously with unconscious events). Unlike Mele (2009), who did not openly acknowledge the role of unconscious decisions or events, Rosenthal (2002) argued that we must abandon an essentialist or commonsensical view of consciousness entailed by the belief that "no mental state counts as being conscious unless the individual who is in that state is conscious of the state" (p. 218). Crucially, he asserted that we need to conceptualize a mental state in dynamic terms — that is, a mental state can be conscious at one time and non-conscious at another time. By this account, the train of readiness potentials culminating in a motor act can be seen as a transformation whereby a prior non-conscious state is turned into a conscious state.

More importantly, with respect to empirical research, the implications of Libet's experiments remained relevant because other researchers were able to replicate his basic finding of early emergence of RPs prior to motor decisions based on different instruments and/or stimuli (see, e.g., Keller and Heckhausen, 1990; Pockett and Purdy, 2010; Travena and Miller, 2002). In general, the other Libet-type experiments showed the same pattern of results as that of Libet et al. (1982, 1983) regardless of differences in stimuli and the type of motor response (e.g., pressing keys): the emergence of the RP representing decision onset was found to be either close to or more than 250 ms before the time of action execution. Particularly noteworthy was Haggard and Eimer's (1999) discovery of the lateralized readiness potential (LRP), a special form of the readiness potential that occurred about 800 ms before movement initiation. Haggard and Eimer (1999) showed that the LRP occurred significantly earlier in trials with early awareness of movement initiation than in trials with late awareness of movement initiation, implicating that the processes underlying the LRP may have causal roles to play in initiating our awareness of movements.

Furthermore, a modified Libet-type experiment conducted with functional magnetic resonance imaging (fMRI) by Lau, Rogers, Haggard, and Passingham (2004) pinpointed the pre-supplementary area (pre-SMA) [the rostral portion of BA 6] as the site that is tightly associated with the generation of spontaneous acts. In the experimental condition, the participants gazed at a red dot revolving around a clock face at a rate of 2560 ms per cycle and encoded its location while performing a button press on each trial. Under this condition, the blood-oxygen-level-dependent (BOLD) activation in the pre-SMA was found to be significantly higher than that from the control condition in which the participants made their button presses without attending to the dot's location. Specifically, significant differences in BOLD signals between the two conditions were observed in the first six seconds after the onset of the spontaneous button press. Critically, activation in the dorsal prefrontal cortex (dPFC) [BA 9/46], commonly associated with motor planning, was found to be closely associated with activation in the pre-SMA during the intention phase when participants attended to the dot's location. The authors thereby concluded that activity in the pre-SMA is tightly coupled to an intention to move that involves attending to a moving stimulus. Importantly, this conclusion supported Libet's interpretation of the Type III RP as reflective of conscious intention.

Is Conscious Intention a Veto?

As for Libet's principal proposal of conscious intention functioning as a veto, another Libet-type fMRI study (Brass and Haggard, 2007) gave support to his claim by implicating the dorsal fronto-median cortex (dFMC) [BA 9)] to be involved in action inhibition. In that study, the participants gazed at a Libet clock with a clock hand moving at a rate of 3000 ms per cycle. After one full revolution of the clock hand, participants under the "action" condition had to spontaneously initiate a key press while participants under the "inhibition" condition had to refrain from the act. Both parties had to judge the temporal onset of their decisions after these two phases. Contrasts between the inhibition and action conditions yielded strong activation in the dFMC while the reverse contrast between these two conditions did not yield any significant activation in the pre-SMA and SMA, which suggested that participants prepared their intentional acts equally well under both conditions. Moreover, participants who displayed higher frequencies of inhibition were found to have stronger inhibition-related dFMC activation, and a significant negative correlation was found between activation in both the dFMC and the primary motor cortex (BA 4). The authors concluded that the dFMC could be a specific brain area involved in the inhibition of intentional actions through a top–down signal gating of the neural pathways linking intention to action.

Even though Brass and Haggard's (2007) study seemed to have offered substantial evidence to vindicate the existence of the conscious veto, it did not offer support for Libet's other proposal of the veto operating within a time span of about 150 ms before movement. This time span for vetoing action did not seem to apply to some patients with parietal lobe lesions who reported the onset of intention to be as late as 50 ms prior to action execution (Sirigu et al., 2004).[8] If exercising the conscious veto is as crucial as what Libet purported it to be, those patients would have an almost negligible amount of time to evaluate their spontaneous intentions and inhibit unwanted ones. Yet there were no reports of patients being unable to make and change their motor decisions.

Consequently, this absence of converging evidence for the temporal range in which the conscious veto occurred led other researchers to propose that the motor decision-making time in the Libet experiments could be influenced or modulated by the "attention to the intention to move" (see Lau, 2009; Lau, Rogers, and Passingham, 2006; Pockett and Purdy, 2010). By interpreting Libet's results with referral to the doctrine of *prior entry*, which stipulated that attended stimuli must be perceived prior to unattended stimuli (Shore, Spence, and Klein, 2001), Lau (2009) suggested that in the Libet experiments attention to the intention to move might have acted as an endogenous cuing process, biasing participants' self-reports of response times to be earlier than what were recorded based on their motor movements. This suggestion was backed up by findings from an fMRI study by Lau et al. (2006) that involved the same Libet-type clock as the one utilized previously by Lau et al. (2004). This follow-up study involved a 2 x 2 factorial design with timing and modality as the independent variables. The timing variable involved "timing" and "nontiming" conditions while the modality variable involved "action" (i.e., pressing keys) and "auditory" (i.e., hearing tones) conditions. In all four conditions, the participants gazed at an unnumbered Libet-type clock that had a red dot revolving around its clock face at 2560 ms per cycle and gauged the location or time of the final appearance of the dot. On each trial of the *action timing* condition, the participants observed one revolution of the dot and made a spontaneous button press while noting the location of the revolving dot, which disappeared shortly after the button press. When the red dot reappeared at the clock's center after a variable delay, the participants operated a game pad and moved the dot to where it was located before they pressed the button. The spatial difference between this manually shifted location and the dot's precise location during the onset of the button press was translated into a temporal difference based on the rule of 1° = 7.1 ms. In the *action nontiming* condition, the dot disappeared after just one revolution; it only reappeared at a random location on the clock face after the participant pressed a button. The participants had to remember this location and subsequently relocate the dot from the clock's center after a delay based on

[8]This relatively late onset of the intention to move was also biased by the fact that the patients with parietal lobe damage were much better at judging the time of actual movement than the time in which the intention to move first arose.

the same (aforementioned) procedure. The *auditory timing* and *auditory nontiming* conditions followed the events from the action timing and action nontiming conditions, respectively, except that no spontaneous button presses and relocation of the dot were involved. A tone sounded before the disappearance of the dot at the end of each trial, and the participants reported their estimated time of the tone's onset.

Based on a whole-brain analysis that yielded Fourier series (i.e., general sinusoidal waveforms) of BOLD signal changes over a period of 16 seconds after the onset of time estimates, Lau et al. (2006) found a significant interaction between timing and modality from activation in the cingulate motor area (CMA), an area that falls below the supplementary motor area (SMA) [i.e., the posterior part of the anterior cingulate cortex (ACC)]. Specifically, significantly greater activation in the CMA was derived from contrasting the action timing condition against the action non-timing condition but not from contrasting the auditory timing condition against the auditory nontiming condition. Lau and colleagues also computed a BOLD modulation measure for the CMA and showed that it was negatively and significantly correlated with the perceived time of onset. This implicated that greater CMA activation was associated with greater negative time estimates showing a perception of button presses that occurred earlier than their actual onsets (i.e., the exact time of pressing the buttons). Critically, Lau et al. (2006) did a reanalysis of their previous data on the pre-SMA (collected by Lau et al. [2004]) and found the same negative correlation to exist between the BOLD modulation measure for the pre-SMA and the perceived onset of intention. Based on these findings, Lau and colleagues suggested that the change in CMA activity could be best understood as attentional modulation brought about by the need for in-depth processing of the information required for action execution. The participants were suggested to make use of this increased activity in the CMA to time their movements. As for activity in the pre-SMA, the authors suggested that it was modulated by the amount of attention devoted to the intention to execute a motor act, and that their earlier finding of 228 ms being the average perceived onset before a button press might have been due to early and full attention to the intention to act on the part of their participants. Therefore, neuromodulation in both the CMA and the pre-SMA leading to individual variability in the temporal perception of action onsets could explain why the onset of intention does not always have to emerge within Libet's proposed period of 250 ms prior to action.

In addition, the relevance of attention for motor decision is also supported by findings from a Libet-type experiment by Pockett and Purdy (2010). During "decision" trials in which the participants were instructed to decide on the correct key to press after summing up a pair of numbers displayed at the center of the Libet clock, some participants either did not exhibit any RP or exhibited RPs that tended to start at the same time as their self-reported decision time. Pockett and Purdy (2010) explained these peculiar findings by suggesting that the attention of participants in the time period before action execution was completely taken up by performing the additions. Thus, their attentional focus on the arrival of any spontaneous urge was undermined, leading to the absence of RPs. Importantly, the authors proposed that the RPs found by Libet and colleagues might be more reflective of general readiness or expectancy rather than specific preparation for movement. They argued that Libet-type experiments could demonstrate the ERPs that engendered urge-related movements but could not relate these ERPs to conscious or unconscious decision-making.

Recent Developments

Pockett and Purdy's (2010) argument that antecedent brain signals could not be related to specific motor preparation or the intention to act was countered by findings from a study by Fried, Mukamel, and Kreiman (2011) that applied microelectrode recordings of neuronal activities in the human medial frontal cortex. Fried et al. (2011) adopted Libet et al.'s (1983) paradigm; aside from two changes entailed by: (i) using an analog clock that had a revolving hand instead of one with a light spot moving along the circumference, and (ii) instructing participants to press keys instead of making wrist movements, the main procedures were kept the same. The researchers planted depth electrodes into the frontal and temporal lobes of 12 epileptic patients and recorded extracellular activity from 760 units (264 single units and 496 multiunits) in the medial frontal lobe, comprising the anterior cingulate cortex (ACC) [BA 24/32], the pre-SMA, and the SMA proper, as well as from 259 units in the temporal lobe. Similar to what Libet et al. (1983) found, Fried et al. showed that the onset of conscious awareness — the will or wish (*W*) to elicit a spontaneous motor act — occurred at an average of 193 ms prior to the key press. Notably, during the 400 ms interval prior to *W*, 17% of all the recorded units in the medial prefrontal cortex exhibited changes in firing rate from the baseline firing rate (recorded from 2500 to 1500 ms before *W*), and this proportion was larger than the 13% of recorded units that exhibited changes in firing rate after *W*. In particular, neurons that increased and decreased their firing rates prior to *W* were found. The former was dubbed "increasing" neurons, and the latter "decreasing" neurons. When comparing the response profile of the increasing neurons between the frontal and temporal lobes, the proportion of responsive neurons demonstrating steady increases in firing rate prior to *W* was markedly higher in the medial frontal lobe (comprising the anterior cingulate cortex, the pre-SMA, and the SMA proper) than in the medial temporal lobe. To determine whether the neuronal activity in the medial frontal lobe had any causal relationship with the onset of *W*, Fried and colleagues applied a support vector machine (SVM) classifier (see Hung, Kreiman, Poggio, and DiCarlo, 2005) to discriminate between neuronal activity before *W* and baseline activity in single trials. In machine learning, support vector machines are supervised learning models associated with learning algorithms that are used for categorization and regression analysis. Given a reference or training data set, which in Fried et al.'s study pertained to the recorded activity of a population of neurons chosen at a certain time before the onset of *W*, the support vector machine training algorithm built a model that enabled the categorization of other ensembles of neurons based on whether they were activated before or after *W*. Overall, this linear algorithm predicted the neuronal onset of *W* (i.e., the readiness potential) to occur at an average of 152 ms prior to the self-reported onset of *W*. Fried and colleagues interpreted these findings as suggestive of an "integrate-and-fire" mechanism that integrates the firing of ensembles of medial frontal neurons until a threshold is reached for the emergence of the intention to act. Despite showing that changes in the firing rates of medial frontal neurons could predict the early onset of *W*, the authors remained circumspect and refrained from passing judgment about whether the neuronal changes detected by them in the medial frontal lobe *caused* the emergence of volition. More recent evidence by Schultze–Kraft et al. (2016) showed that the conscious veto could be exerted after the onset of the RP, but it must be exerted within a short period that occurred less than 200 ms before movement onset in order to enable movement cancellation. In other words, this means that conscious control over a spontaneous act cannot be exercised after passing the mark of 200 ms before movement onset — a so-called "point of no return" (Schultze–Kraft et al., 2016). This suggests that the conscious veto is very transient in nature and must be exercised immediately after RP onset in order to abolish any movement of interest.

Critically, Fried et al.'s (2011) findings resonated with a series of fMRI studies conducted by Soon and colleagues (Bode et al., 2011; Soon, Brass, Heinze, and Haynes, 2008; Soon, He, Bode, and Haynes, 2013; for a commentary, see Soon, Allefeld, Bogler, Heinzle, and Haynes, 2014). All of these studies utilized multivoxel pattern analysis (MVPA) on spatiotemporal patterns of brain activity acquired from Libet-type experiments. Multivoxel pattern analysis focuses on multiple volumetric brain pixels ("multivoxels") instead of single volumetric brain pixels ("voxels") and applies relevant pattern classification algorithms to decode the multivoxel patterns of activity occurring at certain timepoints; as such, these patterns were regarded as spatiotemporal in nature (for details about the benefits and technical nuances of MVPA, see Norman, Polyn, Detre, and Haxby, 2006). In the first of these studies, Soon et al. (2008) showed the participants slides of single consonants separated by intervals of 500 milliseconds and instructed the participants to make spontaneous button presses with either their left or right hand whenever they felt the urge to do so. A screen with four consonants appeared after each spontaneous button press, and participants had to select the consonant that corresponded to the moment in which they made their motor decision. By using the same type of SVM classifier that Fried et al. (2011) used to predict the type of spatiotemporal pattern of brain activity associated with either a left or right button press, Soon et al. (2008) showed that neuronal activity in the frontopolar cortex (BA 10) and precuneus/posterior cingulate region (BA 7) preceded the onset of motor decision by as long as seven seconds (!). The duration of seven seconds was particularly surprising and groundbreaking at the time of its discovery because it far exceeded Libet's (1985) 300 ms interval that separated the onset of the RP and the first conscious decision to move. In order to clarify the roles of the frontopolar cortex and the precuneus, Soon et al. (2008) also conducted a control fMRI experiment that instructed participants to decide on making a left or right button press when shown a verbal cue of "select," and to respond after a variable interval when shown another verbal cue of "respond." The classification algorithm predicted neuronal activity in the frontopolar cortex with higher classification accuracy than in the precuneus during the selection phase and showed the reverse trend during the response phase, culminating in a double dissociation. These findings led the authors to suggest the frontopolar cortex as the initiator of unconscious processing of motor decisions and the precuneus as the temporary storage site of the motor decision before it reached consciousness:

> The temporal ordering of information suggests a tentative *causal* model of information flow, where the earliest unconscious precursors of the motors decision originated in frontopolar cortex, from where they influenced the buildup of the decision-related information in the precuneus and later in SMA, where it remained unconscious for up to a few seconds. (Soon et al., 2008, p. 545, italics added)

Regardless of the promising findings, there was a noticeable pitfall in the study, as pointed out by Haynes (2010). This pertained to the ostensibly low level of an average classification accuracy of 60% predicting decisions in the respective cortical sites; even though the 60% accuracy rate was reliable, it was far from perfect. The spatiotemporal resolution provided by a conventional 3T fMRI scanner can only offer a limited amount of the broader information that could be gained based on a more direct measurement of the activity of frontal lobe neurons based on microelectrode recordings, as Fried et al. (2011) subsequently demonstrated. Undeterred by this technical limitation, Soon and Haynes, together with a new team of researchers (Bode et al., 2011), re-conducted their study using ultra-high field fMRI (7T scanner) on the frontopolar cortex and replicated their original findings. Based on improved spatial and temporal resolution, the authors showed that the earliest time at which successful decoding (i.e., classification of neuronal ensembles based on whether or not they cohere with the spatiotemporal patterns associated with an upcoming motor decision) was possible was about 7.5 seconds before a decision was reported to be consciously made, which was slightly earlier than the seven seconds found by Soon et al. (2008). Like the previous study, the classification accuracy for upcoming motor decisions from the frontopolar cortex was not remarkably high during this interval of 7.5 seconds, ranging from 52% to

57%. Nonetheless, the fMRI results were supported by post-experimental surveys collected from the participants, who generally reported that they had been very relaxed and spontaneous in their actions, harboring no specific thoughts during the experiment. The successful replication of the earliest onset of neuronal activity in the frontopolar cortex led the authors to propose this region as a core area for free decisions, one that lies at the top of a hierarchically organized prefrontal functional network.

To test the notion that the frontopolar cortex was indeed involved in the early processing of free decisions that were varied in nature and not limited by motor decisions, a follow-up fMRI study was performed to investigate how the fronto-polar polar cortex and the precuneus were involved in the early processing of voluntary arithmetic decisions that incorporated additions and subtractions (Soon et al., 2013). The decision to add or subtract numbers was seen by the authors as a higher-level and more abstract type of decision compared to the decision to make a spontaneous key press. The participants viewed a series of slides with four digits placed at the corners and a digit at the center with a consonant placed below it. Whenever the participant felt ready to carry out an arithmetic decision, he attended to the digit at the center of the screen and either added or subtracted the centered digit appearing on the next slide. On the third slide, the participant pressed one of the four buttons that corresponded to the spatial locations of the digits at the corners of the slide. Two of those digits conveyed the right answers to the addition and subtraction, respectively; their positions randomly changed from corner to corner on each slide. After making the button press that indicated whether an addition or a subtraction was performed, a slide with four consonants appeared, and the participant selected the consonant that matched the consonant seen on the earliest slide during which he initially made his voluntary decision to perform the arithmetic task. SVM classifiers were used once more, and were trained to distinguish between the spatiotemporal patterns of brain activity related to addition and subtraction. The results showed that a medial frontopolar region (within (BA 10) and a region straddling the precuneus (BA 7) and the posterior cingulate cortex (BA 23/31) encoded the outcome of the impending decision about four seconds prior to its realization. Once more, like in the previous two studies, the classification accuracies were around 60%. Notably, the time-course of classification accuracies partially overlapped with the time-course of activation in the "default mode" network, an interconnected brain system (spanning the fronto-parietal axis) that is usually activated when the individual is generating spontaneous thoughts without focusing on signals from the outside world (e.g., thoughts generated during mind-wandering; see Buckner, Andrews–Hanna, and Schacter, 2008). Critically, the default mode activity and the level of classification accuracy in the frontopolar and precuneus/posterior cingulate regions were found to peak at around the same time of four seconds prior to conscious decision. This overlap was seen by Soon et al. (2013)

as supportive evidence for the notion that preparatory neuronal activities in the frontal and parietal regions were reflective of unconscious processes. Importantly, the authors suggested that the relatively shorter period of neuronal activity generated for the upcoming arithmetic decision might showcase the limitations of unconscious processes in developing and stabilizing more complex representations stemming from abstract intentions. These intentions pertain to arithmetics and other higher-level mental operations.

Discussion

Taken together, we get to see that an assembly of areas in the frontal lobe of the brain — namely the pre-SMA, the SMA (both part of BA 6), the anterior cingulate cortex (BA 24), the medial prefrontal cortex (BA 9), and the frontopolar cortex (BA 10) [in a caudal to rostral direction] — are implicated to be involved in engendering the intention or volition to commit a motor act prior to an individual becoming fully conscious of the intention and then performing the act. The works by Lau and colleagues supported the pioneering works of Libet and colleagues by putting emphasis on the "attention to the intention to act" as the driving "force" that brought about activation in the pre-SMA and anterior cingulate cortex. Fried et al. (2011) applied the fine-grained approach of microelectrode recordings and demonstrated that neurons in these brain regions were indeed active before the time at which one became conscious of the wish to act (*W*). However, they are much more restrained in the interpretations of their findings compared to Soon and colleagues, who appeared to be advocating for a causal trajectory of neural events that stemmed from early neuronal activity in the frontopolar cortex (Bode et al., 2011; Soon et al., 2008). It is crucial to note that Soon and colleagues did not implement the Libet-type CRO clock paradigm, nor did they analyze their data based on conventional indicators of neural activity (i.e., BOLD signals, neuronal firing patterns). Principally, they applied a computationally demanding technique of MVPA to see how the spatiotemporal activity patterns of an ensemble or population of neurons (i.e., clusters of 3D volumetric pixels) captured at a time before the onset of a motor decision could predict the likelihood of the decision's impending occurrence. Surprisingly, they found that such a prediction could occur beyond a 50% chance in as long as seven seconds before a simple act of pressing buttons (Bode et al., 2011; Soon et al., 2008), and in about four seconds before deciding to perform an arithmetic problem by pressing a button (Soon et al., 2013). Despite arguing for unconscious processes as the harbinger of the will to act, with respect to higher-level thinking and reasoning (of which doing mental arithmetic is a part), Soon et al. (2013) were not able to ascertain whether unconscious and conscious representations could be subserved by the same substrates within the frontal and parietal regions or whether such representations could be separated at a finer scale. Therefore, future studies, as Soon et al. (2013) proposed, should consider using tasks that elicit conscious and non-conscious/automatic decisions, along with the training of classification algorithms to predict these two types of decisions based on spatiotemporal patterns of pre-movement neuronal activity. This proposal was consistent with that of Fried et al. (2011), who recommended future investigations of the firing profiles of neurons in the parietal cortex *before* and *after* the emergence of conscious intention, so as to better understand the mechanisms of conscious and unconscious processes.

Unresolved Issues and Future Directions

The aforementioned proposals showed that we are merely at the tip of the iceberg in our modern endeavors to understand the nature of consciousness. Despite the technological benefits offered by modern neuroimaging techniques, we are still unclear about how to differentiate between conscious and unconscious processes and their underlying neural substrates. Critically, the proposals for future research by Soon, Fried, and their colleagues show that it is still too early to interpret what an early onset of neuronal activity prior to the onset of a motor decision truly means. And supposing that unconscious (or subconscious) motor pre-planning indeed occurred seven seconds before the act, when and how would Libet's notion of the conscious veto apply — considering these words of Libet (1999)?

> Some have proposed that even an unconscious initiation of a veto choice would nevertheless be a genuine choice made by the individual and could still be viewed as a free will process (e.g., Velmans, 1991). I find such a proposed view of free will to be *unacceptable*. In such a view, the individual would not consciously control his actions; he would only become aware of an unconsciously initiated choice. He would have no direct conscious control over the nature of any preceding unconscious processes. But, *a free will process implies one could be held consciously responsible for one's choice to act or not to act*.[9] (p. 52, emphases added)

Based on this account, it is possible that Libet would have disagreed with Soon et al.'s, (2008) interpretation of tracing the origin of voluntary acts — a representation of free will in action — to unconscious origins. Consequently, this beckons us to question what being unconscious truly means. Does it refer to a superficial form of unawareness (i.e., a state in which information for global control is either not fully available or fully processed [Chalmers, 1995]) that is divorced from underpinning neurophysiological activity — or does it refer to a full-fledged abandonment of attention or intentionality that eschews any reflection on its phenomenological

[9]Ensuing discussions of "free will" shall follow Libet's conception. Specifically, "free will" shall be defined as "the power of an individual to make free choices, not determined by divine predestination, the laws of physical causality, fate, etc." ["free will, n." (2017, March 25). *Oxford English Dictionary Online*. Retrieved from http://www.oed.com/view/Entry/74438?redirectedFrom=free+will]

contents (i.e., an abandonment of the state/event consciousness, with the subject having no access to her internal mental state [Lycan, 1987, 1996])? There is certainly a yearning for a clearer conception of what the "contents" of a conscious (or unconscious) mental state may be, as shown by the caveat raised by Libet (1999) on the distinction between awareness and its contents:

> Our own previous studies have indicated that *awareness* is a unique phenomenon in itself, distinguished from the *contents of which one may become aware*. For example, awareness of a sensory stimulus can require similar durations of stimulus trains for somatosensory cortex and for medial lemniscus. But the *content* of those awarenesses in these two cases is different, in the subjective timings of sensations (Libet *et al.*, 1979). The content of an unconscious mental process (e.g. correct detection of a signal in the brain *without any awareness* of the signal) may be the same as the content *with awareness* of the signal. But to become aware of that same content required that stimulus duration be increased by about 400 msec (see Libet *et al.*, 1991). (Libet, 1999, p. 53, emphases added)

By this account, Libet implied brain-related activity was an instance of the "content of which one may become aware." With regard to the studies of Soon and colleagues, this content could refer to the classification accuracies of the neuronal ensembles. The classification accuracies shown immediately before and after the motor decision matched each other at approximately the same level (Soon et al., 2008), and buttressed Libet's idea of invariant content irrespective of the state of awareness.[10] Hence, the possibility remains for future research to endorse MVPA classification accuracies as potential *observable* representations of the contents of awareness that is distinct from the condition of "being aware." Furthermore, from a conceptual standpoint, Libet's contents of conscious awareness may be conceived as analogous to the contents of higher-order perception engendered by the evaluation of our perceptual experiences (cf. Armstrong, 1968; Lycan, 1996).[11] Since the attainment of any "higher-order" status implies a hierarchy of levels or events, this analogy begs the question of whether being aware could be understood as a subtle progression of mental events leading to the emergence of conscious awareness.

In order to examine what "being aware" connotes, it may be pertinent for future researchers to consider the process of voluntary movement as arising out of several stages of awareness, flowing from (i) a state of total non-awareness of impending movement, to (ii) a state of intentionality contingent on the availability of probes/

---

[10] For a graphical illustration of the matching of the classification accuracies, the reader is advised to refer to Figure 2 in Soon et al.'s (2008) article.

[11] This notion ascribes to the theory of higher-order perception that was first proposed by Locke (1690). This theory is also called the "inner sense theory," stipulating that a higher-order, non-conceptual, and intentional state can engender a phenomenally conscious state via a faculty of "inner sense." This "inner sense" refers to our innate capacity to construct mental representations based on first-order perceptual awareness (i.e., the availability of information from the sensory modalities).

cues, to (iii) a state of meta-awareness (akin to the onset of the conscious will [*W*] based on Libet's paradigm, emerging around 250 ms before movement, which is slightly beyond a full-fledged state of awareness following movement onset), and to (iv) a final point of no return after which the motor act cannot be vetoed (i.e., 200 ms before movement, according to Schultze–Kraft et al., 2016; see also Matsuhashi and Hallett, 2008; Smallwood and Schooler, 2006). However, owing to the fact that this proposal of awareness progression (Matsuhashi and Hallett, 2008) stemmed from a veto paradigm (i.e., vetoing finger extensions after hearing tones that coincided with the conscious intention to move) that deviated substantially from Libet's CRO clock paradigm and engendered a much earlier time at which the thought to move arose (1.42 seconds before movement, more than a second earlier than Libet's *W*), it cannot be applied unequivocally to explain the findings of Libet et al. (1982, 1983). Even though the onset of Type I RP (about 1000 ms before movement) recorded in Libet et al.'s studies would have conformed to Matsuhashi and Hallett's (2008) proposed stage of intentionality, the ensuing onset of Type II RP, interpreted as a marker of an absence of pre-plans, would not have fit well with the same stage. This is due to the absence of any auditory signals in the CRO clock paradigm that would have served as potential probes or cues for awareness. Nonetheless, the vetoing of pre-plans to act based on hearing tones does have an advantage over Libet's paradigm in that it relies on real-time decisions rather than on post-event subjective recall and potential latency lapses in reading the dot's position on the CRO clock. Therefore, it would be best to approach Matsuhashi and Hallett's (2008) stages of awareness progression as a paradigm-specific model that addresses the timing of the intention to move in terms of the conscious veto.[12]

Interestingly, an operation of the veto paradigm in reversed mode (i.e., making immediate responses instead of abolishing them) can be seen from a *Libetus Interruptus* paradigm that was designed to test the validity of an accumulator stochastic decision model of the neural decision to move (Schurger, Sitt, and Dehaene, 2012). This paradigm was derived from the CRO clock paradigm and was similar in all aspects except that it further required participants to respond spontaneously (i.e., pressing a button immediately) whenever they heard a clicking sound. The findings based on this paradigm showed that the neural decision to move came at a time closer to the actual movement (about 150 ms before movement) compared to Libet's *W* (about 250 ms before movement) and was preceded by a gradual negative-going voltage deflection that reflected the buildup of the mounting urge to move. When comparing neural decision times of movement obtained from the *Interruptus* and veto paradigms to Libet's *W*, the former paradigm yielded a noticeably smaller temporal difference (around 100 ms) than the latter paradigm (around 1.2 seconds). The same pattern of temporal differences could be seen

[12]This type of theoretical model ought to account for analogous patterns of findings replicable under similar experimental conditions, in which the task stimuli and instructions are kept largely invariant.

when the neural decision times obtained by Bode et al. (2011), Soon et al. (2008, 2013), Lau et al. (2004, 2006), and Fried et al. (2011) were compared to Libet's *W*. The first set of comparisons (involving studies by Bode, Soon, and colleagues) yielded differences on the scale of several seconds (for as large as 6.75 seconds when considering Soon et al.'s [2008] MVPA findings) whereas the latter set of comparisons (involving studies by Lau, Fried, and colleagues) yielded differences on the scale of tenths of milliseconds. In conjunction, these pieces of evidence showed that close variants of Libet's CRO clock paradigm (Fried et al., 2011; Lau et al., 2004, 2006; Schurger et al., 2012) generated less discrepancy in terms of pre-movement neural decision times than other voluntary decision paradigms that did not implement the same testing interface (Bode et al., 2011; Matsuhashi and Hallett, 2008; Soon et al., 2008, 2013). This suggests that the qualitative aspects of an experimental paradigm should be considered when constructing any model or theory that attempts to explain the neural precursors of any conscious decisions or acts. Such aspects centrally pertain to the cues/probes (if any) for conscious decision-making and movement, the modality of stimuli presentation, task-related instructions, demand characteristics, and their relevant control procedures.[13]

More importantly, based on the premise that analogous experimental paradigms/tasks are more likely to yield results that can be framed under a common theory compared to dissimilar or distinct paradigms, it may be worthwhile for future investigators of pre-movement neuronal activity to construct models or theories that are paradigm-specific.[14] Each of these models/theories, being centered on a distinctive paradigm, is likely to contribute to a more nuanced elucidation of the relationship between pre-movement neuronal signals and the contents or stages of awareness. Principally, they should serve the purpose of explaining the inconsistent occurrences of *W* that were found based on different experimental paradigms. Over time, an accumulation of such theories offers the potential to generate a standard or integrated framework that could explain the same neurophysiological phenomenon across different experimental contexts or modes of testing.

Conclusion

We have come a long way since Libet's first experiments demonstrating the presence of cortical activity prior to voluntary acts, and have obtained greater insights into the underlying neural activities, localized to the frontal and medial regions of the brain, through modern neuroscience techniques. However, a complete picture

---

[13] As defined by Orne and Whitehouse (2000, pp. 469–470), demand characteristics refer to "the totality of cues and mutual expectations which inhere in a social context…which serve to influence the behaviour and/or self-reported experience of the research receiver."

[14] This is contingent on the similar paradigms/tasks being comparable in terms of operation, tools/equipment for task presentation, and modes of data analysis.

of the relationship between a motor act and its neural precursors remains elusive, and more work can be conducted to relate different stages of awareness (or non-awareness) to the different patterns of neuronal activity preceding a movement, preferably with regard to different types of paradigms devised for investigating such pre-movement activity. Ultimately, much more work needs to be done to uncover the hidden mysteries of our consciousness before we launch a thorough intellectual discussion of what being conscious truly means. Until we do so, it is important that we abstain from adopting a bipolar stance regarding the origin of our conscious thoughts and behavior — that is, to characterize these origins as either totally deterministic (i.e., aligned with universal physical laws) or totally non-deterministic (i.e., conforming to imperceptible phenomena that violate physical laws) [Libet, 2004]. As Nahmias (2011) rightly points out, determinism is typified by prior events causing present events founded upon universal physical laws, and should *not* be taken to mean that a person's beliefs, desires, and decisions have no purpose for what one tries to do. Through the survey studies by Nahmias, Morris, Nadelhoffer, and Turner (2005, 2006), the "problem" in linking determinism and free will together has been ascribed to individual differences in the comprehension of whether or not our beliefs, desires, and decisions can be eschewed (or "bypassed," as Nahmias et al. [2005, 2006] termed it) when generating actions. Determinism does not have to be regarded as incompatible with free will so long as we do not endorse a fatalistic view of the world (i.e., a whatever-happen-will-happen mentality) [Nahmias, 2011; Nahmias et al., 2005, 2006; Nahmias and Murray, 2010].[15] If future consciousness researchers and theorists are willing to consider the multifarious differences in belief systems that different organisms endorse with respect to their perceptions of reality, it is very likely that we shall gradually learn to reconcile deterministic events with our deliberations and decisions that mark the cornerstones of free will.[16]

Regardless of the direction chosen for further research, it would be prudent to heed the following advice:

> My conclusion about free will, one genuinely free in the nondetermined sense, is that its existence is at least as good, if not a better, scientific option than is its denial by natural law determinist theory. Given the speculative nature of both determinist and nondeterminist theories, why not adopt the view that we do have free will (until some real contradictory evidence appears, if it ever does)? Such a view would at least allow us to proceed in a way that accepts and accommodates our own deep feeling that we do have free will. (Libet, 2004, p. 156)

---

[15]Conversely, people could be induced to think that determinism and free will are incompatible when they are instructed to believe that they live in a totally deterministic universe, in which each decision they make must happen in a predetermined way (Nahmias et al., 2005, 2006; Nahmias and Murray, 2010).

[16]In the real world, the need for reconciliation between the laws of determinism and the notions of free will is of invaluable import with regard to law and justice. Legal responsibility is inadvertently tied to free will, and a denial of free will in favor of deterministic elements or events could cause unfavorable impediments or difficulties in the evaluation of eyewitness testimonies and the sentencing of criminals (see Rychlak and Rychlak, 1997).

Perhaps this sums up Libet's legacy — that we must neither conform to a dogmatic view of consciousness nor lose faith in our quest to discover the origin and mechanisms of consciousness. Many untrodden paths in consciousness research awaits us, and we must journey on with courage and open hearts.

References

Armstrong, D. (1968). *A materialist theory of the mind*. London: Routledge.

Baars, B. (1988). *A cognitive theory of consciousness.* Cambridge: Cambridge University Press.

Block, N. (1995). On a confusion about the function of consciousness. *Behavioral and Brain Sciences*, 18, 227–247.

Bode, S., He. A. H., Soon, C. S., Trampel, R., Turner, R., and Haynes, J. D. (2011). Tracking the unconscious generation of free decisions using ultrahigh field fMRI. *PLoS ONE*, *6* (6): e21612. doi: 10.1371/journal.pone.0021612

Brass, M., and Haggard, P. (2007). To do or not to do: The neural signature of self-control. *Journal of Neuroscience, 27*, 9141–9145.

Breitmeyer, B. G. (1985). Problems with the psychophysics of intention. *Behavioral and Brain Sciences, 8*, 539–540.

Buckner, R. L., Andrews–Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1–38.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2*, 200–219.

Crick, F., and Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, *8*, 97–107.

Crick, F., and Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, *6*, 119–126.

Davidson, D. (1963). Actions, reasons, and causes. *The Journal of Philosophy*, *60*, 685–700.

Fried, I., Mukamel, R., and Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron*, *69*, 548–562.

Glynn, I. M. (1990). Consciousness and time. *Nature, 348*, 477–479.

Gomes, G. (1998). The timing of conscious experience: A critical review and reinterpretation of Libet's research. *Consciousness and Cognition, 7*, 559–595.

Gomes, G. (2002). The interpretation of Libet's results on the timing of conscious events: A commentary. *Consciousness and Cognition, 11*, 221–230 [Discussion, 308–325].

Haggard, P., and Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, *126*, 128–133.

Haynes, J. D. (2010). Beyond Libet: Long-term prediction of free choices from neuroimaging signals. In W. Sinnott–Armstrong and L. Nadel (Eds.), *Conscious will and responsibility: A tribute to Benjamin Libet* (pp. 85–96). New York: Oxford University Press. doi: 10.1093/acprof:oso/9780195381641.003.0009

Hung, C. P., Kreiman, G., Poggio, T., and DiCarlo, J. J. (2005). Fast readout of object identity from macaque inferior temporal cortex, *Science*, *310*, 863–866.

Keller, I., and Heckhausen, H. (1990). Readiness potentials preceding spontaneous motor acts: Voluntary vs. involuntary control. *Electroencephalography and Clinical Neurophysiology*, *76*, 351–361.

Kornhuber, H. H., and Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflüigers Archiv für die Gesamte Physiologie des Menschen und der Tiere*, *284*, 1–17.

Lau, H. (2009). Volition and the function of consciousness. In N. Murphy, G. Ellis, and T. O'Connor (Eds.), *Downward causation and the neurobiology of free will* (pp. 153–169). Chennai: Springer Verlag.

Lau, H. C., Rogers, R. D., Haggard, P., and Passingham, R. E. (2004). Attention to intention. *Science*, *303*, 1208–1210.

Lau, H. C., Rogers, R. D., and Passingham, R. E. (2006). On measuring the perceived onsets of spontaneous actions. *Journal of Neuroscience*, *26*, 7265–7271.

Levy, N. (2005). Libet's impossible demand. *Journal of Consciousness Studies*, *12*, 67–76.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*, 529–566.

Libet, B. (1989). Conscious subjective experience vs. unconscious mental functions: A theory of the cerebral processes involved. In R. M. J. Cotterill (Ed.), *Models of brain function* (pp. 35–50). New York: Cambridge University Press.

Libet, B. (1993). The neural time factor in conscious and unconscious events. In G. R. Block and J. Marsh (Eds.), *Experimental and theoretical studies of consciousness* (pp. 156–179). Chichester: John Wiley and Sons.

Libet, B. (1999). Do we have free will? *Journal of Consciousness Studies*, *6*, 47–57.

Libet, B. (2000). Time factors in conscious processes: Reply to Gilberto Gomes. *Consciousness and Cognition*, *9*, 1–12.

Libet, B. (2002). The timing of mental events: Libet's experimental findings and their implications. *Consciousness and Cognition*, *11*, 291–299.

Libet, B. (2004). *Mind time: The temporal factor in consciousness*. Cambridge, Massachusetts: Harvard University Press.

Libet, B., Gleason, C. A., Wright, E. W., Jr., and Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). *Brain*, *106*, 623–642.

Libet, B., Wright, E. W., Jr., and Gleason, C. A. (1982). Readiness-potentials preceding unrestricted 'spontaneous' vs. pre-planned voluntary acts. *Electroencephalography and Clinical Neurophysiology*, *54*, 322–335.

Locke, J. (1690). *An essay concerning human understanding* (first edition). London: Thomas Bassett.

Lycan, W. (1987). *Consciousness*. Cambridge, Massachusetts: MIT Press.

Lycan, W. (1996). *Consciousness and experience*. Cambridge, Massachusetts: MIT Press.

Matsuhashi, M., and Hallett, M. (2008). The timing of the conscious intention to move. *European Journal of Neuroscience*, *28*, 2344–2351.

Mele, A. R. (2009). *Effective intentions: The power of conscious will*. New York: Oxford University Press.

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83, 435–456.

Nahmias, E. (2011). Intuitions about free will, determinism, and bypassing. In R. Kane (Ed.), *The Oxford handbook of free will* (second edition, pp. 555–576). New York: Oxford University Press.

Nahmias, E., Morris, S., Nadelhoffer, T., and Turner 1, J. (2005). Surveying freedom: Folk intuitions about free will and moral responsibility. *Philosophical Psychology*, *18*, 561–584.

Nahmias, E., Morris, S. G., Nadelhoffer, T., and Turner, J. (2006). Is incompatibilism intuitive? *Philosophy and Phenomenological Research*, *73*, 28–53.

Nahmias, E., and Murray, D. (2010). Experimental philosophy on free will: An error theory for incompatibilist intuitions. In J. Aguilar, A. Buckareff, and K. Frankish (Eds.), *New waves in philosophy of action* (pp. 189–216). Basingstoke: Palgrave-Macmillan.

Norman, K.A., Polyn, S.M., Detre, G.J., and Haxby, J.V., (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*, 424–430.

Orne, M. T., and Whitehouse, W. G. (2000). Demand characteristics. In A. E. Kazdin (Ed.), *Encyclopedia of psychology* (pp. 469–470). Washington, D.C.: American Psychological Association.

Pockett, S., and Purdy, S. C. (2010). Are voluntary movements initiated preconsciously? The relationships between readiness potentials, urges, and decisions. In W. Sinnott–Armstrong and L. Nadel (Eds.), *Conscious will and responsibility: A tribute to Benjamin Libet* (pp. 34–46). New York: Oxford University Press. doi: 10.1093/acprof:oso/9780195381641.001.0001

Rosenthal, D. M. (2002). The timing of conscious states. *Consciousness and Cognition*, *11*, 215–220.

Rychlak, R. J., and Rychlak, J. F. (1997). Mental health experts on trial: Free will and determinism in the courtroom. *West Virginia Law Review*, *100*, 1–38.

Schultze–Kraft, M., Birman, D., Rusconi, M., Allefeld, C., Görgen, K., Dähne, S., Blankertz, B., and Haynes, J. D. (2016). The point of no return in vetoing self-initiated movements. *Proceedings of the National Academy of Sciences*, *U.S.A.*, *113*, 1080–1085.

Schurger, A., Sitt, J. D., and Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences*, *U.S.A.*, *109*, E2904–E2913.

Searle, J. (1992). *The rediscovery of the mind.* Cambridge, Massachusetts: MIT Press.

Searle, J. R. (1993). The problem of consciousness. *Consciousness and Cognition*, *2*, 310–319.

Shore, D. I., Spence, C., and Klein, R. M. (2001). Visual prior entry. *Psychological Science*, *12*, 205–212.

Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A., and Haggard, P. (2004). Altered awareness of voluntary action after damage to the parietal cortex. *Nature Neuroscience*, *7*, 80–84.

Smallwood, J., and Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*, 946–958.
Soon, C. S., Allefeld, C., Bogler, C., Heinzle, J., and Haynes, J. D. (2014). Predictive brain signals best predict upcoming and not previous choices. *Frontiers in Psychology*, *5*, 1–3. doi: 10.3389/fpsyg.2014.00406
Soon, C. S., Brass, M., Heinze, H. J., and Haynes, J. D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*, 543–545.
Soon, C. S., He, A. H., Bode, S., and Haynes, J. D. (2013). Predicting free choices for abstract intentions. *Proceedings of the National Academy of Sciences, U.S.A.*, *110*, 6217–6222.
Trevena, J. A., and Miller, J. (2002). Cortical movement preparation before and after a conscious decision to move. *Consciousness and Cognition*, *11*, 162–190.
Wegner, D. (2002). *The illusion of conscious will*. Cambridge, Massachusetts: MIT Press.
Wegner, D. M., and Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, *54*, 480–492.

BOOKS RECEIVED FOR REVIEW

Actual Causality. Joseph Y. Halpern. MIT Press, Cambridge, Massachusetts, 2016. $50.00 hard, 240 pages.

An Advanced Guide to Psychological Thinking: Critical and Historical Perspectives. Robert Ausch. Lexington Books, Lanham, Maryland, 2015. $90.00 hard, 257 pages.

Advances in Experimental Philosophy and Philosophical Methodology. Jennifer Nado (Editor). Bloomsbury, New York, 2016. $112.00 hard, 192 pages.

Advances in Religion, Cognitive Science, and Experimental Philosophy. Helen de Cruz and Ryan Nichols (Editors). Bloomsbury, New York, 2016. $112.00 hard, 232 pages.

The Ancient Origins of Consciousness: How the Brain Created Experience. Todd E. Feinberg and Jon M. Mallatt. MIT Press, Cambridge, Massachusetts, 2016. $35.00 hard, 392 pages.

Animal Electricity: How We Learned That the Body and Brain Are Electric Machines. Robert B. Campenot. Harvard University Press, Cambridge, Massachusetts, 2016. $39.95 hard, 352 pages.

Becoming Human: The Ontogenesis, Metaphysics, and Expression of Human Emotionality. Jennifer Greenwood. MIT Press, Cambridge, Massachusetts, 2016. $40.00 hard, 272 pages.

Beyond Physicalism. Edward Kelly, Adam Crabtree, and Paul Marshall (Editors). Rowan and Littlefield, Lanham, Maryland, 2015. $60.00 hard, 622 pages.

The Biological Basis of Mental Health (third edition). William T. Blows. Routledge, New York, 2016. $57.95 paper, 384 pages.

The Boundaries of Babel: The Brain and the Enigma of Impossible Languages (second edition). Andrea Moro. MIT Press, Cambridge, Massachusetts, 2016. $30.00 paper, 328 pages.

Celebrity and Entertainment Obsession: Understanding Our Addiction. Michael S. Levy. Rowan and Littlefield, Lanham, Maryland, 2015. $40.00 hard, 231 pages.

Cognitive Pluralism. Steven Horst. MIT Press, Cambridge, Massachusetts, 2016. $54.00 hard, 376 pages.

Cognitive Unconscious and Human Rationality. Laura Macchi, Maria Bagassi, and Riccardo Viale (Editors). MIT Press, Cambridge, Massachusetts, 2016. $54.00 hard, 401 pages.

Consciousness in the Physical World: Perspectives on Russellian Monism. Torin Alter and Yujin Nagasawa (Editors). Oxford University Press, New York, 2015. $74.00 hard, 462 pages.

Critical Flicker Fusion: Psychoanalysis at the Movies. William Fried. Karnac Books, Cornwall, United Kingdom, 2016. $39.95 paper, 160 pages.

Disclosing the World: On the Phenomenology of Language. Andrew Inkpin. MIT Press, Cambridge, Massachusetts, 2016. $43.00 hard, 400 pages.

Epistemic Friction: An Essay on Knowledge, Truth, and Logic. Gila Sher. Oxford University Press, New York, 2016. $50.00 hard, 320 pages.

Everyone Dies Young: Time Without Age. Marc Augé (translated by Jody Gladding). Columbia University Press, New York, 2016. $17.00 paper, 112 pages.

Felt Time: The Psychology of How We Perceive Time. Marc Wittmann. MIT Press, Cambridge, Massachusetts, 2016. $24.95 hard, 184 pages.

Free Will and Action Explanation: A Non-Causal, Compatibilist Account. Scott Sehon. Oxford University Press, New York, 2016. $74.00 hard, 248 pages.

Gender and Social Hierarchies: Perspectives from Social Psychology. Klea Faniko, Fabio Lorenzi–Cioldi, Oriane Sarrasin, and Eric Mayor (Editors). Routledge, New York, 2016. $52.95 paper, 193 pages.

The Given: Experience and its Content. Michelle Montague. Oxford University Press, New York, 2016. $60.00 hard, 272 pages.

Giving a Damn: Essays in Dialogue with John Haugeland. Zed Adams and Jacob Browning (Editors). MIT Press, Cambridge, Massachusetts, 2016. $50.00 hard, 384 pages.

Governing Behavior: How Nerve Cell Dictatorships and Democracies Control Everything We Do. Ari Berkowitz. Harvard University Press, Cambridge, Massachusetts, 2016. $29.95 hard, 240 pages.

Hallucination-focused Integrative Therapy: A Specific Treatment that Hits Auditory Verbal Hallucinations. Jack A. Jenner. Routledge, New York, 2016. $44.95 paper, 186 pages.

Hidden in Plain Sight: The Social Structure of Irrelevance. Eviatar Zerubavel. Oxford University Press, New York, 2015. $99.00 hard, $24.95 paper, 216 pages.

In Their Right Minds: The Lives and Shared Practices of Poetic Geniuses. Carole Brooks Platt. Imprint Academic, Exeter, United Kingdom, 2015. $29.95 paper, 201 pages.

Kantian Conceptual Geography. Nathaniel Jason Goldberg. Oxford University Press, New York, 2015. $74.00 hard, 271 pages.

Literature and the Metaphoric Universe in the Mind. Nicolae Babuts. Transaction Publishers, New Brunswick, New Jersey, 2015. $59.95 hard, 59 pages.

Microgenetic Theory and Process Thought. Jason W. Brown. Imprint Academic. Exeter, United Kingdom, 2015. $39.00 paper, 212 pages.

The Multiple Realization Book. Thomas W. Polger and Lawrence A. Shapiro. Oxford University Press, New York, 2016. $85.00 hard, $35.00 paper, 240 pages.

A Natural History of Human Morality. Michael Thomasello. Harvard University Press, Cambridge, Massachusetts, 2016. $35.00 hard, 194 pages.

Neural Control of Speech. Frank H. Guenther. MIT Press, Cambridge, Massachusetts, 2016. $63.00 hard, 424 pages.

On Multiple Selves. David Lester. Transaction Publishers, New Brunswick, New Jersey, 2015. $49.95 hard, 198 pages.

The Oxford Handbook of Cognitive and Behavioral Therapies. Christine Maguth Nezu and Arthur M. Nezu (Editors). Oxford University Press, Oxford, United Kingdom, 2016. $175.00 hard, 511 pages.

Panpsychism: Contemporary Perspectives. Godehard Brüntrup and Ludwig Jaskolla (Editors). Oxford University Press, New York, 2016. $74.00 hard, 424 pages.

The Pragmatic Turn: Toward Acrion-Oriented Views in Cognitive Science. Andreas K. Engel, Karl J. Friston, and Danica Kragic (Editors). MIT Press, Cambridge, Massachusetts, 2016. $49.00 hard, 432 pages.

The Psychology of Radicalization and Terrorism. Willem Koomen and Joop Van Der Pligt. Routledge, New York, 2016. $52.95 paper, 284 pages.

Rational Belief: Structure, Grounds, and Intellectual Virtue. Robert Audi. Oxford University Press, New York, 2015. $29.95 paper, 281 pages.

The Rationalty Quotient: Toward a Test of Rational Thinking. Keith E. Stanovich, Richard F. West, and Maggie E. Toplak. MIT Press, Cambridge, Massachusetts, 2016. $39.00 hard, 480 pages.

Reading Italian Psychoanalysis. Franco Borgogno, Alberto Luchetti, and Luisa Marina Coe (Editors). Routledge, New York, 2016. $295.00 hard, 738 pages.

Reason in the World: Hegel's Metaphysics and Its Philosophical Appeal. James Kreines. Oxford University Press, New York, 2015. $74.00 hard, 290 pages.

Reasons Why. Bradford Skow. Oxford University Press, New York, 2016. $70.00 hard, 224 pages.

Recovering from Psychosis: Empirical Evidence and Lived Experience. Stephen Williams. Routledge, New York, 2016. $34.95 paper, 132 pages.

The Research Companion: A Practical Guide for Those in the Social Sciences, Health and Development. Stephen Williams. Routledge, New York, 2016. $52.95 paper, 289 pages.

Reconstructing Strategy: Dancing with the Objectivity God. Saqib Qureshi. Two Harbors Press, Minneapolis, Minnesota, 2015. $29.99 hard, 440 pages.

Rethinking Therapeutic Culture. Timothy Aubry and Trysh Travis. University of Chicago Press, Chicago, 2015. $30.00 paper, 267 pages.

Scurvy: Disease of Discovery. Jonathan Lamb. Princeton University Press, Princeton, New Jersey, 2016. $35.00 hard, 328 pages.

Tackling Poverty and Social Exclusion: Promoting Social Justice in Social Work (third edition). John H. Pierson. Routledge, New York, 2016. $50.95 paper, 231 pages.

Tense Bees and Shell-Shocked Crabs: Are Animals Conscious? Michael Tye. Oxford University Press, New York, 2016. $29.95 hard, 231 pages.

Thought in Action: Expertise and the Conscious Mind. Barbara Gail Montero. Oxford University Press, New York, 2016. $65.00 hard, 288 pages.

Towards Non-Being: The Logic and Metaphysics of Intentionality (second edition). Graham Priest. Oxford University Press, New York, 2016. $45.00 paper, 384 pages.

Unclaimed Experience: Trauma, Narrative, and History. Cathy Carruth. Johns Hopkins Press, Baltimore, Maryland, 2016. $29.95 paper, 195 pages.

Using Hypnosis with Children: Creating and Delivering Effective Interventions. Lynn Lyons. W. W. Norton and Company, New York, 2015. $35.00 paper, 386 pages.

Voice Leading: The Science Behind a Musical Art. David Huron. MIT Press, Cambridge, Massachusetts, 2016. $38.00 hard, 272 pages.

What's Wrong with Morality? A Social-Psychological Perspective. C. Daniel Batson. Oxford University Press, New York, 2015. $98.50 hard, $39.95 paper, 265 pages.

AUTHOR INDEX VOLUME 37

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 price. 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: steven.connelly@indstate.edu

JMB is abstracted or indexed in: *Cultures, Langues, Textes: La Revue de Sommaires; Current Contents (Social and Behavioral Sciences); EBSCOhost; EMBASE/Excerpta Medica; International Bibliography of Book Reviews; International Bibliography of Periodical Literature; Linguistics and Language Behavior Abstracts; Physics Abstracts; ProQuest; 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

Summer and Autumn 2016 Vol. 37 Nos. 3 and 4

CONTENTS